


Multifractal analysis of birdsong and its correlation structureRabindev Bishal,¹ Gabriel B. Mindlin^{2,3} and Neelima Gupte^{1,3}¹*Department of Physics, IIT Madras, Chennai 600036, India*²*Department of Physics, University of Buenos Aires, 1428 Buenos Aires, Argentina*³*Complex Systems and Dynamics Group, IIT Madras, Chennai 600036, India* (Received 16 April 2021; revised 17 November 2021; accepted 20 December 2021; published 18 January 2022)

The time series recordings of typical songs of songbirds exhibit highly complex and structured behavior, which is characteristic of their species and stage of development, and need to be analyzed by methods that can uncover their correlation structure. Here we analyze a typical song of a canary using Hurst exponents and multifractal analysis, which uncovers the correlation structure of typical song segments. These are then compared with the corresponding quantities from shuffled data, which destroys the temporal correlations and iterative amplitude-adjusted Fourier transform (IAAFT) data. It is seen that temporal correlations are responsible for the multifractal behavior seen in the data and that two-point correlations, which are preserved by the transform, are important in the high-fluctuation regime. Higher-order correlations and intersyllabic gaps dominate the behavior of the low-fluctuation regime. These observations are supported by the simplicial characterization of the corresponding time series networks. Complexity measures are also used to analyze the amplitude envelope time series. These indicate that intersyllabic gaps contribute a significant fraction to the complexity of the birdsong. Our method provides a detailed characterization of the data, which can enable the comparison of real and synthetic birdsong and comparisons across stages of development and species. A brief comparison with the song of the zebra finch supports this.

DOI: [10.1103/PhysRevE.105.014118](https://doi.org/10.1103/PhysRevE.105.014118)**I. INTRODUCTION**

The activity of birdsong constitutes an important example of a process in which a neurophysiological process results in complex audible output [1]. The study of the characteristics of the birdsong, and the identification of a system that can produce a synthetic birdsong of the same characteristics, provides important insights in the ways in which the neural architecture in the brain can coordinate with a delicate vocal apparatus that the birds can and must control with high precision [1–5]. It is interesting to note that in the case of many species, birdsong shares with human speech the feature that the acquisition of vocalization matures with age and also requires exposure to a tutor. The tools required for this analysis draw from multidisciplinary areas such as neuroscience, dynamical systems theory, time series analysis, and, recently, machine learning techniques. Different signal processing and time series analysis tools have been used to understand and characterize its spectral features such as time-frequency responses, spectrograms [6], syllabic sequences [7,8], the rhythmic structure of intensity, pitch, and onset timing of notes [9]. Multifractal detrended fluctuation analysis has been used to study the predictable and unpredictable patterns and the fluctuation of the amplitude envelope, pitch, and intensity of the thrush nightingale's song [9].

We note that multifractal detrended fluctuation analysis has been the technique of choice in analyzing nonstationary time series and has been applied in diverse contexts such as analyzing long-time weather records, neuron spiking, heart

rate dynamics, and others. The technique can identify the time correlations in the data but avoids spurious correlations. It can also identify crossover timescales with separate regimes with distinct scaling exponents, e.g., long-range correlations on small scales and other types of correlations or uncorrelated behavior on large scales. The scaling behavior in real data can be even more complex, and different parts of the time series can require different exponents. In some examples, different types of scaling behavior can be observed for many interwoven fractal subsets and require a full range of scaling exponents to characterize their behavior. This kind of detailed description is of great utility in comparing real data with surrogate data, synthetic data from models, or in applications that require detailed tracking of variations. For example, age-related and disease-related changes in the scale-invariant structure of heart rate variability are indicated by changes in the multifractal spectrum. In the case of endogenous brain dynamics and response times, the multifractal spectrum has also become sensitive to cognitive performance. Thus, this analysis can provide important information which can complement information obtained from other techniques.

Here we analyze a typical song of a canary using Hurst exponents and detrended multifractal analysis, which uncovers the correlation structure of typical segments of the complete song. The spectrum of Hurst exponents obtained for individual segments is compared with the spectrum obtained for two types of surrogate data generated from the original data, one generated by shuffling the original data from the time series, and the other generated by the iterative amplitude-adjusted

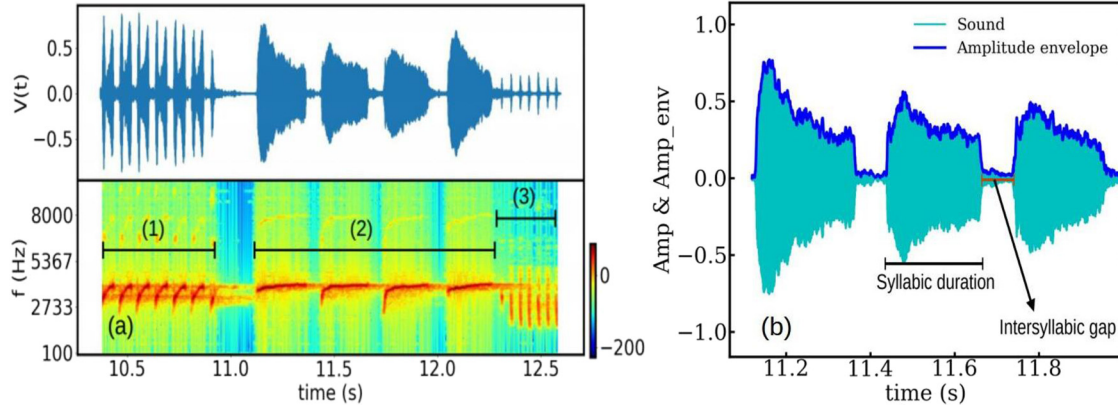


FIG. 1. (a) A typical example of the song of a canary along with the corresponding sonogram. Three types of syllables are present from left to right: an up-sweep (1), a small up-sweep followed by the tonal (2), and a down-sweep (3). (b) The amplitude envelope showing the syllable and the intersyllabic gap (pause) for a canary song segment.

Fourier transform (IAAFT) [10,11] which preserves the power spectrum and the distribution of the data. The temporally shuffled data results in a flat monofractal spectrum for the Hurst exponents, indicating the importance of the temporal correlations. The spectrum and scaling properties for the surrogate data generated by the IAAFT match well with the original data in the high-fluctuation regime but show significant deviations in the low-fluctuation regime. Thus, preserving the power spectrum is sufficient to preserve the characteristics of the time series in the high-fluctuation regime. However, higher-order correlations and the contribution of the intersyllabic gaps become important in the low-fluctuation regime. The simplicial characterization of the time series network constructed from the energy data supports this observation.

We also use different complexity measures to analyze the time series segments. These complexity measures quantify the hierarchical development of complexity and a notion of its time evolution. The analysis set up here paves the way to the comparison of actual birdsong data with simple nonlinear dynamical models [12,13] which can synthesize birdsong, as well as the detailed quantitative comparison of birdsong across species and different stages of development. This is supported by a preliminary analysis of a zebra finch song.

II. METHODS

We outline the methods used to analyze the typical time series obtained from a birdsong. In terms of a sound wave, birdsong is a traveling wave created due to pressure fluctuation originated by the oscillation of labia in the bird's vocal organ, i.e., the syrinx. The typical structure of a birdsong consists of brief vocalizations separated by pauses [Figs. 1(a) and 1(b)]. These brief vocalizations are called syllables [1]. A bird can produce many such brief vocalizations per second (more than 30/s for the canary) with intersyllabic gaps in between. We observe rich spectral and temporal structure even in the small duration of a single vocalization in the sonogram. The temporal evolution of the frequencies defines the richness of the spectral properties of the syllables. The frequencies of these syllables have different kinds of temporal evolution. These evolutions can be of the tonal type (corresponding to a single frequency), harmonics, up-sweeps (with increasing

frequency), down-sweeps (with decreasing frequency), and many other complex frequency modulations. Some of these behaviors can be seen in the signal of the birdsong seen in Fig. 1(a). On the other hand, the amplitude envelope of the birdsong encodes the temporal evolution of the amplitude, rhythmic structures, the duration of the intersyllabic gaps, the onset of syllables, syllabic duration, and many other phenomena. This temporal evolution often shows the fluctuation of the predictable and unpredictable pattern of the song. All these properties vary with different birds. Some birds have very few syllabic patterns, such as the koel (*Eudynamis*), coppersmith barbet (*Megalaima haemacephala*), and many others. Some birds have many different syllabic repetitions, as seen in the songs of lyrebirds (*Menura*), the common nightingale (*Luscinia megarhynchos*), and many others [14]. Here we study the temporal fluctuation and scaling properties of the amplitude envelope of the birdsong of the canary (*Serinus canaria domestica*).

The sound file analyzed here is the song of an adult canary [15]. The age of the bird and the sex of the bird are unknown. The audio file has a sampling rate of 44100 Hz. The song's duration is 16 s, including pauses between the phrases. The audio data contain an isolated syllable (0.78–1.23 s) of very short duration. We have discarded this isolated syllable as it is too short for significant scaling. We have divided the time series by the maximum of the absolute signal so that the signal lies between -1 and 1 . A typical time series for the birdsong is shown in Fig. 1(a). The x -axis refers to the time (s), and the y -axis is the voltage. We use a microphone that maps the pressure change into a voltage signal. The corresponding sonogram is also shown in Fig. 1(a). Figure 1(b) shows the amplitude envelope, identifying the parts which correspond to the amplitude envelope and the intersyllabic gap. Three types of syllables can be seen in Fig. 1(a), viz., an up-sweep, the tonal, and a down-sweep. In the following subsection, we first introduce a general method to determine the amplitude envelope of any given song. Then we discuss the methods and characteristics used to quantify the dynamic patterns of predictability and unpredictability of the amplitude envelope. These include the multifractal detrended fluctuation formalism and a complexity measure. We include a brief discussion of these quantities in this section.

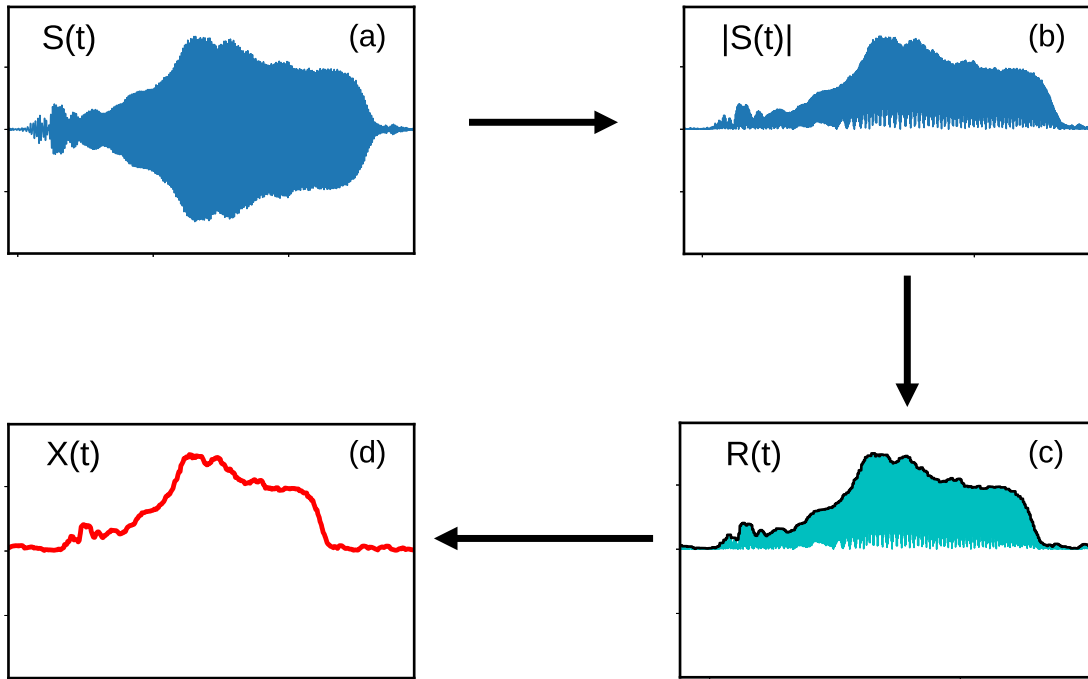


FIG. 2. (a) A typical syllable of the canary song. (b) The absolute value of the time series. (c) The signal after finding the maximum value of the absolute time series. This is identified as $R(t)$. In the case of this analysis, we have considered a window of 10 sample units shifted by half the window size. (d) The signal is then filtered with a low-pass filter with a cutoff frequency of 300 Hz.

A. Amplitude envelope determination

To obtain an amplitude envelope output without any attenuation, we use an approach that combines a peak detection algorithm with a window [16] which shifts. The details of the methods are as follows (Fig. 2). First, for a given time series $S(t)$, we compute the absolute value of the signal $|S(t)|$. Then we divide the absolute value of the signal ($|S(t)|$) into moving windows of fixed width which shift by a certain fraction of the window size. Typically the window is shifted by 0.25 to 0.5 times the window’s width. For each window, we replace all the values with the maximum value of the signal seen in the window. Then we use a low-pass filter with a cutoff frequency to smooth the signal. We have used the (scipy.signal.butter) package from Scipy (Scientific python) library for filtering process. Figure 2(d) shows a prototypical example of the amplitude envelope time series. The scaling property of this time series can be studied using the detrended fluctuation analysis technique defined in the following subsection.

B. Multifractal detrended fluctuation formalism

Real-world time series are, in general, inhomogeneous in their temporal fluctuation of the magnitude. These time series are analyzed using the scaling property of the temporal fluctuations. There are several methods of analysis, such as fluctuation analysis [17–20], wavelet-based approaches [21,22], power spectrum techniques, etc. Among these methods, a frequently used method of analysis is the method of detrended fluctuation analysis (DFA)[18]. This DFA is based on the standard random-walk diffusion analysis that estimates how the standard deviation of the variable grows as a function of the timescale with an exponent H [18]. A signal having no long-range structure, like white noise, has a standard deviation

with an exponent H which does not change with the change in timescale. On the other hand, a signal containing long-range structure (like the waveform of birdsong or speech) has a well-structured fluctuation across different timescales. As a result of this, the standard deviation and its scaling behavior will not be uniform across all timescales. Here different parts of the time series follow different scaling relations with the timescale. In particular, the magnitude of the fluctuations of such time series scales differently with time. A single scaling exponent can not provide complete information on these scaling properties. Therefore, one needs a wide range of scaling exponents to analyze the full features of the time series fluctuations, i.e., these time series are multifractal in nature. In this section, we discuss the framework [17] required to extract the multifractal features from a time series. The multifractal DFA (MF-DFA) is a generalized version of detrended fluctuation analysis [18] and consists of four steps, as follows:

1. Suppose that x_j is a series of length N , with $x_j = 0$ for an insignificant fraction of the series. We find the profile of the time series defined to be the cumulative sum $\{Y(n)\}$ of the data after the subtraction of the mean from each data point:

$$Y(n) = \sum_{j=1}^n [x_j - \langle x \rangle] \quad n = 1, \dots, N. \quad (1)$$

Here $\langle x \rangle = \frac{1}{N}(\sum_{j=1}^N x_j)$ is the average value of the series. The subtraction of $\langle x \rangle$ from each of the data values is not essential as in further steps the cumulative sum of the time series would be detrended by suitable polynomials.

2. Next we divide the profile $Y(n)$ into $k_n = \text{int}(\frac{N}{n})$ nonoverlapping segments of equal lengths n . Since the length (N) of the series is not often a multiple of the timescale n considered, a short part at the end of the profile may remain.

In order to not disregard this part of the series, we repeat the same procedure from the opposite end. Therefore, the total number of segments is $2k_n$ segments.

3. We then find the local trend for each of the $2k_n$ segments by a least-square fit of the profile and determine the variance for $v = 1, \dots, k_n$:

$$F^2(n, v) = \frac{1}{n} \sum_{i=1}^n \{Y[(v-1)n+i] - y_v(i)\}^2. \quad (2)$$

For $v = k_n + 1, \dots, 2k_n$,

$$F^2(n, v) = \frac{1}{n} \sum_{i=1}^n \{Y[N - (v - k_n)n + i] - y_v(i)\}^2. \quad (3)$$

4. Next, we find the average over all the segments to obtain the q th-order fluctuation function defined as

$$F_q(n) = \left\{ \frac{1}{2k_n} \sum_{v=1}^{2k_n} [F^2(n, v)]^{\frac{q}{2}} \right\}^{\frac{1}{q}}. \quad (4)$$

We determine the scaling behavior of the fluctuation functions by analyzing log-log plots of $F_q(n)$ vs timescale n for each value of q . If the series x_i are long-range power-law correlated, $F_q(n)$ increases, for the large values of n , as a power law,

$$F_q(n) \sim n^{h_q}. \quad (5)$$

The quantity h_q is known as the generalized Hurst exponent. For a monofractal time series, h_q is independent of q , since the scaling behavior of the variances $F^2(n, v)$ is identical for all segments v . The averaging procedure in Eq. (4) will give this identical scaling behavior for all values of q . For a positive value of q , the segments having larger values of fluctuations (i.e., larger deviation from the fitting polynomial) will dominate the sum of $F^2(n, v)$ [Eq. (4)]. Thus, for positive values of q , h_q describes the scaling behavior of segments with large fluctuations. For negative values of q , the segments with small fluctuations (i.e. small deviations from the fitting polynomial) will dominate the $F^2(n, v)$. Thus, for negative values of q , h_q describes the scaling behavior of segments with small fluctuations.

The exponent β for the power spectral density [18,19] satisfies the relation $\beta = 2h_2 - 1$. For a long-range correlated signal, the exponent $h_2 > 0.5$, and for a long-range anticorrelated signal, the exponent $h_2 < 0.5$. The exponent for the random walk is $h_2 = 0.5$. The exponent for brown noise has the value $h_2 = 1.5$, and the exponent for pink noise is $h_2 = 1$.

In the case of stationary time series, $h_2 = H$, where H is the Hurst exponent [17,19]. Thus, the power spectral exponent β is related to H via the relation $\beta = 2H - 1$. For nonstationary signals, the Hurst exponent $H = h_2 - 1$. Hence, the exponent β for the power spectral density satisfies the relation $\beta = 2H + 1 = 2h_2 - 1$.

It is also pertinent to mention some of the limitations of the multifractal detrended fluctuation analysis. Since the MFDFA uses local fits to polynomials, the order of the polynomial is important, and a single form or power of the polynomial may not be the appropriate fit for all segments of the time series. Some improvements on the MFDFA method address this aspect [23]. However, the removal of local trends in

MFDFA using discontinuous polynomial fitting can result in pseudofluctuation errors [24].

Additionally, some processes which involve long-range correlations, such as processes based on fractional Gaussian noise, may show spurious exponents in the MFDFA [25]. However, recent methods which have successfully analyzed a fractional Ornstein-Uhlenbeck process and a Lévy distribution provide a clue to the extraction of reliable results in long-range correlated cases [26]. Some examples of long-range monofractal data can show spurious multifractality in the presence of white noise, short-range correlations, and periodic external trends. The effect of these factors on real-world and synthetic monofractal and multifractal time series data has been discussed by Ludescher *et al.* [27].

C. Simplicial analysis

Time series (TS) networks, *viz.*, the networks constructed out of time series data, can be constructed using a variety of algorithms [28]. These TS networks can be analyzed using the concepts of cliques, and simplicial complexes [29]. Here a clique is defined as a complete subgraph, and a simplex is a set of connected nodes. Isolated points are zero simplices; two connected nodes form a one- simplex, and three connected nodes (i.e., a triangle) constitute a two-simplex. If two simplices have $q + 1$ nodes in common, they share a $q -$ face. A q -connected simplicial structure is a sequence of simplices such that each simplex shares a q -face with the successive simplex. Simplices that are q -connected are also connected at all the lower levels. The dimension of a simplicial complex is the dimension of the largest simplex in the structure. The number of q -dimensional simplices forms the q th component of the f -vector, and the $\vec{f} = \{f_0, f_1, \dots, f_{q_{\max}}\}$, where q_{\max} is the dimension of the highest dimensional simplices in the network. The simplicial analysis of several kinds of networks, such as social networks and networks of neurons, has been carried out. In the case of time series networks, the simplicial structure of the TS network reflects the temporal correlations in the time series data.

Many other quantities can also be devised to analyze stationary and nonstationary time series. The quantity used in this paper, *viz.*, the complexity, is defined in the following subsection.

D. Complexity measure

The complexity measure is often associated with the disorder present in the system. It measures the probabilistic hierarchy of a system ranging from minimum to maximum information with a perfectly ordered system, e.g., a perfect crystal, requiring the minimum information for its specification and a completely disordered system, e.g., an ideal gas, where every accessible state is equally probable, requiring the maximum information [30]. Information measures and disequilibrium measures are sometimes defined by defining some distance measure which measures how far the actual states of the system are from the equiprobable states. These can be used to measure the complexity of the system.

For a system with η accessible states, if all the states have the same probability, i.e., $p_i = \frac{1}{\eta}$, the system is equiprobable.

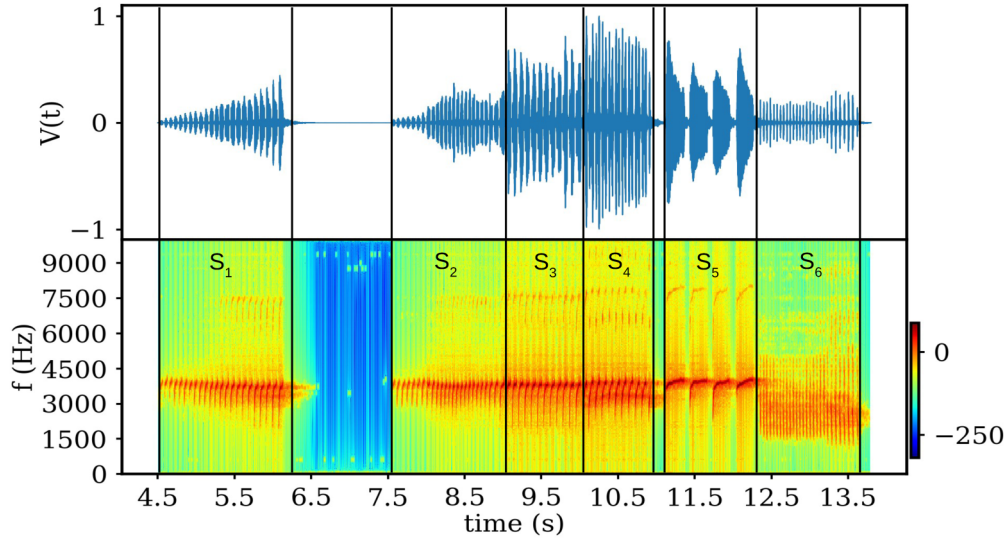


FIG. 3. The canary birdsong along with its sonogram. Segments are defined based on the syllabic patterns in the sonogram and the shape of the amplitude envelope. We show such segments for an adult canary song.

For a general system, the states are not necessarily equiprobable, so in general, $p_i \neq \frac{1}{\eta}$. For any set of p_i , the normalized information entropy is defined as

$$S = -\frac{\sum_{i=1}^{\eta} p_i \log p_i}{\log \eta}, \quad (6)$$

and the disequilibrium is defined as

$$D = \sum_{i=1}^{\eta} \left(p_i - \frac{1}{\eta} \right)^2. \quad (7)$$

Additionally, the López-Ruiz, Mancini, and Calbet (LMC) [30] complexity measure of the system is defined as $C = S \times D$. We will apply these measures to the amplitude envelope of the birdsong data. The details of the data collection have been discussed at the beginning of this section. We outline below the features of the data that have been collected.

III. BIRDSONG DATA: FEATURES AND ANALYSIS FOR THE SONG OF THE CANARY

Figure 3 shows the audio time series of the birdsong of the canary. We have analyzed the song segment-wise. In the following subsection, we discuss the segmentation process of the audio time series, which is carried out as follows.

As shown in Fig. 3, the audio song consists of many repetitive patterns of syllables. The collection of syllables of a similar type makes a phrase. Here this particular song also has many phrases. We find the song's spectrogram, and then, based on the sonogram and the amplitude variation, we segment the song. We also show the different song segments in Fig. 3. The amplitude envelope of a musical signal encodes the rhythmic structure, the amplitude modulation, the duration of notes, the loudness variation, and many other features. These features encode the expressiveness of music. Like music, some birds, particularly songbirds like canaries or nightingales, also produce music-like signals, which have been analyzed using quantifiers like the multifractal spectrum

of the amplitude envelope and its width [9]. Here we use the Hurst exponent to analyze how the amplitude envelope scales with time. We focus on different aspects like nonlinearity, the role of intersyllabic gaps, and the effect of these entities on the signal's overall multifractality. We note that the amplitude envelopes of the speech signals are often nonlinear. We have used surrogate data analysis to understand the effect of nonlinearity on the multifractal property of the amplitude envelope.

A. Methods of creating surrogate data

A basic approach to creating surrogate data with the same scaling law as the original data is to perform a Fourier transform on the time series, preserve the Fourier amplitudes, randomize the Fourier phases, and perform an inverse Fourier transform to create the surrogate series. This method eliminates nonlinearities stored in the Fourier phases, preserving the power spectrum and the autocorrelation (second-order moments) of the original time series [11]. The surrogate data generated in this method do not preserve the probability distribution of the time series and may lead to an erroneous conclusion regarding the nonlinearity of the underlying process. An improved method is the iterative amplitude-adjusted Fourier transform (IAAFT) [10,11] used here to preserve the power spectrum and the distribution of the data. The process of generating surrogate (IAAFT) data and the shuffled data is briefly summarized below, and further details can be found in Refs. [31,32]. See Ref. [33] for the limitations of the IAAFT procedure and improvement on the algorithms.

The algorithm is iterative in nature and adjusts the Fourier spectrum first and then the amplitudes. The algorithm starts with sorting the original time series by the amplitude. Next, the Fourier transform is carried out on the original data set. At each iteration, the Fourier transform of the iterated time series is calculated, and the Fourier coefficients of the original time series replace its coefficients to produce the desired power spectrum. The phases (Fourier phases of the transformed data) are kept unaltered. After this step, the amplitudes of the

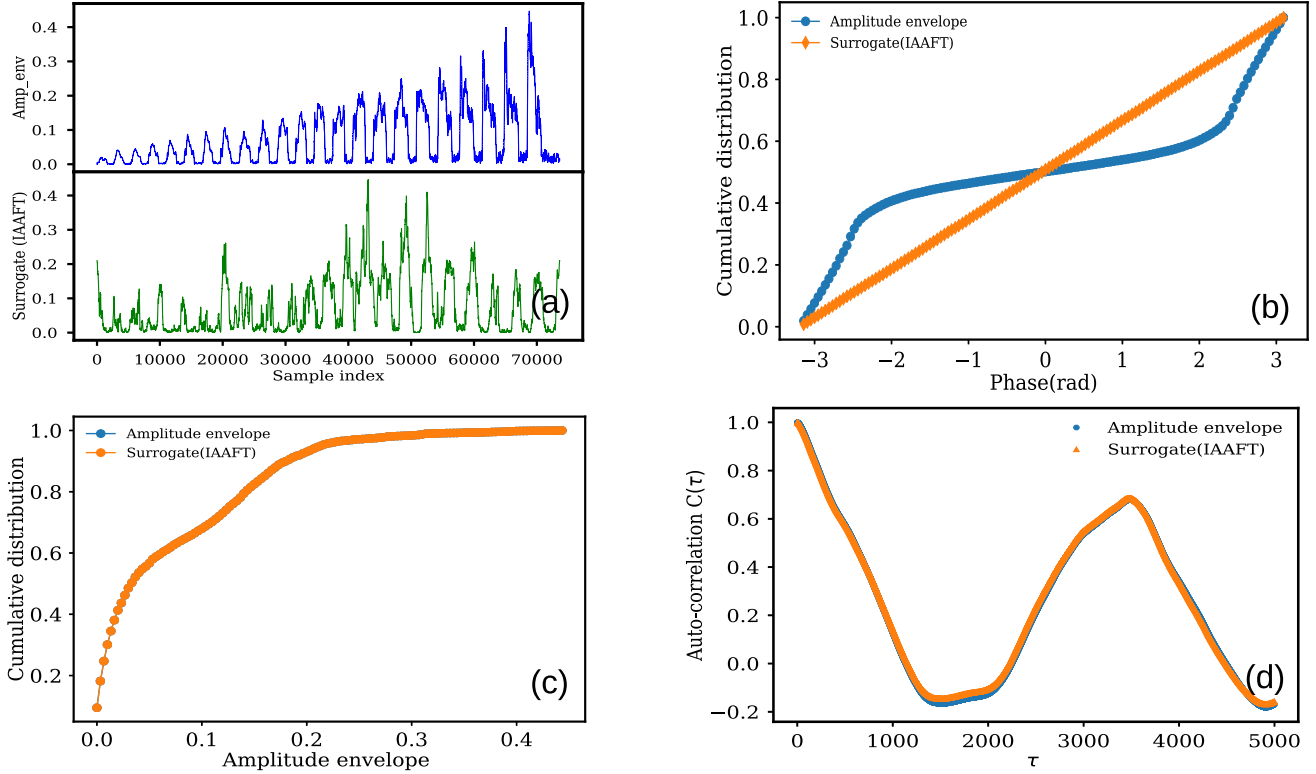


FIG. 4. (a) The amplitude envelope (blue online) and its corresponding surrogate (IAAFT) (green online) for segment 1. (b) The cumulative distribution of the Fourier phases of the original amplitude envelope and corresponding surrogate (IAAFT). It shows the presence of nonuniform nature and correlations in the Fourier phases of the original signal. The surrogate (IAAFT) data have uniformly distributed and uncorrelated Fourier phases. (c) The cumulative distribution of the original amplitude envelope and its corresponding surrogate (IAAFT) for segment 1. It shows identical behavior. (d) The autocorrelation of the original amplitude envelope and its corresponding surrogate.

iterated time series will no longer be the same. Therefore, in the second iterative step, the amplitudes are adjusted by ranking their values and replacing them with the values of the amplitudes at the same rank in the original sorted list. Both iterative steps are repeated until a convergence threshold is reached. In this report, to generate a single surrogate data set for each segment of the canary song, we have used 1000 iterations, and the power spectrum threshold is set at 10^{-8} .

In Fig. 4(a) we have plotted the amplitude envelope of segment 1 as a time series, and its corresponding surrogate data created by the iterated amplitude-adjusted Fourier transform method. The nonuniform distribution of the Fourier phases indicates the presence of nonlinearity in the original amplitude envelope in Fig. 4(b). Figures 4(c) and 4(d) demonstrate the cumulative distributions and autocorrelations of the original amplitude envelope and its corresponding surrogate respectively of segment 1.

We also shuffle the original data to generate a second set of surrogate data to detect the effect of the temporal correlation of the amplitude envelope on the multifractal spectrum. For that, first, we store the original data. Next, we swap each value of the time series chosen randomly from anywhere in the data for each step. On average, the swapping of each data point more than twice makes it a random time series.

For each segment, we have generated two sets of surrogate data with 100 realizations of each set. The first set of surrogate data is generated using the iterative amplitude-adjusted

Fourier randomized method (IAAFT). The IAAFT surrogate data preserve the original data's probability distribution and power spectrum. The second set of surrogate data is the temporally shuffled data. The shuffled data preserve the amplitude envelope's probability distribution but destroy any temporal correlations present in the original data. The IAAFT surrogate data and temporally shuffled data for each segment are used to understand the effect of nonlinear properties and the effect of temporal correlation present in the original time series in terms of the multifractal spectrum. For each segment, the average multifractal spectrum of the surrogate (IAAFT) and the average multifractal spectrum of the shuffled data are compared with the multifractal spectrum of the original amplitude envelope. In the following subsection (III B), we describe the segmentation of the audio time series.

B. Segmentation of the audio data

The audio file of the birdsong consists of many repetitive patterns of syllables. The collection of these syllables of similar types makes a phrase. The audio file of the birdsong considered here also has many phrases. We find the spectrogram of the song, and then, based on the sonogram and the amplitude variation, we segment the song. In Fig. 3, we show different song segments. Table I identifies the different segments, their syllabic nature, and average durations of the

TABLE I. The first column is the segment label, and the second column contains the start and end time of the segments. The third column identifies the syllabic type of the segments, and the fourth column lists the average duration of the syllables present in the segment.

Segment	Start and end time(s)	Syllabic nature	Average syllabic duration(ms)
1	4.54–6.23	Up-sweep followed by tonal	46.88 ± 7.02
2	7.54–9.03	Up-sweep followed by tonal	35.29 ± 6.04
3	9.03–10.03	Up-sweep	63.63 ± 4.04
4	10.03–10.92	Circular up-sweep	63.55 ± 1.83
5	11.12–12.27	Tonal	244.74 ± 2.66
6	12.35–13.66	Down-sweep	18.31 ± 2.38

syllables. We note that the amplitude envelope data can be divided into six segments, numbered 1 to 6.

The data are analyzed as follows. We first obtain the generalized Hurst spectrum in the presence and absence of intersyllabic gaps for the amplitude envelope data and their surrogate. Then we discuss how the complexity changes over the different segments in the presence and absence of intersyllabic gaps. We note here that segments 1, 2, and 3 show behavior that is qualitatively similar, segment 5 shows distinct behavior, and segments 4 and 6 show qualitatively similar behavior in the low-fluctuation regime (particularly with intersyllabic gaps). Hence we discuss the behavior of

segments 1, 5, and 6 in the subsequent subsections. The behavior and qualitative resemblance of segments 2 and 3 to segment 1 and segment 4 and its comparison with segment 6 are analyzed in Appendices A 1, A 2, and A 3, respectively.

C. Segment 1

The sonogram and the amplitude envelope of the segment are shown in Figs. 5(a) and 5(b). Figures 5(c) and 5(d) show the detrended fluctuation of the amplitude envelope profile in the presence and absence of intersyllabic gaps. Crossover effects can be seen in both the cases. Finally, the generalized Hurst exponent h_q is plotted as a function of q for both cases in Figs. 5(e) and 5(f).

A clear crossover can be seen in the detrended fluctuation of the amplitude envelope detrended via a first-order polynomial both in the presence and the absence of intersyllabic gaps. The scaling exponent takes the value $\alpha_1 = 1.68$ for scales less than the crossover timescale ($n = 2420$), indicating a strong long-range correlation detrended fluctuation of the amplitude envelope profile over these scales. For large timescales, i.e., after the crossover timescale, we see the signatures of anticorrelated behavior [Fig. 5(c)]. The intersyllabic gaps have significant effects on the multifractal spectrum and the crossover point of the segment. In the absence of intersyllabic gaps, the nature of the detrended dynamics of the fluctuations does not change, but the crossover occurs at a much lower timescale ($n \approx 1000$) [Fig. 5(d)] For the generalized multifractal spectrum, the scaling features for the large fluctuation region ($q > 0$) of amplitude envelope and the corresponding surrogate data are significantly similar

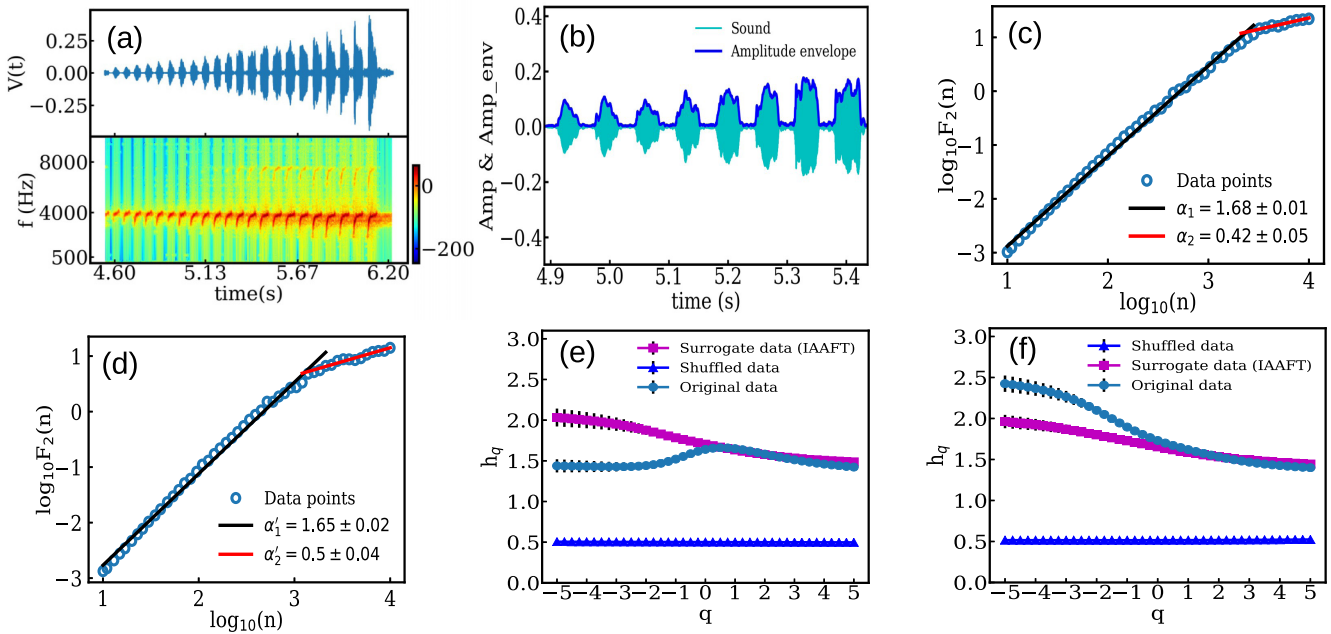


FIG. 5. (a) The sonogram of segment 1. (b) The amplitude envelope along with the song segment. (c) The crossover in the detrended fluctuation of the amplitude envelope profile in the presence of the intersyllabic gaps. The slope h_2 in the first scaling region (timescales smaller than the crossover point) is denoted by α_1 . The slope h_2 in the second scaling region (timescales greater than the crossover point) is denoted by α_2 . (d) The crossover in the detrended fluctuation of the amplitude envelope profile in the absence of intersyllabic gaps. (e) The h_q vs q spectrum of the amplitude envelope of the song segment with intersyllabic gaps. (f) The h_q vs q of the amplitude envelope of the song segment without intersyllabic gaps.

TABLE II. The Kolmogorov-Smirnov (KS) statistics of the generalized Hurst exponent for the original amplitude envelope and the surrogate data (IAAFT) in the presence and absence of the intersyllabic gaps. The columns correspond to the KS statistics ($D_{ks}(q < 0)$) for low-fluctuation region, corresponding probability $p(q < 0)$ of accepting the null hypothesis, KS statistics ($D_{ks}(q > 0)$) for high-fluctuation region, and its corresponding probability $p(q > 0)$, respectively.

Segment	With intersyllabic gap				Without intersyllabic gap			
	$D_{ks}(q < 0)$	$p(q < 0)$	$D_{ks}(q > 0)$	$p(q > 0)$	$D_{ks}(q < 0)$	$p(q < 0)$	$D_{ks}(q > 0)$	$p(q > 0)$
1	1.00	0.00	0.35	0.17	0.71	0.00	0.35	0.17
2	0.90	0.00	0.40	0.08	0.81	0.00	0.40	0.08
3	1.00	0.00	0.25	0.57	0.33	0.20	0.30	0.34
4	1.00	0.00	0.20	0.83	1.00	0.00	0.10	1.00
5	1.00	0.00	0.45	0.03	0.48	0.02	0.50	0.01
6	0.95	0.00	0.10	1.00	0.43	0.04	0.15	0.98

(Table II). The behavior of the h_q vs q for the shuffled data shows that the probability distribution does not affect the multifractal spectrum. It suggests that the multifractal properties of the amplitude envelope are due to the temporal correlation, not its probability distribution.

The low-fluctuation regime ($q < 0$) shows distinctly different scaling properties than the corresponding surrogate data. The significant similarity between the generalized Hurst spectrum of the original amplitude envelope and its corresponding surrogate in the large fluctuation ($q > 0$) part shows that the large fluctuation regime has a correlation structure similar to the surrogate data (IAAFT), which preserves properties up to second order. The significant difference between the generalized Hurst spectrum in the low-fluctuation part ($q < 0$)

and the surrogate data shows that the low-fluctuation regime is dominated by the higher-order correlation properties of the amplitude envelopes. We also note the greater agreement between the multifractal spectrum for surrogate IAAFT data and the original data in the presence of intersyllabic gaps. The behavior of segments 2 and 3 is similar, as discussed in detail in Appendix A 1 and A 2. Segment 5 shows distinct behavior, as can be seen in Fig. 6, and its characteristics are discussed in the subsequent subsection. Segments 4 and 6 show qualitatively similar behavior for low-fluctuation regimes ($q < 0$ with an intersyllabic gap). The analysis of segment 4 and its qualitative similarities with segment 6 can be found in Appendix A 3.

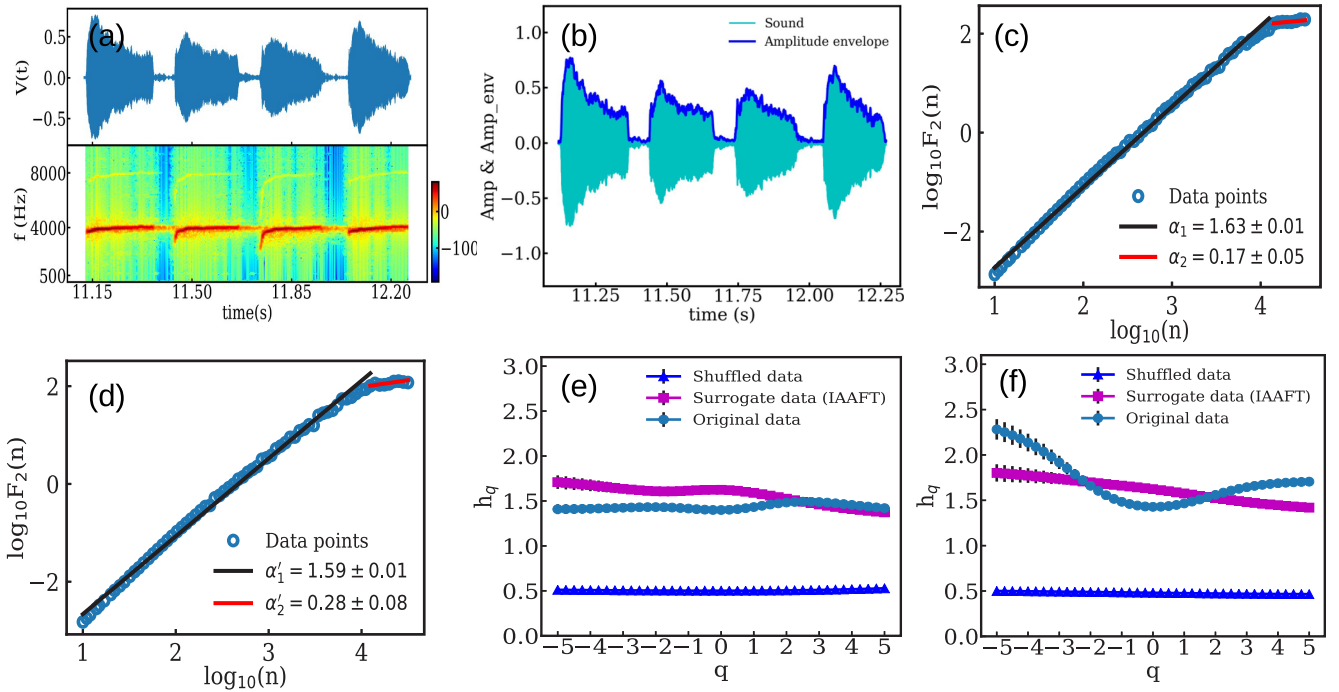


FIG. 6. (a) The sonogram of segment 5. (b) The amplitude envelope along with the song segment. (c) The crossover in the detrended fluctuation of the amplitude envelope profile in the presence of intersyllabic gaps. The slope h_2 in the first scaling region (timescale smaller than the crossover point) is denoted by α_1 . The slope h_2 in the second scaling region (timescales greater than the crossover point) is denoted by α_2 . (d) The crossover in the detrended fluctuation of the amplitude envelope profile in the absence of intersyllabic gaps. (e) The multifractal spectrum h_q vs q spectrum of the song segment's amplitude envelope with intersyllabic gaps. (f) The MF spectrum h_q vs q for the segment for the amplitude envelope without intersyllabic gaps.

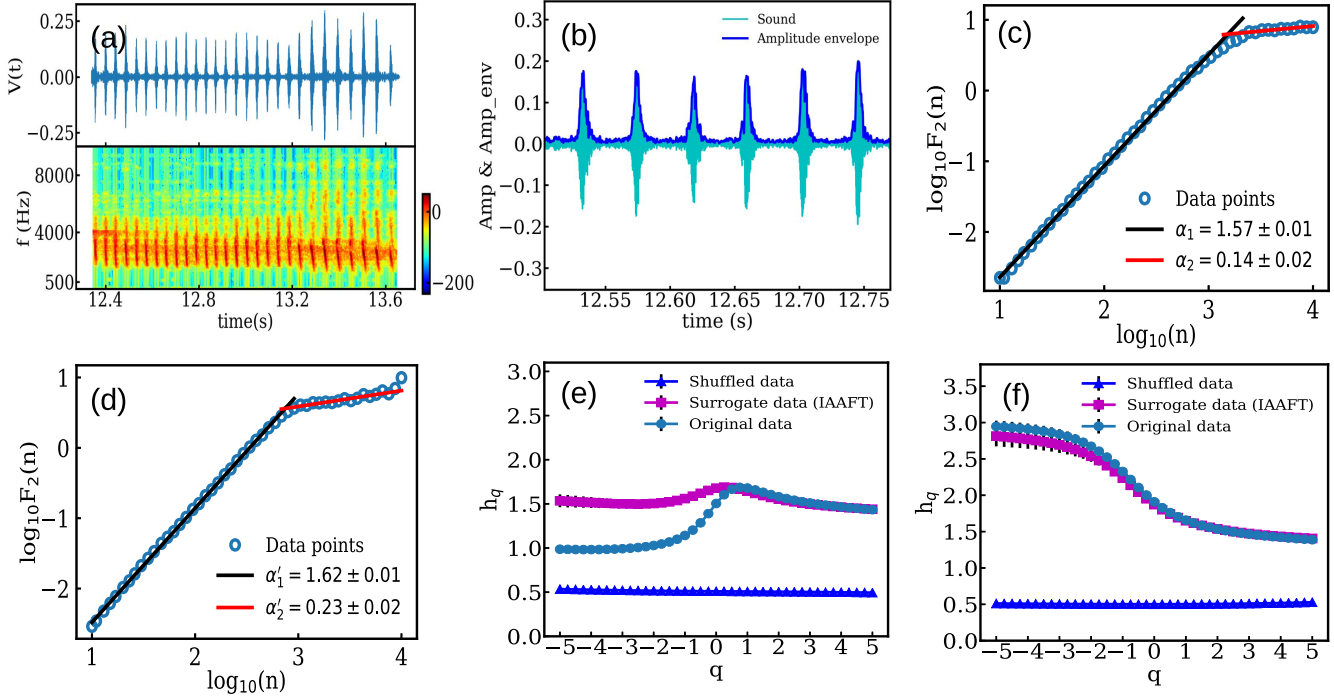


FIG. 7. (a) The sonogram of segment 6. (b) The amplitude envelope along with the song segment. (c) The crossover in the detrended fluctuation of the amplitude envelope profile in the presence of the intersyllabic gaps. The slope h_2 in the first scaling region (timescales smaller than the crossover point) is denoted by α_1 . The slope h_2 in the second scaling region (timescales greater than the crossover point) is denoted by α_2 . (d) The crossover in the detrended fluctuation of the amplitude envelope profile in the absence of intersyllabic gaps. (e) h_q vs q spectrum of the amplitude envelope of the song segment with intersyllabic gaps. (f) h_q vs q for the segment for the amplitude envelope without intersyllabic gaps.

D. Segment 5

This segment consists of four syllables [Fig. 6(a)]. Each of the syllables has tonal nature. We have observed crossover in the detrended (first-order polynomial) fluctuation of the amplitude envelope [see Fig. 6(b)] in the presence and the absence of the intersyllabic gaps in the timescale (n) larger than 10^4 [Figs. 6(c) and 6(d)]. The scaling exponent of the detrended fluctuation of the amplitude envelope profile indicates the presence of a strong long-range correlation [Figs. 6(c) and 6(d)]. There is no significant effect of the intersyllabic gaps on the multifractal spectrum and the crossover point. The scaling features for the large fluctuation region ($q > 0$) of the amplitude envelope and its corresponding surrogate data are significantly different in the generalized multifractal spectrum. The same is true for the low-fluctuation region ($q < 0$). Thus for segment 5, we conclude that both the large and small fluctuation regions show significant contributions from the high-order correlations of the amplitude data.

E. Segment 6

This segment consists of 28 down-sweep syllables [Fig. 7(a)]. The frequency range of these syllables is from 3500 to 1500 Hz. The behavior seen here is similar to that seen for segments 1 to 5, i.e., crossover in the detrended (first-order polynomial) fluctuation of the amplitude envelope [Fig. 7(b)] is seen in the presence and the absence of intersyllabic gaps with the scaling exponent $\alpha_1 = 1.57$ for scales less than the crossover timescale ($n = 1500$) and anticorrelated behavior

for the scales larger than this scale [Fig. 7(c)]. In the absence of intersyllabic gaps, the crossover scale changes to a lower value $n = 740$ [Fig. 7(d)]. In the generalized multifractal spectrum, the scaling features for the large fluctuation region ($q > 0$) amplitude envelope and its corresponding surrogate data are significantly similar (Table II). It is interesting to note that that the surrogate IAAFT data are in good agreement with the original data over the entire q range in the absence of gaps.

F. The comparison of segments (canary song)

The comparison of the characteristics seen in different segments of the canary song is as follows:

(1) There are some significant similarities in the analysis of the phrases. We have observed pronounced crossover in the detrended fluctuations for all the segments. For segments 1 – 4 and 6, the crossover occurs within the range of timescales $10^2 < n < 10^4$, except for the fifth segment, where the crossover occurs at a timescale (n) larger than 10^4 . In all these cases, i.e., for all segments, the crossover corresponds to a change from correlated behavior (i.e., $h_2 > 0.5$) at timescales (n) less than the crossover scale to anticorrelated behavior ($h_2 < 0.5$) at higher timescales (greater than the crossover timescale).

(2) We observe that for all the segments, the crossover positions in $F_2(n)$ are related to the average time period (τ) [Fig. 8(a)] of the repetitive patterns in the time series and are proportional to crossover positions in $F_2(n)$ [Fig. 8(b)]. Figure 8(b) includes the effect of intersyllabic gaps. In the

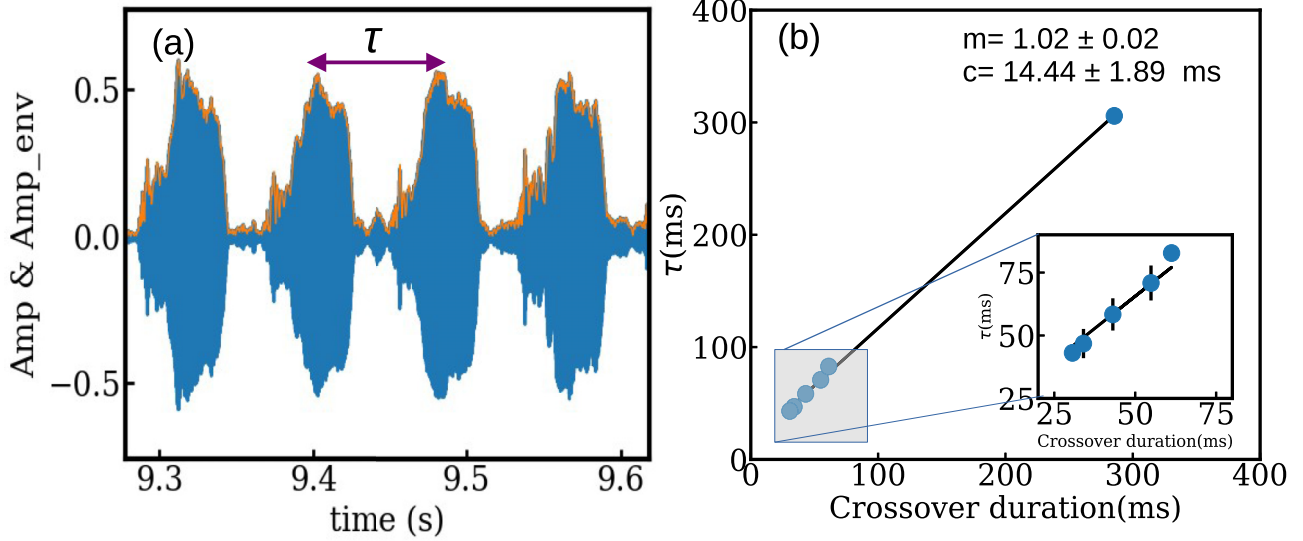


FIG. 8. (a) The time period in the repetitive patterns in the syllabic sequence of segment 3 with intersyllabic gaps is shown. (b) The plot of the average period (τ) of the syllabic sequence in the segments and the crossover durations from the fluctuation [$F_2(n)$] is shown. The fitted line ($y = mx + c$) has slope $m = 1.02 \pm 0.02$ and the intercept $c = 14.44 \pm 1.89$ ms.

absence of intersyllabic gaps, the time periods decrease, and hence the crossover positions shift to smaller values, as is seen for all the segments.

Consistent with the above, the generalized fluctuations [$F_q(n)$] of all the segments show monofractal (minimal multifractal width) properties in the timescale larger than their corresponding crossover points and multifractal behavior below it.

Additionally, all the segments, show significant similarity between the h_q spectrum for the original amplitude envelope and that of the corresponding surrogate data (IAAFT), where the surrogate is tailored to preserve the two-point correlation, in the large fluctuation ($q > 0$) region.

Segments 1–3 have similar multifractal characteristics, whereas segments 4 and 6 have similar multifractal spectra for ($q < 0$) qualitatively. The effects of the intersyllabic gaps in the multifractal spectrum are prominent for the fourth segment and sixth segment compared to those seen in other segments.

We note that the temporal patterns of bronchial pressure and the bird's labial tension determine the birdsong's syllabic structure and frequency evolution. The basic biomechanics action required to produce a syllable is a cycle in the bronchial pressure and the muscle tension [1,34]. The average time period (τ) in the song and the crossover duration in the $F_2(n)$ corresponds to the timescales of different cycles and dynamics involved in the bronchial pressure and the muscle tension required to produce the song.

We have observed that the filtering has a significant effect on the amplitude envelope determination on the multifractal properties. The cutoff frequency determines the detailed temporal structure of the amplitude envelope. The values used here are 300 Hz (for the canary song) and 1000 Hz (for the zebra finch song), respectively. In the MFDFA study, the maximum timescale and common scaling regions in $F_q(n)$ also have to be identified correctly.

G. Kolmogorov-Smirnov statistics for the segments

Table II compares the Kolmogorov-Smirnov statistics for the extent to which the multifractal spectrum of the original data and the IAAFT surrogate data agree with each other in the $q < 0$ and $q > 0$ regimes. The quantity listed is D_{ks} , which is the largest value of the distance between the cumulative distributions of the h_q calculated from the original data set and the surrogate data set. The two distributions are accepted as being similar if this distance D_{ks} does not exceed $D_{critical}(\gamma)$, where γ is the significance level, set at $\gamma = 0.05$ here, which gives $D_{critical} = 0.429$ [35]. The p -values listed indicate the probability with which the similarity between the two distributions is accepted. It can be seen that the KS test confirms our visual observation that there is good agreement between the IAAFT surrogate and the original data in the $q > 0$, i.e., the large fluctuation regime, which does not carry over to the small fluctuation regime. We note that the IAAFT data preserve the autocorrelation property of the original data. Clearly, at the low-fluctuation level, the effects of higher-order correlations and the gaps contribute significantly to the information contained in the original signal, leading to the difference observed between the original and surrogate data.

We observe that for all the segments, the scaling properties of the detrended fluctuation of the large fluctuation region ($q > 0$) are significantly similar with its corresponding surrogate data sets except for the fifth segment. In low-fluctuation regions ($q < 0$), the surrogate and the original data are significantly different in their h_q vs q spectrum except for the third segment (without gaps).

The comparison with shuffled data, which destroy the temporal correlation property but preserve the histogram of the original amplitude envelope, shows that the shuffled data are completely monofractal and are nowhere similar to the multifractal spectrum of the original data. This suggests there is no influence on the probability distribution of the amplitude

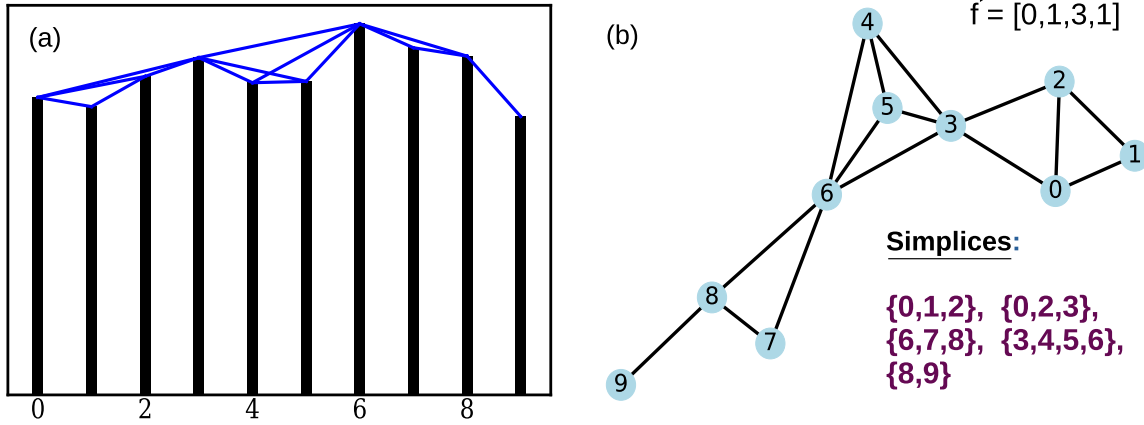


FIG. 9. (a) Connections among the nodes using the visibility algorithm for a small part of energy data, the square of the amplitude. (b) Time series network formed according to visibility algorithm and its different dimensional simplices.

envelope in the multifractal features seen in the detrended fluctuation of the amplitude envelope. We also observe that the nature of the low-fluctuation regions of the time series significantly affects the multifractal spectrum, particularly for the ($q < 0$) region. Most of the low-fluctuation parts correspond to the intersyllabic gaps. If the fluctuation nature of the gaps is multifractal, we get the usual multifractal spectrum. Here, by the usual multifractal spectrum, we mean that the spectrum where $h_q(q > 0)$ is less than $h_q(q < 0)$. On the other hand, if the intersyllabic parts have monofractal like behavior (small multifractal width), we get a spectrum where $h_q(q > 0)$ is greater than the $h_q(q < 0)$ [Fig. 7(e)].

In the case of segment 5 the low p -value seen for $q > 0$ indicates that the h_q spectrum of the original amplitude envelope and the average h_q spectrum of its corresponding IAAFT surrogates seen in Fig. 6(e) are distinctly dissimilar. This implies a substantial contribution of higher-order correlations to the large fluctuation region of segment 5.

H. Source of multifractality

We have discussed the multifractal detrended fluctuation results of the amplitude envelope profile in the previous subsection for individual segments of the song. There are two possible sources of multifractal features [17]. One is the temporal correlation for the window-wise detrended fluctuation, and the other is the long-tailed probability distribution of the amplitude envelope of the time series. To identify the reason for the multifractal property, we have shuffled the amplitude envelope data. The shuffled data have no temporal correlations but preserves its probability distribution. We have observed that the h_q vs q for the shuffled data is now a constant line of $h_q \approx 0.5$. On the other hand, the original amplitude envelope and its IAAFT data show multifractal behavior, although their details differ from each other, as seen in the segment-wise analysis. Thus, the temporal correlation of different magnitudes of the window-wise detrended fluctuation over multiple timescales leads to the multifractal property in the fluctuation of the time series.

I. Simplicial analysis

The simplicial analysis of TS networks has been used in a variety of contexts, including the characterization of

different dynamical regimes of evolving dynamical systems, which have different kinds of short-term correlations [29]. Here the simplicial structures reveal the structure of the short-time correlations in the segments of the birdsong. We demonstrate this via the comparison of birdsong data and the time series of the Lorenz attractor.

Here we use the topological characterizer \vec{f} vector for quantifying the TS networks constructed for the Lorenz attractor (z time series) and the energy, i.e., the square of the amplitude, of segment 3 of the canary birdsong. This quantity retains the variation of the amplitude envelope but is computationally more efficient. We also compare the \vec{f} vector for the original time series and its corresponding surrogate (IAAFT). As earlier, the amplitude-adjusted phase randomized surrogate (IAAFT) time series and the original time series have similar two-point autocorrelation and identical probability distribution.

The simplicial analysis of TS networks has been used in a variety of contexts, including the characterization of different dynamical regimes of evolving dynamical systems, which have different kinds of short-term correlations [29]. Here the simplicial structures reveal the structure of the short-time correlations in the segments of the birdsong. We demonstrate this via the comparison of birdsong data and the time series of the Lorenz attractor.

We construct a TS network from time series data using the visibility algorithm [28]. The visibility graphs are formed by considering the time series data points as nodes, and a link is established between two nodes if there is no obstruction in the line of sight of these two nodes. Two nodes, (y_i, t_i) and (y_j, t_j) , are connected to each other if all other intermediate data points (y_r, t_r) satisfy the following conditions:

$$y_j > y_r + \frac{(y_j - y_i)}{(t_j - t_i)}(t_j - t_r). \tag{8}$$

Figure 9 shows the connections between the nodes using the visibility algorithm for a small part of the energy data and the corresponding network. The TS network is now analyzed using the concepts of cliques and simplicial complexes. We use the Born-Kerbosch algorithm to find all the cliques or simplices from the adjacency matrix (TS network). As said

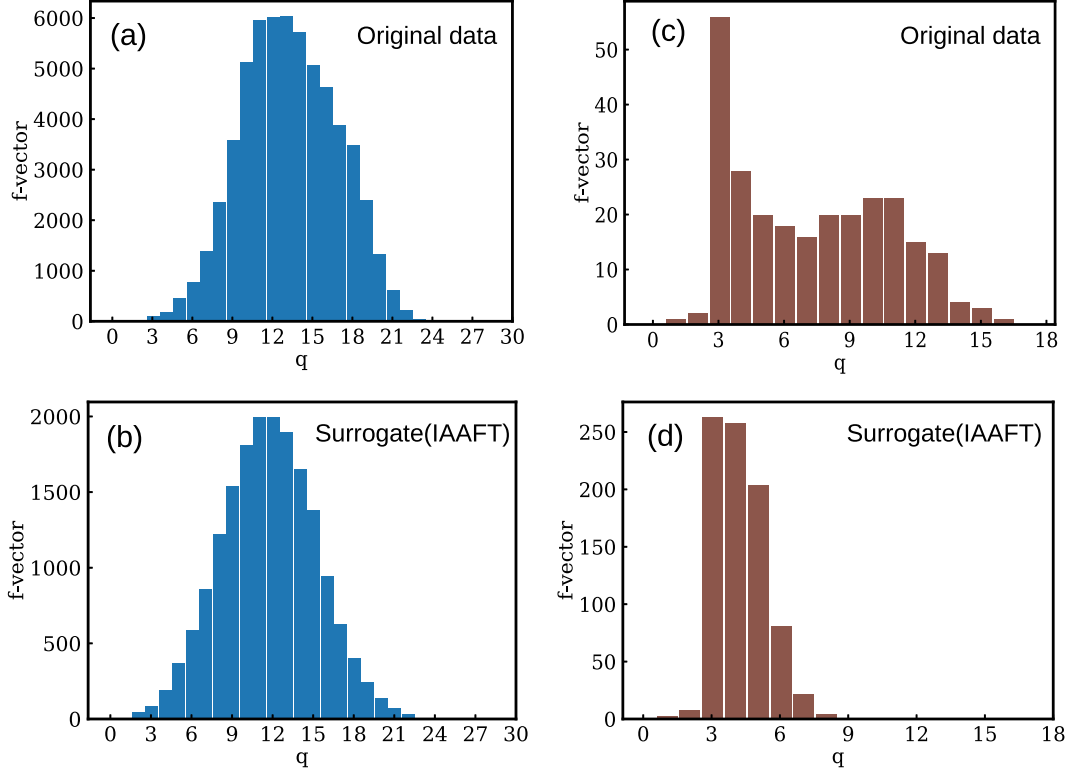


FIG. 10. (a) \vec{f} vector of the energy data of segment 3. The hop and window size are 10 and 40, respectively. (b) \vec{f} vector of the surrogate (IAAFT) of the energy of segment 3. The length of the data is 1000. (c) \vec{f} vector for the Lorentz attractor time series (z time series). The parameters $\rho = 28$, $\sigma = 10$ and $\beta = \frac{8}{3}$. The number of data points $N = 1000$. (d) \vec{f} vector for IAAFT surrogate data of Lorentz attractor. We have considered 10 realizations of the IAAFT surrogate data.

earlier, cliques are complete subgraphs, and simplices are sets of connected nodes, with isolated nodes, two nodes connected by a link, and three nodes connected by links being examples of zero-, one-, and two-dimensional simplices. It is clear from the figure that the simplicial structure reflects the short-term correlations in the system. Figure 9(a) shows the triangles and tetrahedra present in this part of the data. The number of q -dimensional simplices form the q th component of the f -vector, where the $\vec{f} = \{f_0, f_1, \dots, f_{q_{\max}}\}$, and q_{\max} is the dimension of the highest dimensional simplices in the network.

In Figs. 10(a) and 10(b) we have plotted the components of the \vec{f} vector for the energy data of segment 3 and its surrogate IAAFT data. It can be seen that the number of higher-order simplices is reduced for the surrogate (IAAFT) data, indicating that the original data contains higher moments of temporal correlation than the surrogate data. We demonstrate a similar phenomenon for the TS network constructed out of the z variable of the Lorenz attractor. Figures 10(c) and 10(d) show f_q , i.e., the numbers of different dimensional simplices for the Lorenz attractor and its amplitude-adjusted phase randomized surrogate (IAAFT) data. We also observe that the phase randomized surrogate data have more low-dimensional simplices than the original time series. This reduction of the higher dimensional simplices implies a higher order of temporal correlation in the original time series, which gets eliminated due to Fourier phase randomizations. The topological characterizer, *viz.*, the f -vector, indicates the presence of the higher-order correlation structure in the birdsong

data. This complements the conclusion of the Hurst exponent section, where we see that the low-fluctuation regime is dominated by correlations of higher order than the second-order moments, which are retained by the surrogate data.

J. Complexity measure result

In this section, we analyze the López-Ruiz, Mancini, and Calbet (LMC) [30] complexity for each birdsong segment. The consideration of the intersyllabic gaps in the amplitude envelope makes the amplitude envelope of the segments more complex [Fig. 11(a)]. In Table III we see that segment 3 and segment 4 have a lower value of complexity.

In the presence of intersyllabic gaps, the Shannon entropy (S) [Eq. (6)] has small variation across the different segments except segment 6 (Table III). The disequilibrium (D) [Eq. (7)] values of segment 3, segment 4, and segment 5 have low values (Table III). It indicates that the amplitude envelopes of these segments have a more uniform distribution compare to other segments. As a consequence of this, these segments have lower complexity compared to other segments.

We note that the Shannon entropy (S) (Eq. (6)) has no significant variation across all the segments without the intersyllabic gaps (Table III). The significant change in the Shannon entropy (S) of segment six indicates the effect of the intersyllabic gaps in the Shannon entropy amplitude envelope. The disequilibrium (D) of individual segments reduces significantly. This significant change also shows how the

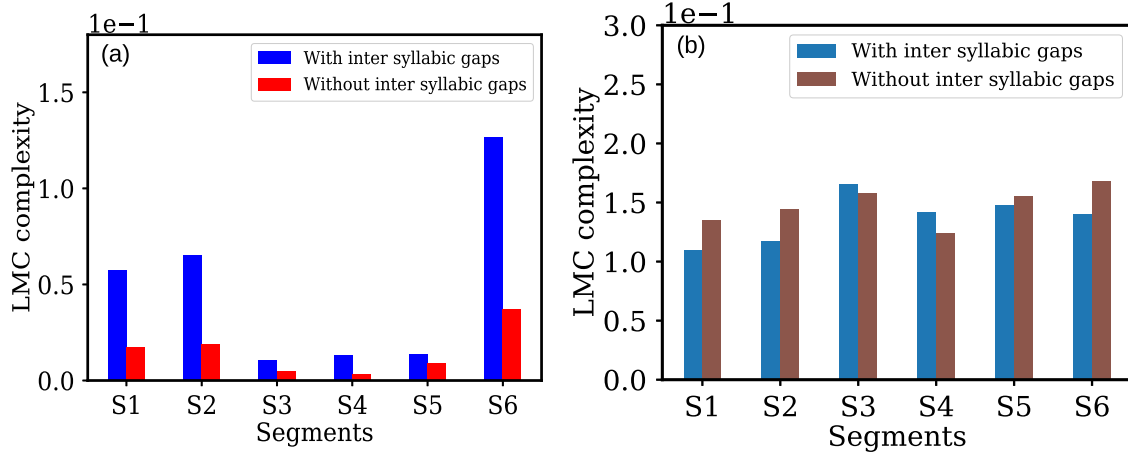


FIG. 11. (a) The segment-wise measure of LMC complexity of the amplitude envelope. The blue (red) bar corresponds to the LMC complexity for the amplitude envelope in the presence (absence) of the intersyllabic gaps. The number of bins is 200, and the data lie between 0 and 1. (b) Segment-wise LMC complexity results in the magnitude of the amplitude envelope’s increment of the presence and absence of the intersyllabic gaps (Table III). The number of bins is 200, and the data lie between 0 and 0.01.

intersyllabic gaps affect the disequilibrium of the amplitude envelope. Consequently, the product of the disequilibrium (D) and Shannon entropy (S) decreases. In Table III we show the LMC complexity for the magnitude of the amplitude envelope increment series. The eight and ninth column show the amplitude envelope increment’s complexity in the presence and absence of intersyllabic gaps. These two columns indicate no significant difference in the LMC complexity across the segments in the amplitude envelope level. It also implies that there is no effect of the intersyllabic gaps across all the segments in the LMC complexity of amplitude envelope increment’s magnitude.

We note that the complexity measures are simple and constitute useful quantifiers of the importance of intersyllabic gaps in the birdsong data. However, the MFDFA contributes a more detailed method of analysis of the timescales and correlations of the problem.

IV. BIRDSONG DATA: FEATURES AND ANALYSIS FOR THE SONG OF THE ZEBRA FINCH

Another commonly studied songbird species is the zebra finch. The birdsong of the zebra finch conforms to a singing

pattern that consists of a single short but complex song. This differs from the singing pattern of the canary, which contains far more controlled tonal sequences. We carry out a multifractal detrending analysis of the zebra finch song to illustrate the differences with the results obtained for the analysis of the song of the canary.

A. Data source

We have carried out our analysis for an example of a zebra finch (*Taeniopygia guttata*) song sourced from a YouTube video [36]. We have extracted the audio data using the 4K Video downloader software. The duration and the sampling rate of the signal are 49 s and 48 000 Hz, respectively. The original signal is a two-channel signal. We have used the Praat software to convert it to a mono channel signal. Next, we have filtered the signal using a bandpass filter (20–20 kHz) in the Praat software.

We have chosen the peak detection algorithm outlined in Sec. II B for the amplitude envelope determination for the zebra finch song. For this process, we have chosen a window of six sample units (n) and a shift of one third the window size, and a low-pass filter with a cutoff frequency of 1000 Hz.

TABLE III. Segment-wise complexity measure for the amplitude envelope and magnitude of amplitude envelope’s increment time series respectively in the presence and absence of the intersyllabic gaps. S , D , and C correspond to the Shannon entropy, disequilibrium, and LMC complexity measure, respectively.

Segment	Amplitude envelope time series						Magnitude of amplitude envelope increment time series	
	With intersyllabic gap			Without intersyllabic gap			With intersyllabic gap	Without intersyllabic gap
	S	D	$C = S \times D$	S	D	$C = S \times D$	$C = S \times D$	$C = S \times D$
1	0.746	0.077	0.057	0.864	0.02	0.017	0.109	0.135
2	0.757	0.086	0.065	0.895	0.021	0.019	0.117	0.145
3	0.908	0.012	0.011	0.941	0.005	0.005	0.166	0.158
4	0.897	0.015	0.013	0.959	0.003	0.003	0.142	0.124
5	0.891	0.015	0.013	0.918	0.01	0.009	0.147	0.155
6	0.574	0.221	0.127	0.847	0.044	0.037	0.140	0.168

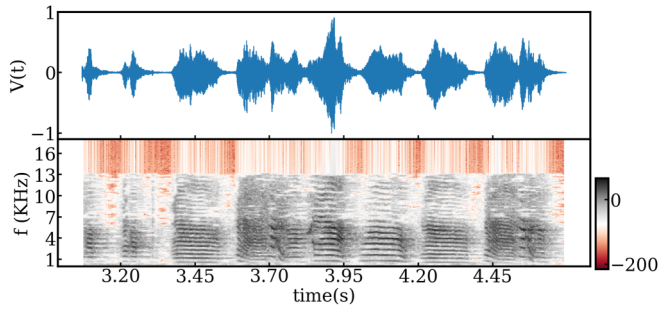


FIG. 12. Audio signal and the sonogram of segment 1. Harmonic structures, different frequency gaps in the harmonics, and their spectral energies are prominent in the syllables and motifs.

We have carried the analysis of two segments of the zebra finch song for comparison with the canary birdsong. We note that this is a proof of principle analysis, and a more detailed comparison of the two species will be carried out in future work.

B. Segment 1 (zebra finch song)

We first discuss the frequency content of the segment shown in Fig. 12. The duration of this segment is from 3.07 to 4.697 s. The sonogram clearly shows the harmonic structures

(see Fig. 12) present in the syllables [6,37]. The corresponding amplitude envelope is shown in Fig. 13(a), the underlying fundamental time period is identified in Fig. 13(b) and Fig. 13(c).

The detrended fluctuation analysis for the quantity $F_2(n)$ is shown in Fig. 13(c). As noted earlier, the behavior of this quantity identifies the scaling of the two-point correlation in the amplitude envelope on different timescales. We observe three crossovers in Fig. 13(c). The first crossover at ($n \approx 10^{1.8} = 63$) indicates the change from a short-ranged correlated structure to anticorrelated behavior for the amplitude envelope. This first crossover point corresponds to the fundamental time period of the original signal. We have computed the time duration to the first crossover point, which is 1.312 ms. The inverse of this time corresponds to the frequency $f_0 = 762$ Hz approximately, which is comparable to the fundamental frequency observed in the sonogram (Fig. 12). We have computed the fundamental time period from the original signal [Fig. 13(b)], and we observe that the fundamental frequency in the original song and the frequency computed using the first crossover duration are in reasonable agreement. It is, therefore, likely that the first crossover duration corresponds to the time period of the fundamental oscillation of the labia of the zebra finch.

We observe the second crossover in the $F_2(n)$ at approximately $10^{2.5} \approx 316$ sample units (n) in Fig. 13(c). The smaller-scale signal in the amplitude envelope within first

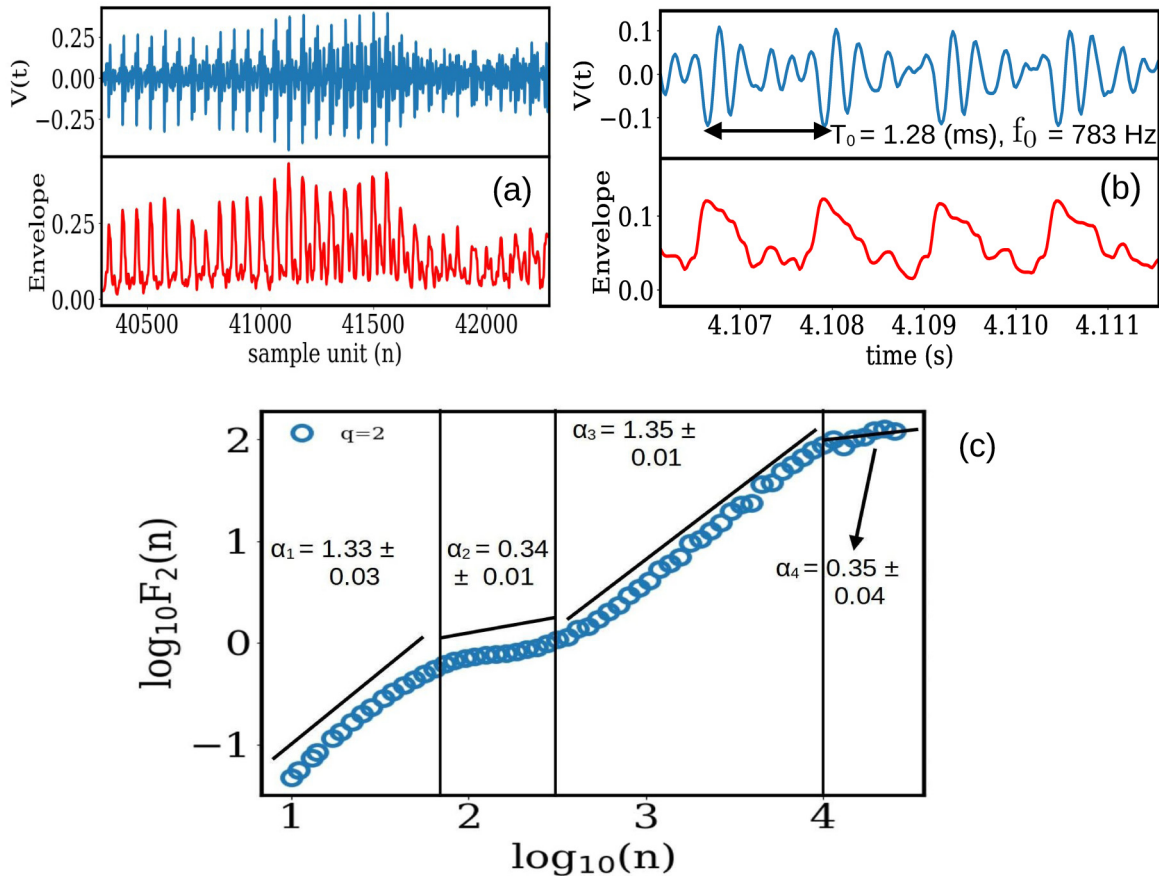


FIG. 13. (a) Some part of the signal (segment 1) and its amplitude envelope. (b) One part of the original signal showing the fundamental time period (top). The corresponding amplitude envelope (bottom). (c) Three crossovers have been shown. The vertical lines correspond to the different crossover regions.

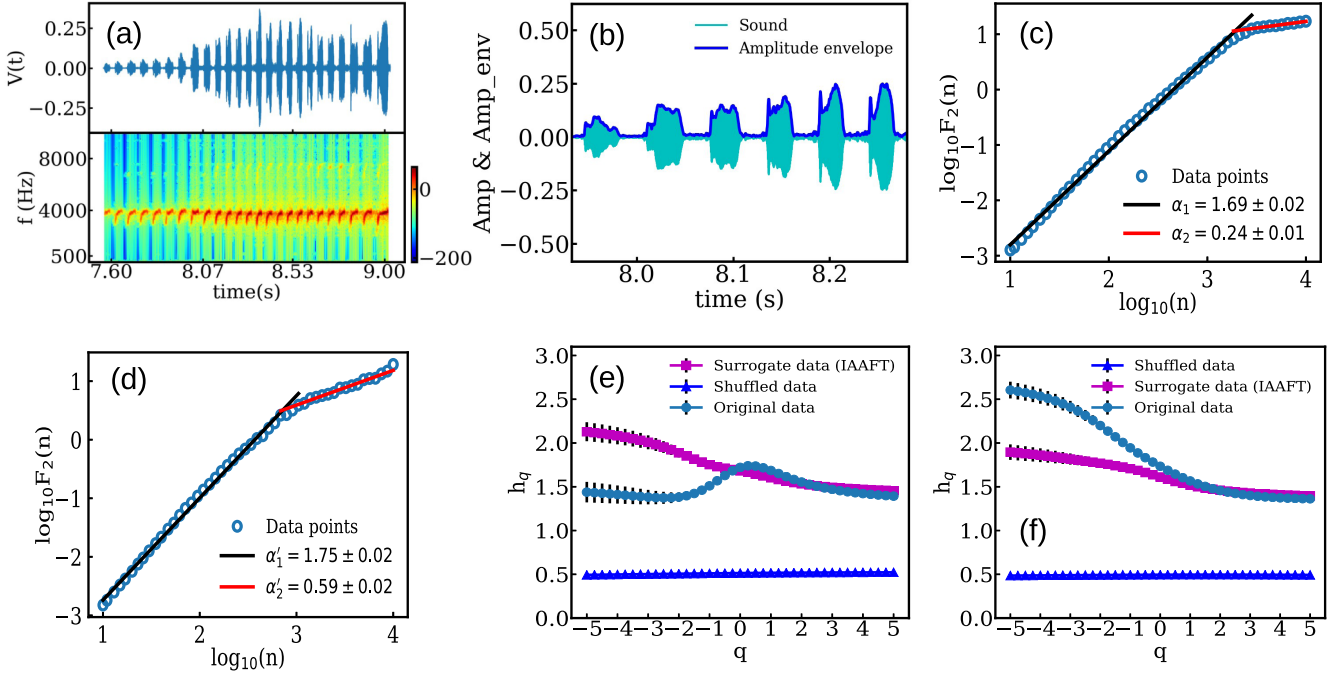


FIG. 14. (a) The sonogram of segment 2. (b) The amplitude envelope along with the song segment. (c) The crossover in the detrended fluctuation of the amplitude envelope profile in the presence of intersyllabic gaps. The slope h_2 in the first scaling region (timescales smaller than the crossover point) is denoted by α_1 . The slope h_2 in the second scaling region (timescales greater than the crossover point) is denoted by α_2 . (d) The crossover in the detrended fluctuation of the amplitude envelope profile in the absence of intersyllabic gaps. (e) The multifractal spectrum h_q vs q spectrum of the song segment's amplitude envelope with intersyllabic gaps. (f) The multifractal spectrum h_q vs q of the segment's amplitude envelope without intersyllabic gaps.

and second crossover point shows the anticorrelation property ($\alpha_2 = 0.34 < 0.5$); whereas for timescales larger than the second crossover point, the quantity shows correlation behavior with exponent ($\alpha_3 = 1.35 > 0.5$) up to the timescale [sample unit (n)] 10^4 . For much larger timescales ($n > 10^4$), the amplitude envelope shows anticorrelated behavior ($\alpha_4 = 0.35 < 0.5$) [Fig. 13(c)]. This large-scale anticorrelation is due to the effect of the repetitive nature of the syllables and the motifs present in the segment (Fig. 12).

The generalized fluctuation curves $F_q(n)$ are plotted in [Fig. 17(a)] of Appendix B 1 for different q -values ($q = -5, -2, 2, 5$), and clearly indicate five distinct regions (I–V, respectively). Each scaling region seen here is very short, and good power-law behavior is not seen over scaling regions of reasonable length for any of the regions except region IV. Nevertheless, a discussion of each region, the extent to which it follows power-law behavior, and the surrogate analysis of each region can provide valuable insights for the analysis of the song. Hence, we carry out the multifractal analysis of the zebra finch song in Appendix B. This analysis is carried out for both segments 1 and 2. Appendices B 1 and B 2 carry out the multifractal analysis of segment 1 (zebra finch song). Appendix B 3 carries out the corresponding analysis of segment 2 for both $F_2(n)$ and $F_q(n)$. We note that both segment 1 and segment 2 show similar behavior for both quantities. The details of the comparison are as follows:

(1) We have observed similar scaling properties in the detrended fluctuation of the amplitude envelope of segment 1 and segment 2 of the zebra finch song. The first and second crossover positions and scaling properties are similar to those

seen for segment 1 [Figs. 13(c) and 19(c)]. We observe the difference in the large-scale ($n > 10^4$) scaling properties in these two segments. This is identified as region V in Fig. 17(a).

(2) The comparison of the multifractal properties and that of the surrogate for segment 1 (zebra finch song) is discussed in Appendices B 1 and B 2, and in Appendix B 3 for segment 2, up to the accuracy observed in the scaling regions. We observe a difference in the multifractal and two-point correlation structures for the longer timescales ($n > 10^4$). In this timescale, the effect of the repetitive patterns of the syllables and motifs are prominent [27]. The analysis shows a smaller correlation exponent.

Since this is a brief comparison, we have not discussed the intersyllabic gaps in the multifractal properties in this context. We hope to include it in a separate study.

V. THE COMPARISON OF THE CANARY AND ZEBRA FINCH SONG

The following important features were observed in the zebra finch and canary birdsongs via the detrended fluctuation analysis of the amplitude envelope and the sonogram:

(1) The sonogram indicates that zebra finch syllables have a more varied spectrum of frequencies compared to the song of the canary. The spectral energies in the zebra finch song are distributed over a more extensive frequency range for the zebra finch song (from few 100 Hz to 13 kHz) (Figs. 12 and 18); whereas for the song of the canary, the spectral energies are confined to the smaller frequency range (1–6 kHz approximately) [Figs. 5(a), 6(a) and 7(a)] and

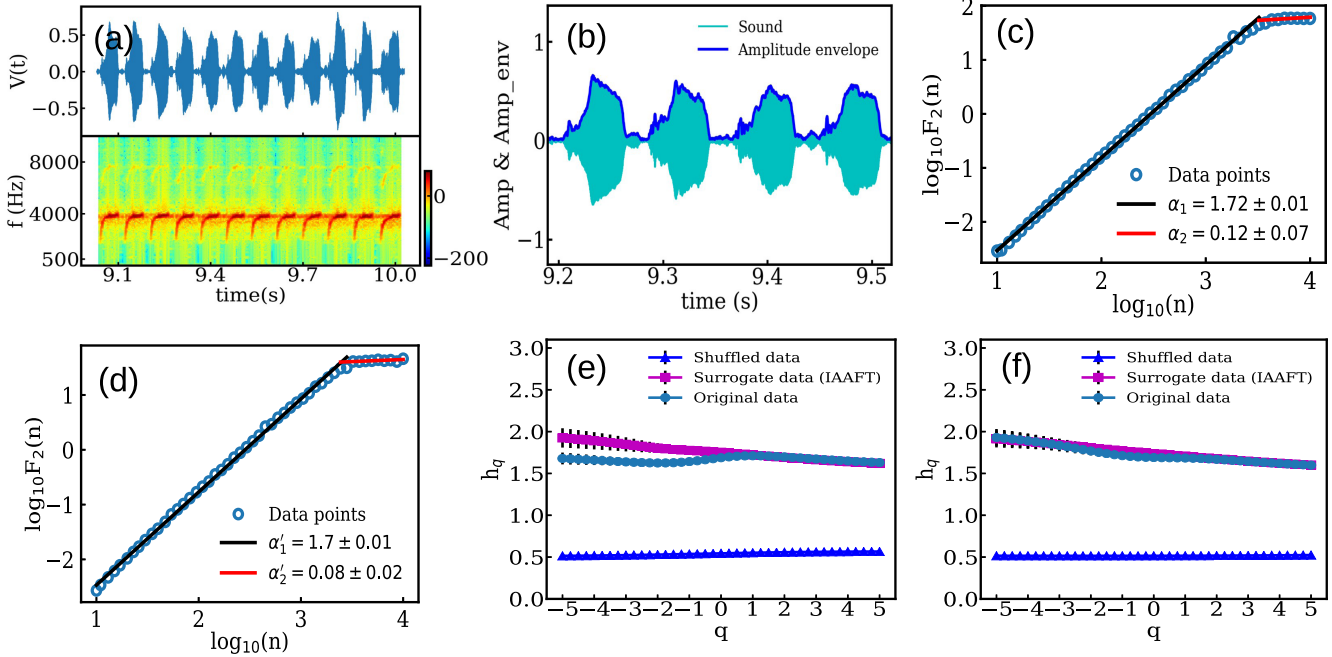


FIG. 15. (a) The sonogram of segment 3. (b) The amplitude envelope along with the song segment. (c) The crossover in the detrended fluctuation of the amplitude envelope profile in the presence of intersyllabic gaps. The slope h_2 in the first scaling region (timescales smaller than the crossover point) is denoted by α_1 . The slope h_2 in the second scaling region (timescales greater than the crossover point) is denoted by α_2 . (d) The crossover in the detrended fluctuation of the amplitude envelope profile in the absence of intersyllabic gaps. (e) h_q vs q spectrum of the song segment’s amplitude envelope with intersyllabic gaps. (f) h_q vs q for the segment’s amplitude envelope without intersyllabic gaps.

[Figs. 14(a) and 16(a)]. We observe harmonicity in the spectrogram of the zebra finch song (Figs. 12 and 18); whereas the canary song shows distinct up-sweep, tonal, and down syllables. Thus, the canary song is purer and indicative of

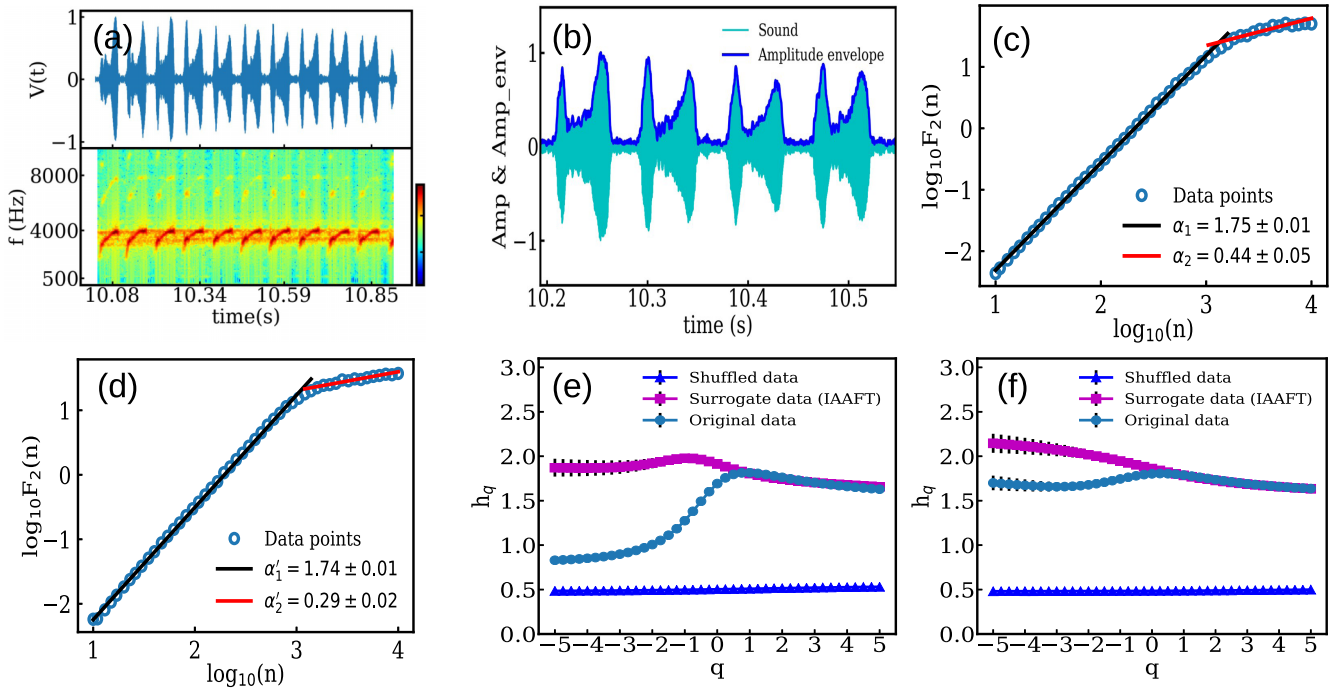


FIG. 16. (a) The sonogram of segment 4. (b) The amplitude envelope along with the song segment. (c) The crossover in the detrended fluctuation of the amplitude envelope profile in the presence of intersyllabic gaps. The slope h_2 in the first scaling region (timescale smaller than the crossover point) is denoted by α_1 . The slope h_2 in the second scaling region (timescales greater than the crossover point) is denoted by α_2 . (d) The crossover in the detrended fluctuation of the amplitude envelope profile in the absence of intersyllabic gaps. (e) h_q vs q spectrum of the amplitude envelope of the song segment with intersyllabic gaps. (f) h_q vs q of the segment’s amplitude envelope without intersyllabic gaps.

better controlled vocal characteristics than the zebra finch song.

(2) The $F_2(n)$ vs n plots of the amplitude envelope of the zebra finch reveal three crossover points (i.e. four regions) [Figs. 13(c) and 19(c)] for each segment, whereas those for the canary only show one (i.e. just two regions) [Figs. 5(c), 6(c), and 7(c)]. For the zebra finch, the first crossover point is seen at small scales, viz., $n \approx 63$, a value which reflects the fundamental time period (T_0) of the song. The crossover behavior here is a change from small-scale correlations to anticorrelated behavior. The second crossover point corresponds to the change of the amplitude envelope from anticorrelated behavior to correlated behavior at large scales. The third crossover point is determined by the effect of the timescale of the repetitive occurrence of the syllables and motifs in the zebra finch song [Figs. 13(c) and 19(c)]. In contrast, for the canary song, we observe a single pronounced crossover from correlated scaling behavior, to anticorrelated behavior, at a crossover point which indicates the timescale of the repetitive patterns of the syllables in the segments.

(3) The syllabic durations and intersyllabic gaps within a particular phrase of a canary song are less variable among themselves. The repetition of syllabic gaps makes a periodic trend which leads to a pronounced crossover in the $F_2(n)$. The scaling exponent in the timescale larger than the crossover point shows anticorrelation in the two-point correlation structure [27].

To summarize, multifractal detrended fluctuation analysis is an effective tool for songbirds like canaries (*Serinus canaria domestica*), nightingales (*Luscinia megarhynchos*), warblers (Parulidae), finches (Fringillidae), mynas (*Acridotheres tristis*), and many others. These songbirds demonstrate good vocal control and therefore have more structured and complex songs. On the other hand call birds usually produce isolated syllables and less complex structured signals, as we observe in species like coppersmith barbets (*Megalaima haemacephala*), great tit birds (*Parus major*), and others. These call signals often have very large intersyllabic gaps. Sometimes, these gaps have zero amplitude. The amplitude envelope of these gaps makes the generalized fluctuation F_q unstable for $q < 0$. The correlation structures in the birdcall can be analyzed after concatenating the syllables, which corresponds to the correlation structure of the syllabic sequence, not in the actual bird call.

VI. CONCLUSIONS

To conclude, our analysis shows that the detrended fluctuation of the amplitude envelope of the song of the canary shows multifractal behavior. The temporal shuffling of these data destroys the multifractal features, confirming the contribution of the temporal correlations to the observed multifractal behavior. The comparison of the Hurst exponent spectrum with that for surrogate data generated by the IAAFT transforms shows a good match for most segments in the high-fluctuation regime, and a poor match with the spectrum in the low-fluctuation regime, indicating that higher-order correlations contribute significantly in the low-fluctuation regime. Intersyllabic gaps contribute significantly to low-fluctuation regimes, indicating their importance for the typical characteristics of the birdsong.

These results are supported by the simplicial characterization of the TS networks. Complexity measures give results consistent with this analysis. We note that our methods permit a fine-scale characterization of the time series under analysis, by which a very detailed comparison can be made between sets of interest to identify the similarities and differences between them. The brief analysis of the zebra finch song carried out here supports this. Hence, our methods can contribute to the quantitative comparison of vocal characteristics of songbirds across species, as well as to the comparison of vocal skills in a given species at different stages of development.

ACKNOWLEDGMENT

The authors thank IIT Madras for the CoE project SP20210777DRMHRDDIRIIT.

APPENDIX A: MULTIFRACTAL ANALYSIS OF AMPLITUDE ENVELOPE FOR CANARY SONG

The MF DFA analysis and the correlation structure of the amplitude envelope of the song of the canary have been discussed in detail in Sec. III. These are analyzed in detail for segments 1, 5, and 6 in Secs. III C–III E, respectively. In this Appendix we discuss segments 2 and 3, whose behavior is qualitatively similar to segment 1 as well as segment 4, whose behavior resembles that of segment 6 in the low-fluctuation part.

Section IV also carries out a proof of principle analysis of the amplitude envelope of the zebra finch song. This includes the analysis of the second moment $F_2(n)$ and the conclusions drawn from this analysis. We note that the zebra finch song is characterized by the presence of marked higher harmonics, and as a result, the multifractal moments $F_q(n)$ contain many regions where approximate scaling is seen over very short segments or timescales n . Nevertheless, these segments contain complex information which is important for the birdsong. Hence we discuss this information in the Appendix B. We will discuss in detail the behavior of $F_q(n)$ vs n for segments 1 and 2 of the zebra finch, in Appendix B together with the approximate fits to scaling behavior over short segments and their error bars.

We discuss segment 2 (canary song) as seen in Fig. 14.

1. Segment 2 (canary song)

This segment consists of 25 syllables. Each of the syllables consist of an up-sweep followed by tonal behavior. This part of the song (segment 2) and the corresponding amplitude envelope is shown in Fig. 14(b). We have observed a crossover in the detrended (first-order polynomial) fluctuation of the amplitude envelope both in the presence and the absence of the intersyllabic gaps. The scaling exponent is $\alpha_1 = 1.69$ for scales less than the crossover timescale ($n \approx 1900$). This shows a strong long-range correlation in the detrended fluctuation of the amplitude envelope profile. For scales higher than the crossover scale, the anticorrelation property is seen [Fig. 14(c)].

As in the case of segment 1, the crossover behavior is similar in the presence and absence of intersyllabic gaps, but the crossover scale shifts to much lower values in the

absence of intersyllabic gaps [Figs. 14(c) and 14(d)]. For this case, the scaling exponent is $\alpha'_1 = 1.75$ in the scale less than the crossover timescale ($n \approx 700$) which is smaller, as expected. For the generalized multifractal spectrum, the scaling features for the large fluctuation region ($q > 0$) of the amplitude envelope, and its corresponding surrogate data are significantly similar [Fig. 14(e)]. The low-fluctuation parts ($q < 0$) show scaling properties dissimilar to the corresponding surrogate data in the presence of intersyllabic gaps [Fig. 14(e)]. The similarity between the generalized Hurst spectrum of the original amplitude envelope and its corresponding surrogate in the large fluctuation ($q > 0$) part shows that the large fluctuation regions have a correlation structure like the surrogate, i.e. the two-point correlations dominate here. The significant difference between the generalized Hurst spectrum in the low-fluctuation part ($q < 0$) shows that the low-fluctuation regime is dominated by higher-order correlation structure in the original amplitude envelope. We note that this is also pronounced in the absence of intersyllabic gaps [Fig. 14(f)].

It is also clear that the multifractal behavior of the amplitude envelope has a significant contribution from its temporal correlations, and not from the probability distribution of the amplitudes [Figs. 14(e) and 14(f)].

Segment 3 (canary song) as seen in Fig. 15 also shows similar behavior (with intersyllabic gaps). We discuss this below.

2. Segment 3 (canary song)

This segment consists of 12 syllables [Fig. 15(a)]. Each of the syllables has an initial up-sweep and is followed by tonal behavior. Figure 15(b) shows the part of the song (segment 3) and its amplitude envelope. We have observed crossover in the detrended (first-order polynomial) fluctuation of the amplitude envelope in the presence and the absence of the intersyllabic gaps. The scaling exponent is $\alpha_1 = 1.72$ for scales less than the crossover timescale ($n \approx 2700$) which indicates a strong long-range correlation in detrended fluctuation of the amplitude envelope profile [Fig. 15(c)]. We observe the anticorrelation property on timescales greater than the crossover [Fig. 15(c)]. As seen in other cases, the intersyllabic gaps have an effect on the multifractal spectrum and the segment's crossover point. Without the intersyllabic gaps, the nature of the detrended fluctuations' dynamics does not change, but the crossover occurs at a much lower timescale ($n \approx 2450$) [Fig. 15(d)].

In the generalized multifractal spectrum (with intersyllabic gaps), the scaling features for the large fluctuation region ($q > 0$) amplitude envelope, and its corresponding surrogate data are significantly similar as in the earlier cases but are quite different for the low-fluctuation $q < 0$ regime [Fig. 15(e)]. This indicates the importance of the two-point correlation structure for positive q and that of the higher-order correlation components for negative q . In the absence of the intersyllabic gaps, the multifractal spectrum shows significant similarities in the correlation structure in the low-fluctuation part ($q < 0$) of the original data and the corresponding surrogate data [Fig. 15(f)].

We note that segments 2 and 3 discussed here show behavior which is similar to the behavior seen in segment 1 (canary) in the presence of intersyllabic gaps, as discussed in Sec. III C. Segment 4 as seen in Fig. 16, however, shows different behavior. We discuss this below.

3. Segment 4 (canary song)

This segment has 11 syllables [Fig. 16(a)]. Figure 16(b) shows part of the song (segment 4) and its amplitude envelope. We have observed crossover in the detrended (first-order polynomial) fluctuation of the amplitude envelope in the presence and the absence of the intersyllabic gaps. The scaling exponent is $\alpha_1 = 1.75$ seen in the scales less than the crossover timescale ($n \approx 1350$) indicates a strong long-range correlation in detrended fluctuation of the amplitude envelope profile [Fig. 16(c)]. We observe the anticorrelation property on the larger timescales [Fig. 16(c)]. We note that the crossover between the two regions now takes place in a more gradual fashion, compared to the first three segments.

The intersyllabic gaps have a significant effect on the multifractal spectrum and crossover point of the segment. As before, without the intersyllabic gaps, the nature of the detrended fluctuation dynamics do not change qualitatively, but the crossover occurs at a much lower timescale ($n \approx 1100$). Again, the scaling features for the generalized multifractal spectrum, the scaling features for the amplitude envelope, and its corresponding surrogate data are significantly similar for the large fluctuation region, ($q > 0$), and not so for the $q < 0$ (small fluctuation region) [Fig. 16(e)]. The effect of the intersyllabic gaps is seen far more strongly here than in the earlier case [Fig. 16(f)]. We note that the difference between the original and IAAFT data in the small fluctuation regions are much more prominent in the presence of the intersyllabic gaps. It is therefore possible to conjecture that the gaps contribute strongly to the higher-order correlations in the amplitude envelope here. This contribution is much smaller in the other segments. We note that similar behavior is seen in segment 6 (canary song) of Fig. 7 and has been discussed in detail in Sec. III E.

This completes the analysis of all segments of the song of the canary, which is one of the cleanest examples of birdsong based on sustained frequencies. To contrast the behavior seen here with another case, we have discussed the case of the zebra finch in Sec. IV. This song of the zebra finch contains the presence of multiple higher harmonics which contribute their own signatures to the MF DFA analysis. Section IV B discusses the behavior of $F_2(n)$ for the first segment of the the zebra finch song. Here we discuss the behavior of $F_q(n)$, its behavior in different regions, and the accuracy with which it approximates scaling behavior in each region.

APPENDIX B: FURTHER DETAILS OF MULTIFRACTAL ANALYSIS OF ZEBRA FINCH SONG

Section IV demonstrates the detrended fluctuation ($F_2(n)$) of the amplitude envelope of segment 1 of the zebra finch song, which shows four different scaling regions. In contrast, the generalized detrended fluctuations $F_q(n)$ vs n , for the same segment, shows five different scaling regions in the log-log

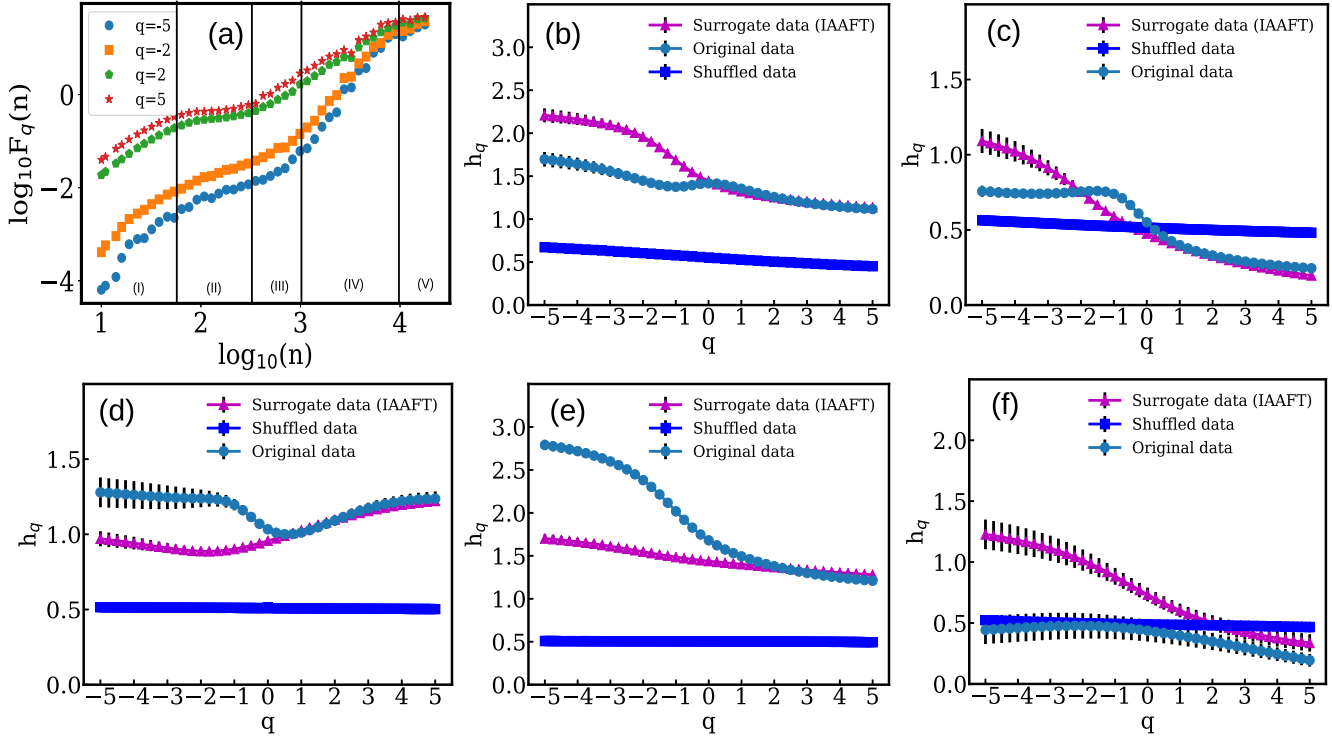


FIG. 17. (a) The fluctuation of the amplitude envelope for q values $(-5, -2, 2, 5)$ bottom to top. Five different scaling regions (I–V, respectively) have been shown. (b–f) h_q vs q of the original data and corresponding surrogate (IAAFT) and the shuffled data for different scaling regions shown in Fig. 17(a), respectively. We have considered the power spectrum threshold (PSD) to be 2×10^{-7} for all the IAAFT surrogate data generation processes [10]. The average has been taken over 50 surrogate data sets (IAAFT) and 50 shuffled data sets.

scale. We note that regions 3 and 4 can be regarded as a single scaling region for $q = 2$, and has been shown as a single region (region 3) for the plot of $F_2(n)$ in Fig. 13(c). The other values of q show the difference clearly [Fig. 17(a)]. None of these regions [except region IV in Fig. 17(a)] show good scaling behavior over reasonably long scaling regions for any of the generalized fluctuations $[F_q(n)]$ in a strict sense. Despite this, a discussion of each region and the extent to which it can approximate reasonable scaling behavior can provide insights into the correlations in the amplitude envelope of the zebra finch song. We, therefore, describe the multifractal detrended fluctuation analysis of the amplitude envelope for segment 1 of the zebra finch song in the subsequent section. In the case of segment 2 of the zebra finch song, we describe both the detrended fluctuation analysis (i.e., the $F_2(n)$ case) and the multifractal analysis $F_q(n)$ case in this section.

1. Multifractal detrended fluctuation analysis of segment 1(zebra finch song)

The generalized fluctuation curves $F_q(n)$ plotted in [Fig. 17(a)] for different q -values ($q = -5, -2, 2, 5$), indicate five common scaling regions (I–V, respectively).

2. Scaling behavior (zebra finch song segment 1)

(1) Region I corresponds to a very small timescale ($1 < \log_{10} n < 1.8$). In this timescale, the average h_q of the surrogate data (IAAFT) spectrum is significantly similar to the $F_q(n)$ for the large fluctuation region ($q > 0$). On the other

hand, the exponents of the multifractal spectrum of the shuffled data are not equal to 0.5 in the $q < 0$ regime [Fig. 17(b)] indicates the minimal scale correlation present in the low-fluctuation parts of the shuffled data. The common scaling region for all the $F_q(n)$ is not long enough, and the scaling exponents of the power-law fitted lines for $q = -5, -2, 2, 5$ are $1.69 \pm 0.08, 1.45 \pm 0.05, 1.26 \pm 0.03,$ and, 1.12 ± 0.04 respectively.

(2) In the second region, we find approximate power-law behavior for a small window ($1.8 < \log_{10} n < 2.42$) for all the $F_q(n)$. The average h_q of the surrogate data (IAAFT) for $q > 0$ is similar to the multifractal spectrum of the original data for $q > 0$, whereas for $q < 0$, we observe a significant difference between them [Fig. 17(c)]. The shuffled data show

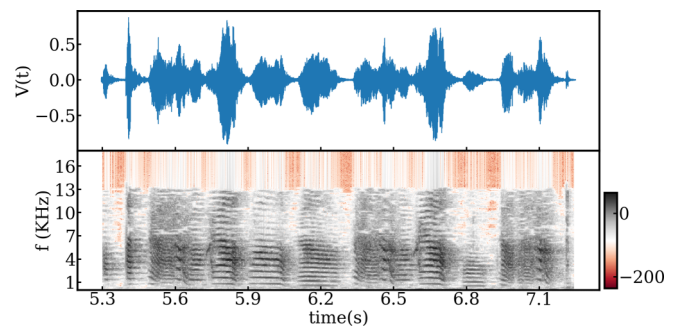


FIG. 18. Spectrogram of segment 2. Harmonic structures are prominent in the syllables and motifs. Different frequency gaps in the harmonics and spectral energies syllable-wise are visible.

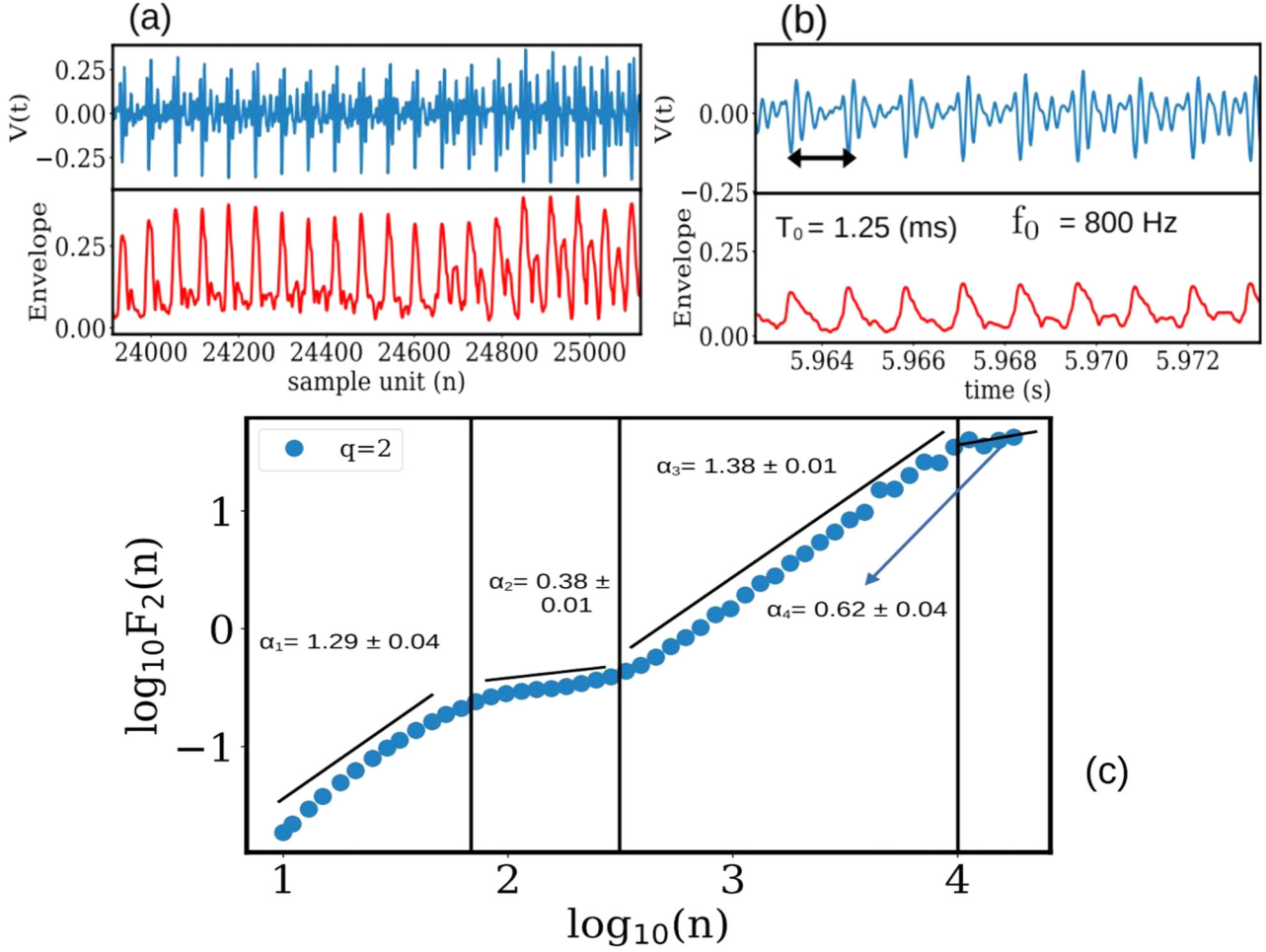


FIG. 19. (a) Part of the signal and its amplitude envelope. (b) Part of the original signal showing the fundamental time period (top). The amplitude envelope (bottom). (c) Three crossovers have been shown in $F_2(n)$ vs n graph. The vertical lines correspond to the different crossover regions for segment 2.

a minimal multifractal width around 0.5 [Fig. 17(c)]. This constant line around 0.5 indicates that the multifractality is due to the temporal order of the amplitude envelope. Like region I, the common scaling region for all the $F_q(n)$ is not long enough. Therefore the presence of a multifractal property is not a robust argument for this scale. The scaling exponents of the power-law fitted lines for $q = -5, -2, 2, 5$ are $0.76 \pm 0.04, 0.75 \pm 0.01, 0.33 \pm 0.01, 0.24 \pm 0.01$, respectively.

(3) Region III has a scaling range ($2.42 < \log_{10} n < 3$). Over this timescale, we observe significant deviation between the average h_q spectrum of the surrogate (IAAFT) data and the original amplitude envelope for $q < 0$. Thus three point and higher-order correlations are dominant for $q < 0$ [Fig. 17(d)]. The scaling exponents of the power-law fitted lines for $q = -5, -2, 2, 5$ are $1.28 \pm 0.1, 1.24 \pm 0.06, 1.09 \pm 0.03, 1.24 \pm 0.05$ respectively. This region can also be regarded as a crossover region.

(4) Region IV has a long scaling range ($3 < \log_{10} n < 4$). In this region, we find better long-range scaling compared to all other regions. The scaling exponents of the power-law fitted lines for $q = -5, -2, 2, 5$ are $2.79 \pm 0.06, 2.38 \pm 0.04, 1.38 \pm 0.01, 1.21 \pm 0.02$, respectively. In this timescale, we observe significant deviation between the average h_q spectrum of the surrogate (IAAFT) data and

the original amplitude envelope for $q < 0$. The effect of the higher-order correlation (nonlinearity) in the original amplitude envelope is prominent in the low-fluctuation regions ($q < 0$) [Fig. 17(e)]. We observe good agreement between the scaling properties between the IAAFT surrogate and the original data in the large fluctuation regions [Fig. 17(e)].

(5) In region V the amplitude envelope of the original data has anticorrelation properties [Fig. 17(f)]. The scaling exponents of the power-law fitted lines for $q = -5, -2, 2, 5$ are $0.44 \pm 0.12, 0.48 \pm 0.1, 0.39 \pm 0.07, 0.19 \pm 0.05$, respectively. In this timescale, the effect of the repetitive patterns of the syllables and motifs present in the segment are prominent. As a consequence of the repetitive nature of the motifs and the syllables, the amplitude envelope is anticorrelated in nature [27]. We observe disagreement between the multifractal spectrum of the original amplitude envelope and its corresponding surrogate data (IAAFT). In this context, we observe that the surrogate data generated using the IAAFT method do not preserve the two-point correlations in this timescale. One reason for this could be the presence of phase correlations as has been reported in the IAAFT surrogate data in Ref. [33] for the case of the AGN (active galactic nuclei) time series. Further analysis is required to confirm this for the birdsong data.

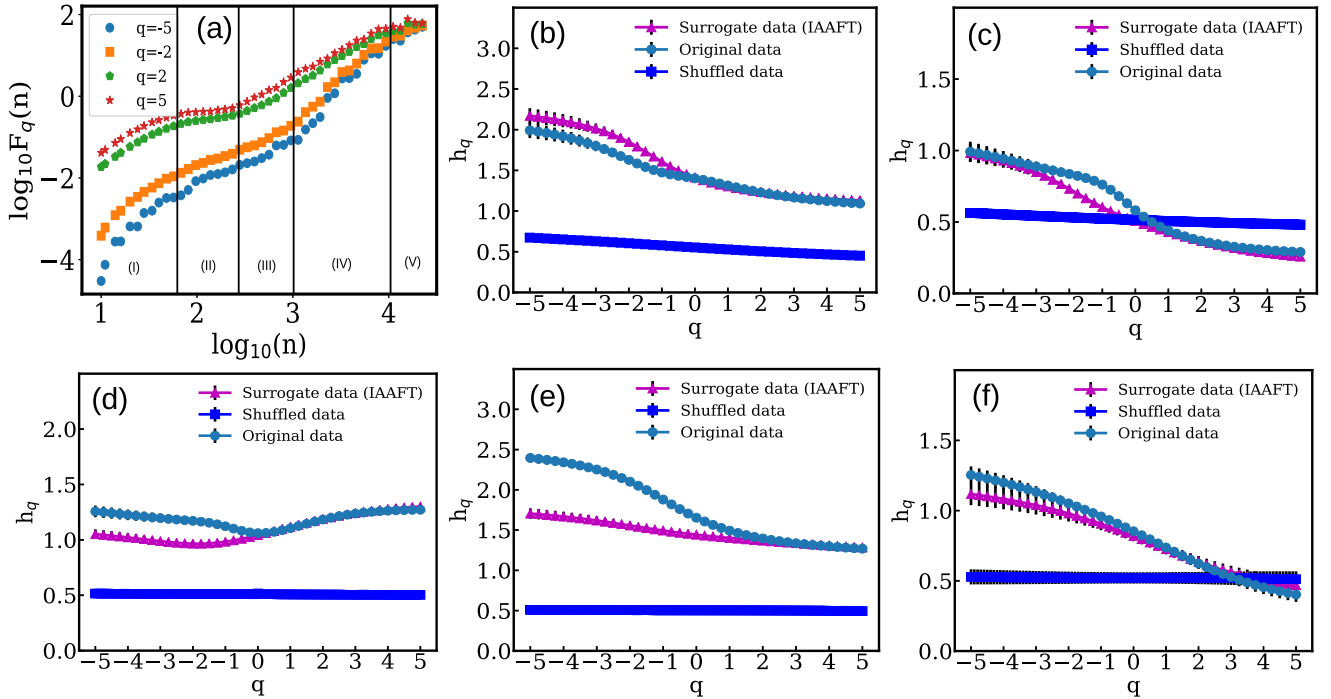


FIG. 20. (a) The fluctuation of the amplitude envelope for q values $(-5, -2, 2, 5)$ bottom to top. Five different scaling regions (I–V, respectively) have been shown. (b–f) h_q vs q of the original data and corresponding surrogate (IAAFT) and the shuffled data for different scaling regions shown in panel (a), respectively. We have considered the power spectrum threshold (PSD) as 2×10^{-7} for all the IAAFT surrogate data generation processes [10]. The average has been taken over 50 surrogate data sets (IAAFT) and 50 shuffled data sets.

3. Segment 2 (zebra finch song)

In the main text, we have discussed the detrended fluctuation $F_2(n)$ for the amplitude envelope of segment 1 of the zebra finch song in Sec. IV. Here we describe the detrended fluctuation analysis ($F_2(n)$) followed by a multifractal detrended fluctuation analysis of the amplitude envelope of segment 2 of the zebra finch song. The duration of the second segment is from 5.295 to 7.25 s. As in the case of the first segment, the sonogram of this segment shows the harmonic nature of its syllables and motifs [6]. The first crossover in Fig. 19(c) corresponds to the fundamental frequency of the song (Fig. 18). The first crossover corresponds to a change from small-scale correlations to anticorrelated behavior in the amplitude envelope. The first crossover duration corresponds to the time period 1.312 ms [Fig. 19(c)]. The corresponding frequency is 762 Hz; whereas the direct measurement of the fundamental time period ($T_0 = 1.25$ ms, $f_0 = 800$ Hz). As discussed earlier for the first segment, we can speculate that the first crossover point observed in the detrended fluctuation analysis of the amplitude envelope corresponds to the fundamental time period of the labial oscillations in the zebra finch song. Note that this direct fundamental time period from the song should be measured separately for different syllables as it can vary for each syllable. We have observed the second crossover in the $F_2(n)$ at approximately $10^{2.5} \approx 316$ sample units (n) [Fig. 19(c)]. For smaller scales (i.e., the signal between the first and second crossover points), the amplitude envelope shows the anticorrelation property ($\alpha_2 = 0.38 < 0.5$); whereas for timescales larger than the second crossover point, the envelope shows correlated behavior

($\alpha_3 = 1.38 > 0.5$) up to the timescale (sample unit (n)) 10^4 . For much larger timescales ($n > 10^4$), the amplitude envelope has shown weakly correlated behavior ($\alpha_4 = 0.62 > 0.5$) [Fig. 19(c)]. Note that the correlation strength depends on the exponent values. (The extent to which the exponent exceeds the value 0.5 is regarded as a measure of how well the signal is correlated.) The weak correlation observed in the large timescale ($n > 10^4$) is due to the repetitive nature of the syllables and motifs present in the segment [27].

For the generalized fluctuation curves [Fig. 20(a)] for different q values ($q = -5, -2, 2, 5$), the results are similar to the results seen in the first segment. In particular, there are five regimes as seen in the case of the first segment, and the observations for each segment are similar (qualitatively) to those seen for the first segment. We note here that the basic time period (crossover positions) seen in both segments is of the same order.

The features for the scaling regions are the following:

(1) We have observed similar scaling properties in the detrended fluctuation of the amplitude envelope of segment 1 and segment 2 of the zebra finch song. The first and second crossover positions and scaling properties for segment 2 are similar to those seen for segment 1. We observe the difference in the large-scale ($n > 10^4$) scaling properties in these two segments. This is identified as region V in Fig. 20(a).

(2) The multifractal properties of the original and the IAAFT surrogate and shuffled data show similar results up to region IV as seen for segment 1. In region I the h_q spectrum of the shuffled in $q < 0$ region is not equal to 0.5, which is due to the very small scale correlations present in the

amplitude envelope after shuffling. Figures 17(c) and 17(d), and Figs. 20(c) and 20(d) show similar multifractal properties in their corresponding timescale of the amplitude envelope for both segments. Figure 17(e) and Fig. 20(e) both are in good agreement for the correlation structure of the large fluctuation part.

(3) We observe a difference in the multifractal and two-point correlation structures [Fig. 17(f)] for the longer timescales ($n > 10^4$). These timescales are identified as region V for both segments [Fig. 17(a) and Fig. 20(a)]. There are distinct differences in the h_q spectrums of the original data for both segments. In this timescale, the effect of the repetitive patterns of the syllables and motifs are prominent [27]. As a consequence of the repetitive nature of the motifs and the syllables, the correlation exponent is smaller. We note dissimilarities in the h_q vs q spectrum between the original data and the corresponding IAAFT data in both fluctuation parts ($q > 0$ and $q < 0$). IAAFT surrogate and original data should preserve the two-point correlation properties. A significant difference in the detrended fluctuation exponent ($q = 2$)

indicates the presence of phase correlation in the IAAFT surrogate data. The phase correlation in the IAAFT surrogate data may lead to a spurious conclusion regarding the nonlinear properties of the original time series. The presence of phase correlation has been shown in the IAAFT surrogate data in Ref. [33] in the context of AGN (active galactic nuclei) time series.

It is helpful to identify a common scaling region for all segments. For this, it is important to find the error bars associated with power-law fits for the q values considered. We note here that the scaling properties for low-fluctuation regions ($q < 0$) are poorer than those for the large fluctuation regions, as can be seen from the larger error bars in these regions [Fig. 17(d)]. Here we have observed that region I and region V as seen in Figs. 20(a) and 17(a) show larger error bars especially for the low-fluctuation parts ($q < 0$). This analysis is in the presence of intersyllabic gaps. Since this is a proof of principle analysis, we have not discussed the contribution of intersyllabic gaps to the multifractal properties in this context. We hope to include it in a separate study.

-
- [1] G. B. Mindlin and R. Laje, *The Physics of Birdsong* (Springer, Verlag, Berlin, Heidelberg, 2005).
- [2] T. Gardner, G. Cecchi, M. Magnasco, R. Laje, and G. B. Mindlin, *Phys. Rev. Lett.* **87**, 208101 (2001).
- [3] A. Amador and G. B. Mindlin, *Chaos* **18**, 043123 (2008).
- [4] R. Laje and G. B. Mindlin, *Phys. Rev. Lett.* **89**, 288102 (2002).
- [5] C. T. Herbert, S. Boari, G. B. Mindlin, and A. Amador, *Chaos* **30**, 053134 (2020).
- [6] S. Saar and P. P. Mitra, *PLoS ONE* **3**, e1461 (2008).
- [7] J. E. Markowitz, E. Ivie, L. Kligler, and T. J. Gardner, *PLoS Comput. Biol.* **9**, 1 (2013).
- [8] T. Sainburg, B. Theilman, M. Thielk, and T. Q. Gentner, *Nat. Commun.* **10**, 3636 (2019).
- [9] T. C. Roeske, D. Kelty-Stephen, and S. Wallot, *Sci. Rep.* **8**, 4570 (2018).
- [10] T. Schreiber and A. Schmitz, *Phys. Rev. Lett.* **77**, 635 (1996).
- [11] T. Schreiber and A. Schmitz, *Physica D* **142**, 346 (2000).
- [12] G. B. Mindlin, *Chaos* **27**, 092101 (2017).
- [13] T. W. Troyer, *Nature (London)* **495**, 56 (2013).
- [14] P. Marler and H. Slabbekoorn, *Nature's Music: The Science of Birdsong* (Elsevier, San Diego, 2004).
- [15] Laboratorio de Sistemas Dinámicos, Kernel description, <http://www.lsd.df.uba.ar/> (2020).
- [16] C. Jarne, A heuristic approach to obtain signal envelope with a simple software implementation, [arXiv:1703.06812](https://arxiv.org/abs/1703.06812).
- [17] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. Stanley, *Phys. A: Stat. Mech. Appl.* **316**, 87 (2002).
- [18] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, *Chaos* **5**, 82 (1995).
- [19] M. S. Movahed, G. R. Jafari, F. Ghasemi, S. Rahvar, and M. R. R. Tabar, *J. Stat. Mech.: Theory Exp.* (2006) P02003.
- [20] R. Hardstone, S.-S. Poil, G. Schiavone, R. Jansen, V. Nikulin, H. Mansvelder, and K. Linkenkaer-Hansen, *Front. Physiol.* **3**, 450 (2012).
- [21] P. C. Ivanov, M. G. Rosenblum, C.-K. Peng, J. Mietus, S. Havlin, H. E. Stanley, and A. L. Goldberger, *Nature (London)* **383**, 323 (1996).
- [22] P. C. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, *Nature (London)* **399**, 461 (1999).
- [23] J. L. Morales-Martínez, I. Segovia-Domínguez, I. Q. Rodríguez, F. A. Horta-Rangel, and G. Sosa-Gómez, *Phys. A: Stat. Mech. Appl.* **565**, 125611 (2021).
- [24] X. Zhang, G. Zhang, L. Qiu, B. Zhang, Y. Sun, Z. Gui, and Q. Zhang, *Water* **11**, 891 (2019).
- [25] P. Carpena, M. Gómez-Extremera, C. Carretero-Campos, P. Bernaola-Galván, and A. V. Coronado, *Entropy* **19**, 261 (2017).
- [26] L. R. Gorjão, G. Hassan, J. Kurths, and D. Witthaut, *Comput. Phys. Commun.* **273**, 108254 (2022).
- [27] J. Ludescher, M. I. Bogachev, J. W. Kantelhardt, A. Y. Schumann, and A. Bunde, *Phys. A: Stat. Mech. Appl.* **390**, 2480 (2011).
- [28] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, *Proc. Nat. Acad. Sci. U. S. A.* **105**, 4972 (2008).
- [29] M. Chutani, N. Rao, N. Nirmal Thyagu, and N. Gupte, *Chaos* **30**, 013109 (2020).
- [30] R. López-Ruiz, H. Mancini, and X. Calbet, *Phys. Lett. A* **209**, 321 (1995).
- [31] N. Stergiou, *Nonlinear Analysis for Human Movement Variability* (CRC Press, Boca Raton, FL, 2018).
- [32] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge University Press, Cambridge, 2003).
- [33] C. R ath, M. Gliozzi, I. E. Papadakis, and W. Brinkmann, *Phys. Rev. Lett.* **109**, 144101 (2012).
- [34] M. A. Goldin and G. B. Mindlin, *Chaos* **23**, 043138 (2013).
- [35] D. E. Knuth, *Art of Computer Programming, Vol. 2: Seminumerical Algorithms* (Addison-Wesley Professional, New York, 2014).
- [36] The Animal Around TheWorld, <https://www.youtube.com/watch?vKtJeHeD7k8>, 8th March 2020.
- [37] N. H. Prior, E. Smith, S. Lawson, G. F. Ball, and R. J. Dooling, *Sci. Rep.* **8**, 6212 (2018).