Journal of Comparative Physiology A (2024) 210:443–458
https://doi.org/10.1007/s00359-023-01664-4

RESEARCH

Supervised learning algorithm for analysis of communication signals in the weakly electric fish *Apteronotus leptorhynchus*

Dávid Lehotzky[1] · **Günther K. H. Zupanc**[1]

Received: 28 March 2023 / Revised: 22 July 2023 / Accepted: 25 July 2023 / Published online: 13 September 2023

Abstract

Signal analysis plays a preeminent role in neuroethological research. Traditionally, signal identification has been based on pre-defined signal (sub-)types, thus being subject to the investigator's bias. To address this deficiency, we have developed a supervised learning algorithm for the detection of subtypes of chirps—frequency/amplitude modulations of the electric organ discharge that are generated predominantly during electric interactions of individuals of the weakly electric fish *Apteronotus leptorhynchus*. This machine learning paradigm can learn, from a 'ground truth' data set, a function that assigns proper outputs (here: time instances of chirps and associated chirp types) to inputs (here: time-series frequency and amplitude data). By employing this artificial intelligence approach, we have validated previous classifications of chirps into different types and shown that further differentiation into subtypes is possible. This demonstration of its superiority compared to traditional methods might serve as proof-of-principle of the suitability of the supervised machine learning paradigm for a broad range of signals to be analyzed in neuroethology.

Keywords Signal analysis · Artificial intelligence · Supervised learning · Chirping behavior · Weakly electric fish · *Apteronotus leptorhynchus*

Introduction

Signals as vehicles for transmission of information from a sender to a receiver play a pivotal role in animal communication (Bradbury and Vehrencamp 2011). Broadcasting of signals is mediated by a variety of sensory channels, such as visual, acoustic, tactile, chemical, and electric. Diversity of signals, either within one sensory modality or by activation of several sensory channels, enables animals to use different signals for different behavioral functions. Within one sensory modality, signal diversity is often achieved by modulation of a generic type of signal. For example, different acoustic signals can be produced by temporal frequency and amplitude modulations, and even rather subtle differences can have profoundly different functional effects (Schwartz et al. 2007; Feng et al. 2009; Hechavarría et al. 2020).

While acoustic signals are displayed intermittently only (although sometimes for prolonged periods at high rates), some electric fishes produce a generic form of electric signal continuously throughout life. This group includes the brown ghost knifefish (*Apteronotus leptorhynchus*), a species of the taxonomic order Gymnotiformes that has been intensively studied as model organisms in ethology and neuroethology.

Apteronotus leptorhynchus generates such continuous electric discharges with its electric organ composed of modified axonal terminals of spinal motoneurons (for review see Zupanc and Bullock 2005). The synchronous depolarization of these so-called electrocytes produces electric pulses separated by short inter-pulse intervals. This results in the appearance of a continuous, wave-like signal, commonly referred to as electric organ discharge (EOD). The frequency at which the fish generates the EOD train is determined, in a one-to-one fashion, by the frequency of the neural oscillations of a central pattern generator in the medulla oblongata, the pacemaker nucleus. Within the species-specific frequency range of 650–1000 Hz, males discharge at higher

Handling Editor: Wolfgang Rössler.

Günther K. H. Zupanc
g.zupanc@northeastern.edu

Dávid Lehotzky
d.lehotzky@northeastern.edu

[1] Laboratory of Neurobiology, Department of Biology, Northeastern University, 360 Huntington Ave, Boston, MA 02115, USA

frequencies than females, with little overlap between the sexes (Meyer et al. 1987; Zupanc et al. 2014). Owing to this sexual dimorphism, the EOD contains information about the sex of its sender.

Whereas the species as whole occupies a broad EOD frequency range, the frequency of the discharges of a given individual within this frequency band is highly constant, as indicated by the coefficient of variation [cv = (standard deviation / mean) × 100 (%)], which assumes values of less than 0.2% over 30-min (Eske et al. 2023). Nevertheless, transient modulations may occur, resulting in diversification of the generic EOD signal. The best-characterized type comprises chirps. In isolated individuals of *A. leptorhynchus*, chirps are very rarely produced, on average less than once per 10 min (Engler et al. 2000; Zupanc et al. 2001; Eske et al. 2023). However, during stimulation with the EODs of conspecific fish or with electric signals mimicking such EODs, or after administration of certain drugs, chirp production may increase one-thousand-fold to rates as high as 2 s^{-1} (Zupanc and Maler 1993; Engler and Zupanc 2001; Eske et al. 2023).

Chirps last between some tens and a few hundred milliseconds and involve complex frequency and amplitude modulations. Six distinct chirp types have been identified (Engler et al. 2000; Zupanc et al. 2006). They are defined by differences in duration, extent of the frequency and amplitude modulations, as well as additional features, such as the presence or absence of an undershoot before the frequency returns to baseline levels as evident in time-frequency plots. The usefulness of these features for differentiating different chirp types has been shown in several other studies (Ho et al. 2013a, b; Turner et al. 2007; Oboti et al. 2023). Most notably, by employing this approach, a comparative analysis revealed an enormous diversity of chirp signals in 13 species of apteronotids, which included not only variation across species but also between congeners and populations of the same species (Turner et al. 2007).

In *A. leptorhynchus*, spontaneously produced chirps are predominantly of type 1, whereas most chirps evoked by the EODs of a neighboring fish (or mimics of such electric signals) or by proper pharmacological stimulation belong to the type 2 category (Engler et al. 2000; Zupanc et al. 2006; Eske et al. 2023). Both type 1 and type 2 chirps are rather short (duration approximately 20 ms) but distinct in terms of the degree of frequency increase (400 Hz versus 100 Hz) and amplitude reduction (approximately 50% versus <10%). Longer chirps of type 3–6 are, most typically, generated by older individuals and directed to fish of the other sex.

While chirps can be elicited from either sex, at similar rates, through application of pharmacological agents (Eske et al. 2023), during electric interaction with conspecifics or in response to electric stimuli mimicking a fish's EOD males chirp at much higher rates than females (Zupanc and Maler 1993; Dulka and Maler 1994; Dunlap et al. 1998; Dunlap 2002; Triefenbach and Zakon 2003; Hupé and Lewis 2008). In addition, chirps are optimally evoked by electric stimuli with frequencies within ±10 Hz of the fish's EOD frequency (Engler and Zupanc 2001). Thus, type 2 chirps are typically exchanged by males. Moreover, the chirps produced by two electrically interacting fish are not independent of each other (Zupanc et al. 2006). Instead, the chirps generated by one fish follow the chirps of the other individual with a preferred latency of roughly 500–1000 ms (Zupanc et al. 2006). This 'echo response' may serve a communicatory function during social interactions, such as aggressive encounters.

Traditionally, different chirp types have been identified and quantified by visual inspection of time–voltage and time–frequency plots (e.g., Engler et al. 2000; Engler and Zupanc 2001; Zupanc et al. 2001; Dunlap and Larkins-Ford 2003; Zupanc et al. 2006; Kolodziejski et al. 2007; Hupé and Lewis 2008; Smith and Combs 2008; Dunlap et al. 2011; Gama Salgado and Zupanc 2011; Neeley et al. 2018). In addition, threshold-based algorithms (Bastian et al. 2001; Aumentado-Armstrong et al. 2015; Henninger et al. 2018; Allen and Marsat 2019; Field et al. 2019) and a method based on assumed chirp waveform (Eske et al. 2023) have been used for chirp detection. Whereas these approaches can be successfully employed for the identification of predefined chirp types, the definition of chirp categories is subject to the investigator's bias. Moreover, such approaches do not allow detection of possible additional chirp types that remained unnoticed previously.

To address these deficiencies, we have, in the present study, developed a supervised learning algorithm. Supervised learning is a machine learning paradigm (Bishop 2006) used across many disciplines. Its goal is to learn, from a "ground truth" (GT) data set, a function that assigns proper outputs (in the present study: time instances of chirps and associated chirp types) to inputs (in the present study: time-series frequency and amplitude data). While we demonstrate the suitability of this machine learning paradigm for the unbiased analysis of chirps produced by *A. leptorhynchus*, we propose that similar approaches can be successfully applied to signal analysis in a variety of other ethological and neuroethological systems.

Materials and methods

EOD recording

For the present investigation, time–voltage recordings of the EOD containing chirps generated spontaneously or evoked pharmacologically were analyzed. These data had been collected as part of a previous study examining the effect of urethane anesthesia on EOD frequency and chirping behavior in *A. leptorhynchus* (Eske et al. 2023).

Eight fish (total lengths: median, 116 mm; range 107–143 mm; body weights: median, 2.9 g; range 2.5–4.8 g) were used. Their EOD baseline frequencies varied between 683 Hz and 868 Hz (normalized to frequency values expected at 26 °C, using a Q_{10} of 1.56). The morphological data and EOD frequencies indicate that the fish were approximately 1 year old and included both males and females (Ilieş et al. 2014; Zupanc et al. 2014).

Details of the experiments and the recording technique are given in Eske et al. (2023). Briefly, each fish was kept in an isolation tank in which a cylindrical plastic tube provided shelter. Differential recording of the fish's EOD was done through a pair of stainless-steel electrodes mounted on the inside of the tube. During recording, the two open ends of the tube were closed with a coarse plastic mesh netting to ensure that the fish did not leave the tube.

The EOD of each fish was recorded for 30 min before, and 180 min immediately after, general anesthesia. State of anesthesia was induced by transferring the fish into a glass beaker containing 2.5% urethane dissolved in water from the fish's isolation tank. During the pre-anesthesia session, spontaneous chirps occurred but at very low rates of approximately 1 chirp/30 min. Anesthesia induced a tremendous increase in chirping behavior, resulting, on average, in 1500 chirps during the 30 min immediately following anesthesia.

For the present analysis, the 30-min-pre-anesthesia recordings, and the 180-min-post-anesthesia recordings, of the 8 fish were combined, yielding a total of 1680 min of EOD recording. Employing the supervised learning algorithm, a total of 30,734 chirps were detected in these combined recordings.

Calculation of EOD frequency and amplitude

The sampled voltage data $\left(t_i, v_i\right)$, $i = 1, \dots, M_{\mathrm{v}}$, were exported from Spike 2 and processed in MATLAB version R2021b. These data were filtered in 3-s windows with 2-s overlap using a bandpass filter with frequency band $[0.5, 1.5] \times f_0$, where the fundamental frequency f_0 in each 3-s window was determined based on the power spectrum of the signal using fast Fourier transform and the "findpeaks" function of MATLAB.

Based on the zero-crossings of the filtered signal, we then computed the time, frequency, and amplitude values $\left(T_k, f_k, A_k\right)$ associated with each $k = 1, \dots, M$, oscillation interval (for details, see Appendix A). An example of computed time-series data of frequency and amplitude is shown in Fig. 1.

Chirp detection by supervised learning

"Ground Truth" data set

Data collection

Tuples of equal-time-length time-series data segments

$$\mathbf{S}_{n_s(r-1)+j} = \left(\left\{T_k^{(r)}, f_k^{(r)}, A_k^{(r)}\right\} : T_k^{(r)} \in \left[T_{\text{start}} + (j-1)\Delta T, T_{\text{start}} + j\Delta T\right],\right.$$
$$\left. k = 1, \dots, M-1\right), \quad j = 1, \dots, n_s, \tag{1}$$

were collected from each recording $r = 1, \dots, n_{\mathrm{r}}$, where n_{r} is the total number of EOD recordings, and superscript $\square^{(r)}$ indicates association with recording r. The time length of segments was determined as $\Delta T = \left(T_{\text{end}} - T_{\text{start}}\right)/n_{\mathrm{s}}$. The values of parameters $T_{\text{start}}, T_{\text{end}}, n_{\mathrm{s}}, n_{\mathrm{r}}$, used for the generation of time-series data segments are provided in Table 1.

Using the MATLAB tool shown in Fig. 2, a person previously trained to identify chirps collected all chirp instances from each segment $\mathbf{S}_i$ for all indices $i \in \mathbf{i}_{\mathrm{GT}}$, where the elements of subset $\mathbf{i}_{\mathrm{GT}} \subset \left\{1, \dots, n_{\mathrm{s}} n_{\mathrm{r}}\right\}$, with $n_{\mathrm{GT}} = \left|\mathbf{i}_{\mathrm{GT}}\right|$ (see Table 1), were randomly chosen, without replacement.

Although for each data point only time and frequency values were displayed during data collection (see Fig. 2), the associated amplitude values were also stored in the GT set of chirps

$$\mathbf{G} = \left\{ \left(\left\{ T_{i,j}, f_{i,j}, A_{i,j} \right\} \right)_{j=1}^{l_i} \right\}_{i=1}^{n}, \tag{2}$$

where $\left\{T_{i,j}, f_{i,j}, A_{i,j}\right\}$ is the j-th data point of the i-th GT chirp sample, l_i denotes the number of data points in the i-th sample, and n is the total number of samples.

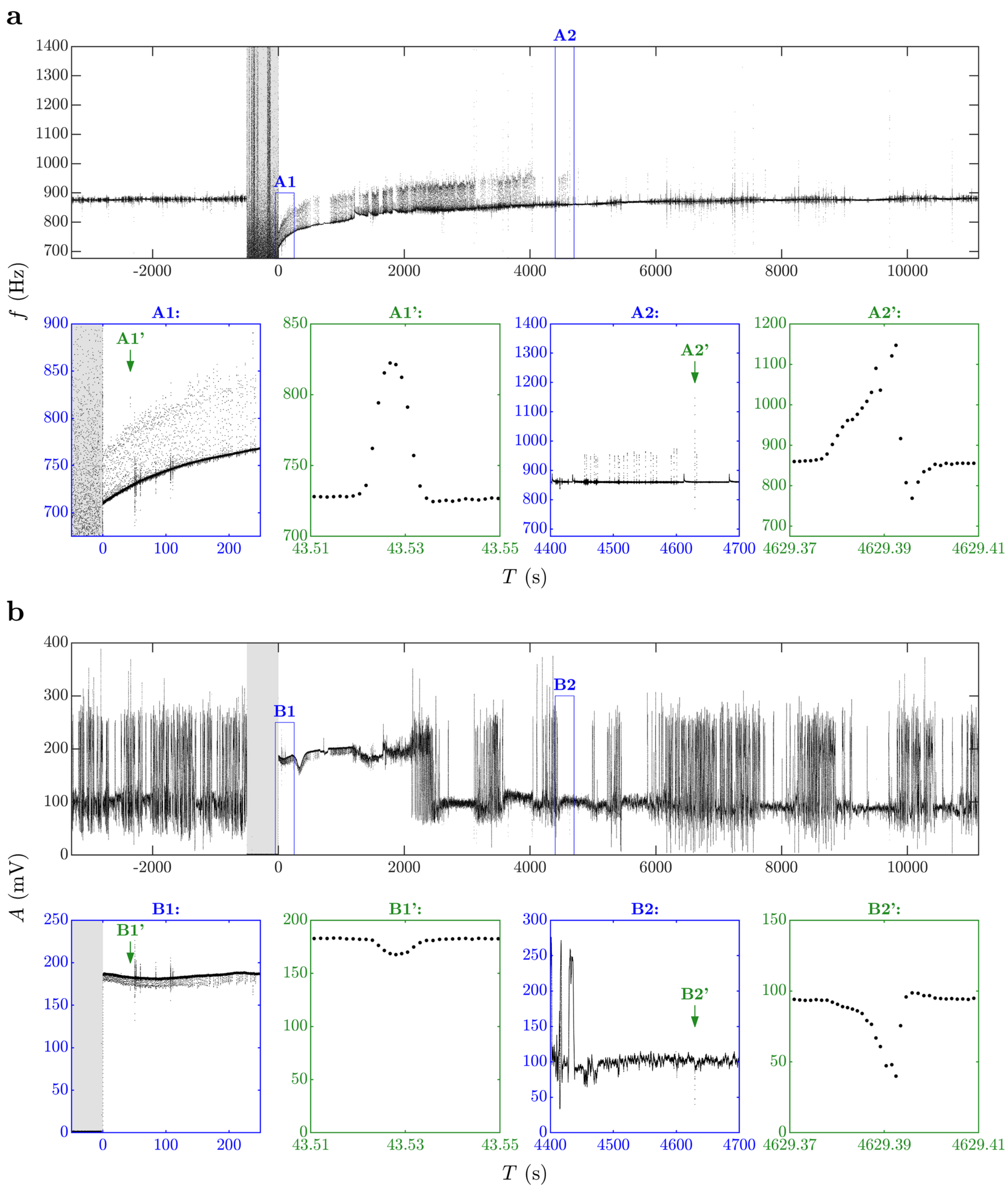
Data processing

The person who collected chirp samples was instructed to include, in each sample, data points prior to and after chirping, associated with the non-modulated, instantaneous "base" frequency of the fish. Hence, we assumed that each sample includes both pre and post-chirp data points and estimated the "base" frequency and amplitude of each sample i as

$$f_{\text{base},i} = \operatorname{median}\left(\left\{f_{i,j}\right\}_{j=1}^{n_{\text{med}}}, \left\{f_{i,l_i-j+1}\right\}_{j=1}^{n_{\text{med}}} \right), \tag{3}$$

$$A_{\text{base},i} = \operatorname{median}\left(\left\{A_{i,j}\right\}_{j=1}^{n_{\text{med}}}, \left\{A_{i,l_i-j+1}\right\}_{j=1}^{n_{\text{med}}} \right), \tag{4}$$

where $n_{\text{med}} < \min_i(l_i/2)$ is an arbitrarily chosen positive integer which we set to $n_{\text{med}} = 10$. We normalized each sample $i = 1, \dots, n$ with respect to the maximum frequency rise according to



$$\varphi_{i,j} = \frac{f_{i,j} - f_{\mathrm{base},i}}{\max\limits_{j \in \{1,\ldots,l_i\}} \left(f_{i,j}\right) - f_{\mathrm{base},i}}, \quad j = 1, \ldots, l_i, \tag{5}$$

and with respect to the base amplitude as

$$a_{i,j} = \frac{A_{i,j} - A_{\mathrm{base},i}}{A_{\mathrm{base},i}}, \quad j = 1, \ldots, l_i. \tag{6}$$

Then, we centered the time values of each sample according to

◂**Fig. 1** EOD frequency f (**a**) and amplitude A (**b**) with respect to time T in a recording involving urethane anesthesia (for details of computation see Sect. "Calculation of EOD frequency and amplitude"). After baseline recording, the tube with the fish was transferred from the home tank to a glass beaker containing 2.5% urethane solution dissolved in aquarium water. As soon as the fish stopped undulating its anal fin and moving its opercula, it was returned to the home tank (arbitrarily defined as time point $T = 0$). The gray bar indicates the time during which the fish was exposed to the anesthetic. Changes in the orientation and position of the fish relative to the recording electrodes result in noisy amplitude signals (pre-anesthesia, and $T >\sim 2000$ s as shown in **b**). The reduction of noise immediately after anesthesia is related to the ceased movement of the fish. Note onset of type 2 chirping at higher rates immediately after anesthesia (**a**/A1, **b**/B1) that persists to approximately $T = 4600$ s after exposure to the anesthetic (**a**/A2, **b**/B2). The recorded signal contains both type 2 (**a**/A1', **b**/B1') and type 1 (**a**/A2', **b**/B2') chirps. The latter is characterized by large rise and negative undershoot in frequency (**a**/A2'), as well as a large drop in amplitude (**b**/B2'). By contrast, the former is characterized by a smaller rise without undershoot in frequency (**a**/A1') and a smaller reduction in amplitude (**b**/B1')

$$\tilde{T}_{i,j} := T_{i,j} - T_{i,j_{\mathrm{cen},i}}, \quad j = 1, \dots, l_i, \tag{7}$$

$$j_{\mathrm{cen},i} = \underset{k}{\operatorname{argmin}}\left(\left|H_{i,k} - \frac{1}{2}\right|\right), \tag{8}$$

$$H_{i,k} = \frac{\sum_{j=1}^{k} h\left(\varphi_{i,j}\right)}{\sum_{j=1}^{l_i} h\left(\varphi_{i,j}\right)}, \quad k = 1, \dots, l_i, \tag{9}$$

where rectifier

$$h\left(\varphi_{i,j}\right) = \frac{\ln\left(1 + e^{\delta(|\varphi_{i,j}| - \bar{\varphi}_i)}\right)}{\delta - \delta\bar{\varphi}_i}, \tag{10}$$

with

$$\bar{\varphi}_i = 4\max\left(\mathrm{sd}\left(\left\{\varphi_{i,k}\right\}_{k=1}^{n_{\mathrm{med}}}\right), \mathrm{sd}\left(\left\{\varphi_{i,l_i-k+1}\right\}_{k=1}^{n_{\mathrm{med}}}\right)\right), \tag{11}$$

was applied for the elimination of noise and to highlight "meaningful" parts of the frequency sample. Here $\mathrm{sd}(\cdot)$ denotes the standard deviation, $\bar{\varphi}_i$ is the cutoff value of normalized frequency associated with sample i and $\delta = 50$ is an arbitrarily chosen smoothing parameter.

Using the empirical cumulative distribution $H_{i,\cdot}$ of rectified frequency values $h\left(\varphi_{i,\cdot}\right)$, we trimmed each sample, such that only the data points j within interval $\tilde{T}_{i,j} \in \left[-3\Delta\tilde{T}_i, 3\Delta\tilde{T}_i\right]$ were kept, with

$$\Delta\tilde{T}_i = \tilde{T}_{i,j_i^+} - \tilde{T}_{i,j_i^-}, \tag{12}$$

$$\begin{aligned} j_i^+ &= \underset{k\in\{1,\dots,l_i\}}{\operatorname{argmin}}\left(\left|H_{i,k} - 0.9\right|\right), \\ j_i^- &= \underset{k\in\{1,\dots,l_i\}}{\operatorname{argmin}}\left(\left|H_{i,k} - 0.1\right|\right). \end{aligned} \tag{13}$$

Note that here $\Delta\tilde{T}_i$ is the difference between the 90% and 10% percentile estimates of the empirical cumulative distribution $H_{i,\cdot}$. The above described data processing method is illustrated in Fig. 3.

Grouping and resampling

Because our supervised learning method requires uniform size among GT samples, we grouped and resampled all GT samples according to the number of data points that formed the individual GT samples.

After trimming, the size of each GT sample was roughly commensurate with the length of the associated chirp. To distinguish between chirps whose duration have different time scales, we divided GT samples into three groups and resampled the members of each r group such that associated samples contained $10^r + 1$ number of points:

$$\begin{aligned} \mathbf{G}_r = \Big\{ & \left\{\left(T_{i,j_{\mathrm{cen},i}+j}, f_{i,j_{\mathrm{cen},i}+j}, A_{i,j_{\mathrm{cen},i}+j}\right)\right\}_{j=-10^r/2}^{10^r/2} : \\ & \left|\left\{\tilde{T}_{i,j} : \tilde{T}_{i,j} \in \left[-3\Delta\tilde{T}_i, 3\Delta\tilde{T}_i\right], 1 \le j \le l_i\right\}\right| \in \left(10^{r-1}+1, 10^r+1\right], \\ & 1 \le i \le n\Big\}, \quad r = 1, 2, 3. \end{aligned} \tag{14}$$

Here we utilized the fact that all data points inside any GT sample can be located within the associated recording's time-frequency-amplitude data. For example, if we know that $T_{i,1}$ and T_q are from the same recording and that $T_{i,1} = T_q$, then we can find any other point j associated with sample i: $\left(T_{i,j}, f_{i,j}, A_{i,j}\right) = \left(T_{q+j-1}, f_{q+j-1}, A_{q+j-1}\right)$.

Note that chirps typically have a duration shorter than 0.5 s, and the highest EOD frequency in *A. leptorhynchus* is approximately 1000 Hz, therefore GT sample groups $\mathbf{G}_r$, $r = 1, 2, 3$, are able to capture the full length of all chirps.

Training

Principal component analysis

After resampling, we recomputed, according to Eqs. 3–6, the normalized frequencies and amplitudes $\left(\varphi_{i,j_{\mathrm{cen},i}+j}, a_{i,j_{\mathrm{cen},i}+j}\right), j = -10^r/2, \dots, 10^r/2$, of each chirp sample i in each GT group $\mathbf{G}_r$. For ease of notation, in the following, we drop the shift $j_{\mathrm{cen},i}$ in the second subscript index.

For each r, we collected from $\mathbf{G}_r$ the normalized frequency and amplitude values

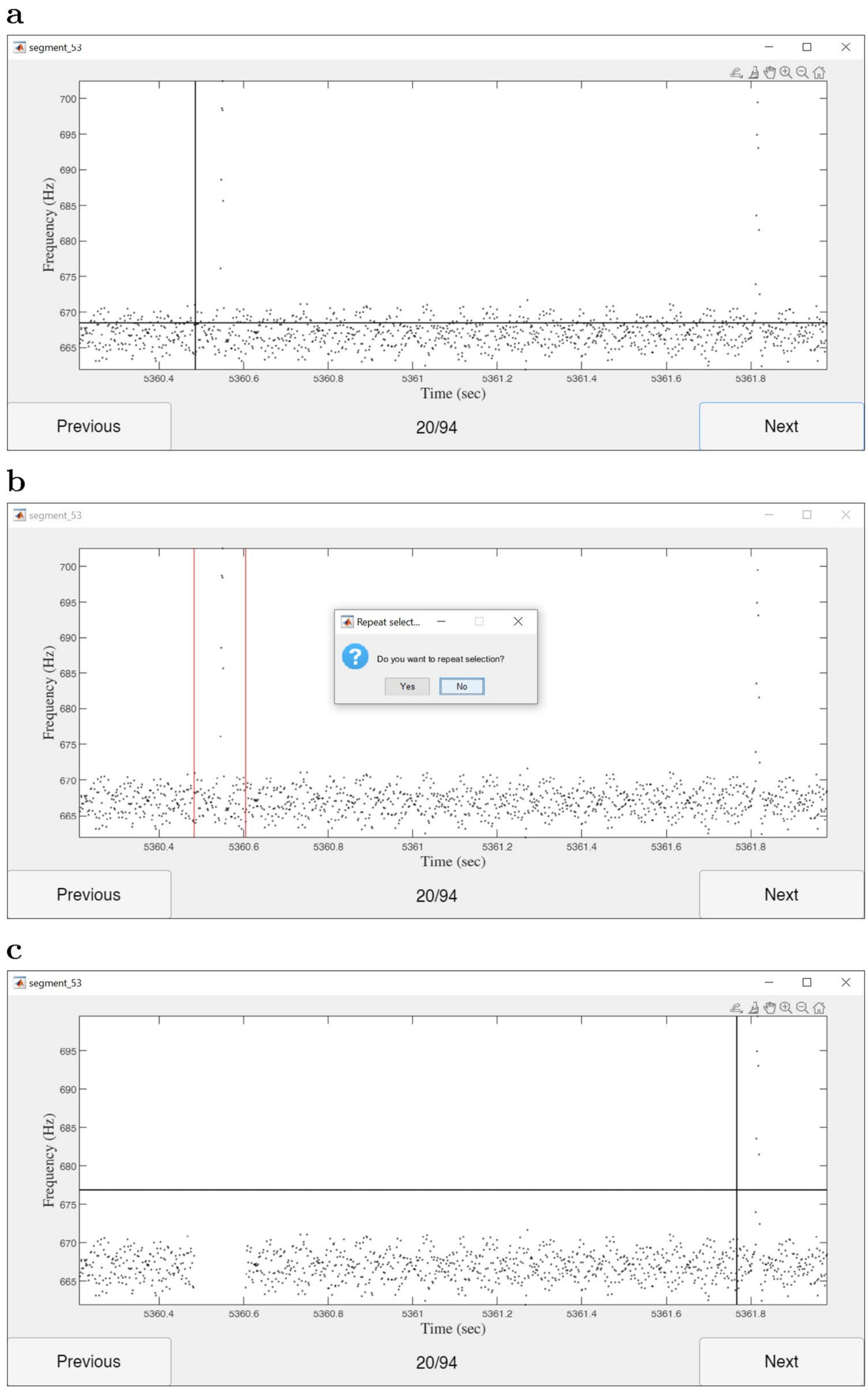


Fig. 2 Matlab tool built for collecting chirp samples from time-series frequency data (black dots). The user can select data points associated with a chirp by moving the cursor (intersection of *black lines* in **a** and **c**) to the two end points of the time interval delimiting the chirp instance. After selecting the time interval (red lines in **b**), the user must confirm the current selection before proceeding to collect further data points (see dialog box in **b**). Following the confirmation of the selection, data points associated with the selected time interval are stored and removed from the displayed data set (**c**). Once all displayed chirp instances have been collected, the user can move to the next (or previous), overlapping, time segment to collect the remaining chirp data points from the time-series frequency data segment

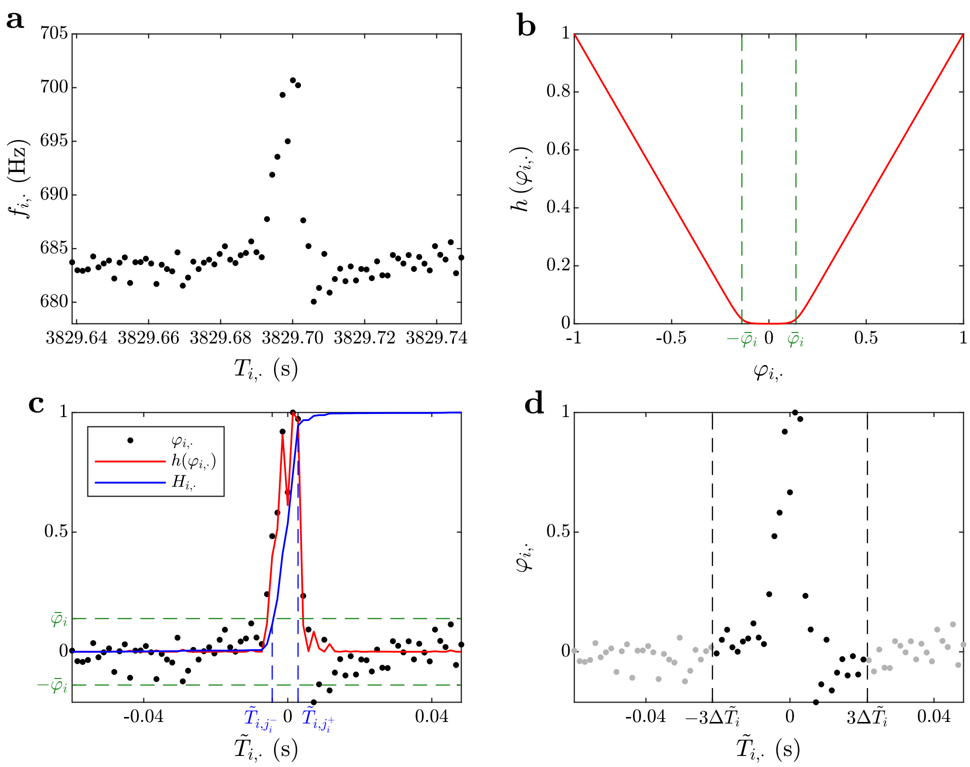


Fig. 3 Processing of "ground truth" samples (see Sect. ""Ground Truth" data set"). Data points $\{(T_{i,j}, f_{i,j})\}_{j=1}^{l_i}$ of the i-th sample are plotted in **a** as *black dots*. The frequency values $\{f_{i,j}\}_{j=1}^{l_i}$ are normalized according to Eq. 5 and passed through the rectifier function (red curve) displayed in **b**. The green dashed lines in **b** and **c** display the cutoff value $\bar{\varphi}_i$ of the rectifier function. The centered and normalized data points $\{(\tilde{T}_{i,j}, \varphi_{i,j})\}_{j=1}^{l_i}$ of the i-th sample (see Eqs. 5–11) are displayed in **c** as black dots together with the rectified normalized frequencies (red curve) and their empirical cumulative distribution (blue curve). The 10% and 90% percentile estimates (blue, dashed lines in **c**) of this cumulative distribution determine the time width of the sample: $\Delta\tilde{T}_i = \tilde{T}_{i,j_i^+} - \tilde{T}_{i,j_i^-}$. The sample is trimmed based on this time width (**d**) such that data points outside interval $\tilde{T}_{i,j} \in [-3\Delta\tilde{T}_i, 3\Delta\tilde{T}_i]$ (delimited by black, dashed lines and marked by *gray dots*) are eliminated

$$\mathbf{f}_r^{(i)} = \left[\varphi_{i,-10^r/2}, \dots, \varphi_{i,10^r/2},\right]^{\mathrm{T}}, \tag{15}$$

$$\mathbf{a}_r^{(i)} = \left[a_{i,-10^r/2}, \dots, a_{i,10^r/2}\right]^{\mathrm{T}}, \tag{16}$$

of each sample i associated with the training set (for details about the training set, see Sect. "Cross-validation") into a matrix $\mathbf{X}_r \in \mathbb{R}^{m_r \times 2(10^r+1)}$ such that

$$\mathbf{X}_r^{\mathrm{T}} = \begin{bmatrix} \mathbf{f}_r^{(1)} & \cdots & \mathbf{f}_r^{(m_r)} \\ \mathbf{a}_r^{(1)} & \cdots & \mathbf{a}_r^{(m_r)} \end{bmatrix}, \tag{17}$$

where m_r is the total number of samples in $\mathbf{G}_r$ associated with the training set. For the further ease of notation, in the following, we drop index r, as well.

We determined the principal components (PCs) $\mathbf{p}_1, \dots, \mathbf{p}_{2(10^r+1)}$, of $\mathbf{X}$ by performing the spectral decomposition of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$. Then we projected the training data set onto the space of the first N PCs, i.e., we computed

$$\mathbf{Y} = \mathbf{X}\mathbf{P}_N, \tag{18}$$

where $\mathbf{P}_N = [\mathbf{p}_1, \dots, \mathbf{p}_N]$.

Gaussian mixture model fitting

We modeled the projected data $\mathbf{Y}^{\mathrm{T}} = \left[\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(m)}\right]$ using the Gaussian mixture model (GMM)

$$\mathbf{y}^{(i)} \sim \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c), \quad c \sim M_C(p_1, \dots, p_C), \tag{19}$$

where $\mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)$ is the multivariate normal distribution of the c-th mixture component with mean $\boldsymbol{\mu}_c \in \mathbb{R}^{N\times 1}$ and covariance $\boldsymbol{\Sigma}_c \in \mathbb{R}^{N\times N}$, while $M_C(p_1, \dots, p_C)$ is a multinomial distribution with C number of categories and mixing proportions $p_1, \dots, p_C$. We estimated the unknown parameters

Table 1 Parameter values used for generating time-series data segments $\mathbf{S}_i$, $i \in \mathbf{i}_{GT}$, from which "ground truth" chirp samples $\mathbf{G}$ were collected

Parameter (unit)	Value
T_{start} (s)	3200
T_{end} (s)	6200
n_s (1)	10
n_r (1)	16
n_{GT} (1)	64

Values of T_{start} and T_{end} were measured in time coordinates fixed to the start of recordings

$\boldsymbol{\Theta} = \{p_1, \dots, p_C, \boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_C, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_C\}$ of this GMM based on data $\mathbf{Y}$ using the "fitgmdist" function of MATLAB.

Elimination of outliers

After fitting the GMM, we assigned each data sample i to the cluster with maximum posterior probability, i.e., we computed the cluster of sample i according to

$$c_i = \underset{c \in \{1,\dots,C\}}{\operatorname{argmax}}(P(c|i)), \tag{20}$$

for each $i = 1, \dots, m$, where $P(c|i)$ is the probability that sample i belongs to cluster c, given the observation $\mathbf{y}^{(i)}$. Then, we computed the coefficient of determination (CoD) of the frequency component of each sample with respect to its assigned cluster mean as

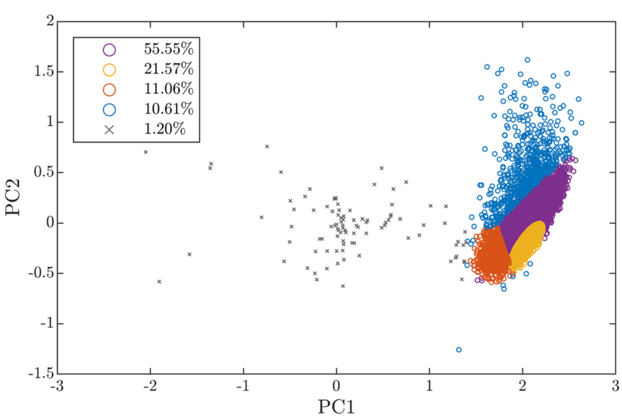


Fig. 4 Training data set projected to the space spanned by the first two principal components (PC1 and PC2). Circles with different colors correspond to clusters identified by the algorithm. Gray crosses correspond to samples in an eliminated cluster. The percentage-wise size of kept (circles) and eliminated (crosses) clusters is indicated at the top left corner, relative to the size of the training set

$$R_i^2 = 1 - \frac{\left\| \mathbf{f}^{(i)} - \bar{\mathbf{f}}_{c_i} \right\|^2}{\left\| \mathbf{f}^{(i)} - \bar{\mathbf{f}}^{(i)} \right\|^2}. \tag{21}$$

Here $\|\cdot\|$ denotes the L2 norm and

$$\left[\bar{\mathbf{f}}_c, \bar{\mathbf{a}}_c\right]^{\mathrm{T}} = \mathbf{P}_N \hat{\boldsymbol{\mu}}_c, \tag{22}$$

$$\bar{\mathbf{f}}^{(i)} = \frac{1}{10^r + 1}\left(\mathbf{1}^{\mathrm{T}}\mathbf{f}^{(i)}\right)\mathbf{1}, \tag{23}$$

with $\mathbf{1}$ being a vector of 1-s.

We eliminated each cluster c for which the 5% percentile of associated CoD values $\{R_i^2 : c_i = c, 1 \le i \le m\}$ was below threshold $\delta_{R^2} = 0.3$. Additionally, we eliminated each cluster c whose size $\left|\{i : c_i = c, 1 \le i \le m\}\right|$ was below threshold $\delta_c = 30$.

Figure 4 illustrates the projected training data $\mathbf{Y}$ from $\mathbf{G}_2$, with parameters $N = 2$ and $C = 5$; note the eliminated cluster.

Detection

Training yields PCs $\mathbf{P}_N$ and GMM

$$\mathbf{y}^{(i)} \sim \mathcal{N}\left(\hat{\boldsymbol{\mu}}_c, \hat{\boldsymbol{\Sigma}}_c\right), \quad c \sim M_{C^*}\left(\tilde{p}_1, \dots, \tilde{p}_{C^*}\right), \tag{24}$$

where $C^* \le C$ is the number of kept clusters, with $\tilde{p}_c = \hat{p}_c / \sum_{q=1}^{C^*} \hat{p}_q$, and $\hat{p}_c, \hat{\boldsymbol{\mu}}_c, \hat{\boldsymbol{\Sigma}}_c$, being the estimated parameters of kept clusters $c = 1, \dots, C^*$.

To detect chirps in recordings, we analyzed data points $\left\{\left(T_{i+j-1}, f_{i+j-1}, A_{i+j-1}\right)\right\}_{j=1}^{10^r+1}, i = 1, \dots, M - 10^r$ in a moving time window containing $10^r + 1$ samples (see Fig. 5a). At each instance i, we computed normalized frequency and amplitude values

$$\mathbf{f}^{(i)} = \left[\varphi_{i,1}, \dots, \varphi_{i,10^r+1},\right]^{\mathrm{T}}, \tag{25}$$

$$\mathbf{a}^{(i)} = \left[a_{i,1}, \dots, a_{i,10^r+1}\right]^{\mathrm{T}}, \tag{26}$$

according to formulas Eqs. 3–6 with $\left(T_{i,j}, f_{i,j}, A_{i,j}\right) = \left(T_{i+j-1}, f_{i+j-1}, A_{i+j-1}\right)$ and $l_i = 10^r + 1$.

Mahalanobis-distance-based detection

At each instance i, our Mahalanobis-distance-based (MDB) detection method first projects the normalized frequency and amplitude data onto the PCs according to

$$\mathbf{y}^{(i)} = \mathbf{P}_N^{\mathrm{T}} \begin{bmatrix} \mathbf{f}^{(i)} \\ \mathbf{a}^{(i)} \end{bmatrix}, \tag{27}$$

then it determines the kept cluster which is most likely to generate $\mathbf{y}^{(i)}$:

$$c_i = \underset{j \in \{1,\dots,C^*\}}{\operatorname{argmax}} (P(j|i)). \tag{28}$$

Afterward, our method computes the Mahalanobis distance

$$d_i = \sqrt{\left(\mathbf{y}^{(i)} - \hat{\boldsymbol{\mu}}_{c_i}\right)^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}_{c_i}^{-1} \left(\mathbf{y}^{(i)} - \hat{\boldsymbol{\mu}}_{c_i}\right)}. \tag{29}$$

For any point generated by kept cluster c_i, realizations d_i^2 follow a chi-squared distribution with N degrees of freedom: $D_i^2 \sim \chi_N^2$.

The MDB method collects all i instances, where the squared Mahalanobis distance is below threshold ε_{d^2} and the maximum frequency rise is above threshold ε_f, into the tuple

$$\mathbf{c}_{\mathrm{MDB}} = \Big(i : d_i^2 < \varepsilon_{d^2}, \max_{1 \le j \le 10^r+1} \left(f_{i,j}\right) - f_{\mathrm{base},i} > \varepsilon_f, \\ i = 1, \dots, M - 10^r - 1 \Big). \tag{30}$$

Each contiguous segment in $\mathbf{c}_{\mathrm{MDB}}$ corresponds to an identified chirp. In each contiguous segment, we associate the identified chirp with the instance i that has lowest distance d_i. Threshold ε_{d^2} is determined based on a chosen level of significance α such that $P\left(D_i^2 < \varepsilon_{d^2}\right) = 1 - \alpha$. The MDB method is illustrated in Fig. 5b.

Coefficient-of-determination-based detection

At each instance i, our coefficient-of-determination-based (CDB) detection method computes the CoD of the frequency component with respect to each kept cluster mean according to

$$R_{i,c}^2 = 1 - \frac{\left\| \mathbf{f}^{(i)} - \bar{\mathbf{f}}_c \right\|^2}{\left\| \mathbf{f}^{(i)} - \bar{\mathbf{f}}^{(i)} \right\|^2}, \quad c = 1, \dots, C^*, \tag{31}$$

using formulae Eqs. 22–23, and assigns instance i to the cluster with highest CoD value:

$$c_i = \underset{c \in \{1,\dots,C^*\}}{\operatorname{argmax}} \left(R_{i,c}^2\right). \tag{32}$$

Afterward, the CDB method collects all instances into the tuple $\mathbf{c}_{\mathrm{CDB}}$ where the CoD value and the maximum frequency rise are both above thresholds ε_{R^2} and ε_f, respectively:

$$\mathbf{c}_{\mathrm{CDB}} = \Big(i : R_{i,c_i}^2 > \varepsilon_{R^2}, \max_{1 \le j \le 10^r+1} \left(f_{i,j}\right) - f_{\mathrm{base},i} > \varepsilon_f, \\ i = 1, \dots, M - 10^r - 1 \Big). \tag{33}$$

Similarly to the MDB method, identified chirps are associated with contiguous segments in $\mathbf{c}_{\mathrm{CDB}}$. In each contiguous segment, the identified chirp is assigned to the instance i that has the highest R_{i,c_i}^2 value. The CDB method is illustrated in Fig. 5c.

Chirp detection based on assumed chirp waveform

In order to assess the performance of the two algorithms detailed above, we chose, as a reference, the time-frequency-shape-based (TFSB) chirp detection algorithm described in (Eske et al. 2023). This algorithm is based on the chirp waveform function

$$\varphi\left(\tilde{T};\tilde{\alpha}\right) = \frac{2e^{\tilde{\alpha}\tilde{T}}}{1 + e^{2\tilde{\alpha}\tilde{T}}}, \tag{34}$$

which is assumed to characterize, during chirps, the normalized frequency φ with respect to time $\tilde{T}$. This function is parameterized by a single parameter $\tilde{\alpha}$ that controls chirp duration (see Fig. 6).

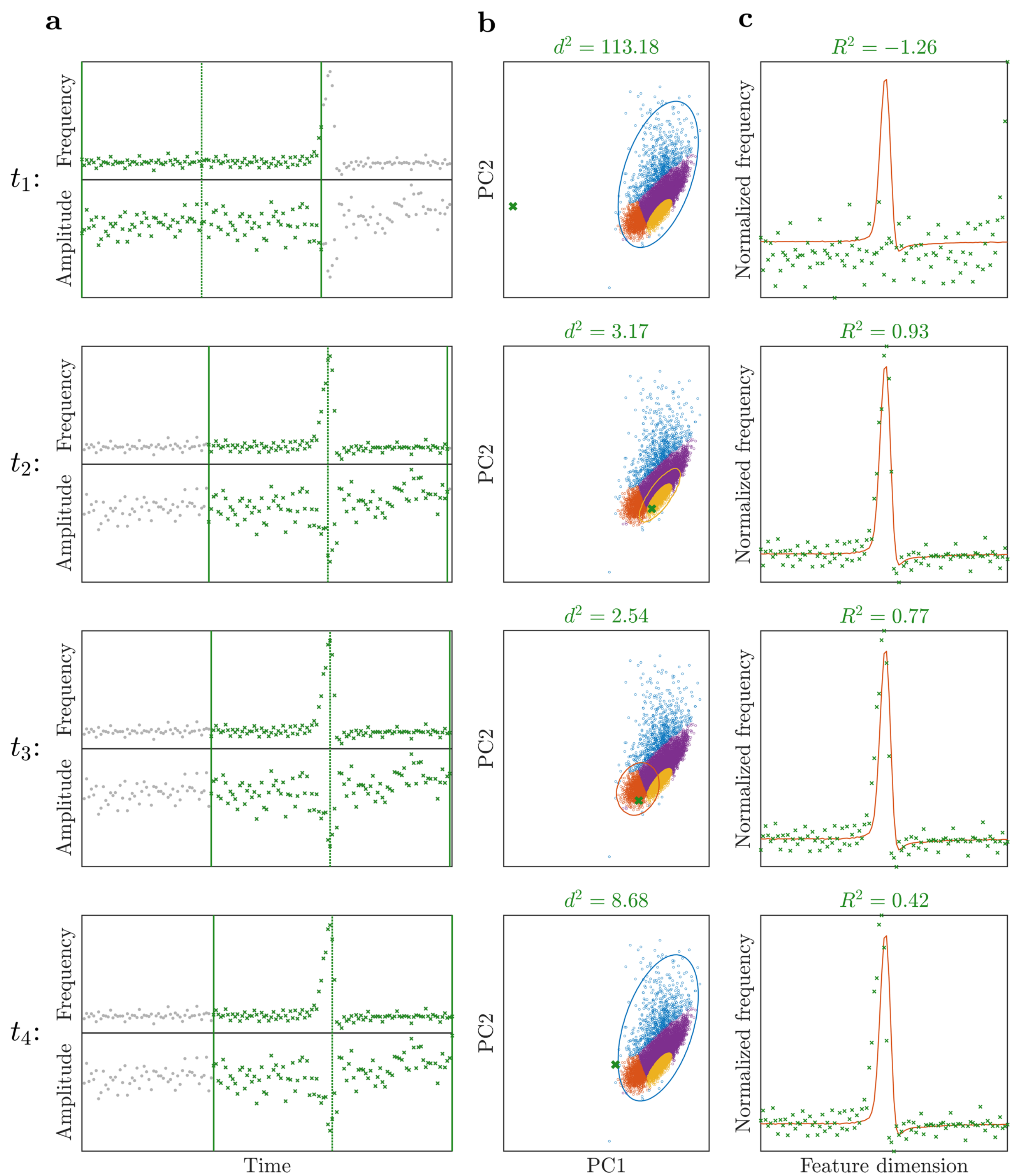
The TFSB algorithm has 7 hyper-parameters, out of which we fixed 5 (see Table 2), and the remaining 2 we determined via cross-validation (see Sect. "Cross-validation").

Cross-validation

To determine the optimal hyper-parameter values $\mathbf{h}_{\mathrm{opt}}$ of detection algorithms, we used k-fold cross-validation. In particular, we randomized indices $i \in \mathbf{i}_{\mathrm{GT}}$ associated with time-series data segments $\mathbf{S}_i$ and split them onto k number of equal-size folds: $\mathbf{i}_{\mathrm{GT},q} \subset \mathbf{i}_{\mathrm{GT}}$, $q = 1, \dots, k$. For each iteration step $q = 1, \dots, k$, of cross validation, a single fold $\mathbf{i}_{\mathrm{GT},q}$ was used as a test set for determining the performance of the algorithm, while the rest of the folds were used as a training set. Note that only the two supervised algorithms were trained (for details, see Sect. "Training"), while the TFSB algorithm did not involve any training (Eske et al. 2023). The performance of each algorithm was determined by computing the false positive and false negative rates for each iteration step $q = 1, \dots, k$, as

$$\mathrm{FP}_q = \frac{\sum_{s \in \mathbf{i}_{\mathrm{GT},q}} \sum_{j=1}^{m_{\mathrm{A},s}} \mathbb{1}\left(\left| \left\{ i : \hat{T}_j^{(s)} \notin \left[T_{i,1}^{(s)}, T_{i,l_i}^{(s)} \right], 1 \le i \le m_{\mathrm{GT},s} \right\} \right| = m_{\mathrm{GT},s} \right)}{\sum_{s \in \mathbf{i}_{\mathrm{GT},q}} m_{\mathrm{A},s}}, \tag{35}$$

$$\mathrm{FN}_q = 1 - \frac{\sum_{s \in \mathbf{i}_{\mathrm{GT},q}} \sum_{i=1}^{m_{\mathrm{GT},s}} \mathbb{1}\left(\left| \left\{ j : \hat{T}_j^{(s)} \in \left[T_{i,1}^{(s)}, T_{i,l_i}^{(s)} \right], 1 \le j \le m_{\mathrm{A},s} \right\} \right| > 0 \right)}{\sum_{s \in \mathbf{i}_{\mathrm{GT},q}} m_{\mathrm{GT},s}}, \tag{36}$$



where $\mathbb{1}(\cdot)$ is the indicator function, $\hat{T}_j^{(s)}$ denotes the j-th time instance of chirps detected by the algorithm in time-series data segment $\mathbf{S}_s$, while $T_{i,1}^{(s)}$ and $T_{i,l_i}^{(s)}$ correspond to the first and last data point of the i-th chirp sample in $\mathbf{G}_r$ collected from data segment $\mathbf{S}_s$. Parameters $m_{\mathrm{A},s}$ and $m_{\mathrm{GT},s}$ denote the total number of chirps detected by the algorithm in $\mathbf{S}_s$, and collected manually from $\mathbf{S}_s$, respectively. The overall performance of the algorithm was determined by averaging over all folds:

$$\overline{\mathrm{FP}}(\mathbf{h}) = \frac{1}{k}\sum_{q=1}^{k} \mathrm{FP}_q(\mathbf{h}), \quad \overline{\mathrm{FN}}(\mathbf{h}) = \frac{1}{k}\sum_{q=1}^{k} \mathrm{FN}_q(\mathbf{h}). \tag{37}$$

◂**Fig. 5** Illustration of the chirp detection methods described in Sect. "Detection". Different rows correspond to different time instants ($t_1 < t_2 < t_3 < t_4$) of the "sliding" time window indicated by vertical *green lines* in **a**. At each time instant, the Mahalanobis-distance-based detection algorithm (**b**) normalizes the data set inside the time window (green crosses in **a**) and projects it to the space spanned by the principal components of the training set (green cross in **b**). If the squared Mahalanobis distance value d^2 associated with this projected point (indicated at the top of each row in **b**) is below the limit of the cluster with highest posterior probability (corresponding to the color-coded ellipse in **b**), then the Mahalanobis-distance-based algorithm may detect a chirp (2nd and 3rd row). At each time instant, the coefficient-of-determination-based algorithm (**c**) normalizes the data set inside the time window and computes its coefficient of determination with respect to each cluster mean. The highest coefficient-of-determination value R^2 among all cluster means is indicated at the top of each row in **c**, and the related cluster mean is plotted (color-coded line in **c**). If this value is above a threshold, then the coefficient-of-determination-based algorithm may detect a chirp

Note that false positive and false negative rates depend on hyper-parameters $\mathbf{h}$. We tuned the hyper-parameters such that for a given maximum tolerated average false positive rate r_{FP}, the average false negative rate is minimized, i.e.,

$$\mathbf{h}_{\mathrm{opt}}\left(r_{\mathrm{FP}}\right) = \underset{\mathbf{h}\in\mathbf{H}\left(r_{\mathrm{FP}}\right)}{\operatorname{argmin}}\left(\overline{\mathrm{FN}}(\mathbf{h})\right), \quad \mathbf{H}(r_{\mathrm{FP}}) = \left\{\mathbf{h} \in \boldsymbol{\Omega} : \overline{\mathrm{FP}}(\mathbf{h}) \le r_{\mathrm{FP}}\right\}, \tag{38}$$

where $\boldsymbol{\Omega}$ is the search domain of hyper-parameters. At the maximum tolerated average false positive rate r_{FP}, the lowest achievable average false negative rate is

$$r_{\mathrm{FN}}\left(r_{\mathrm{FP}}\right) = \overline{\mathrm{FN}}\left(\mathbf{h}_{\mathrm{opt}}\left(r_{\mathrm{FP}}\right)\right). \tag{39}$$

The implemented search domains of hyper-parameters are summarized in Table 3.

Results

Performance of detection algorithms

For the GT group $\mathbf{G}_2$, we computed the lowest achievable average false negative rate r_{FN} of each algorithm at given average false positive rate tolerances r_{FP} (see Fig. 7) according to Eq. 39, using the search domains in Table 3. These results show that the performance of the MDB method is inferior to the CDB and TFSB methods. The CDB method performs better than the MDB and TFSB methods, although, the $r_{\mathrm{FN}}\left(r_{\mathrm{FP}}\right)$ curves of the CDB and TFSB methods are nearly identical (Fig. 7).

Principal components and explained variance

To illustrate waveform components that dominate GT group $\mathbf{G}_2$, we computed its PCs (Fig. 8a, b) and the explained variance in terms of the number of its retained PCs (Fig. 8c). The first PC explains 90% of the variation in $\mathbf{G}_2$ (Fig. 8c). The frequency shape of the first PC (PC1 in Fig. 8a, b) is similar to the chirp waveform of the TFSB method (cf. Fig. 6). This, together with the high percentage of explained variance associated with the first PC, result in a similar performance of the TFSB method and the CDB method (Fig. 7).

Chirp detection

After cross-validation, we trained a model according to Sect. "Training" based on the entire GT data set $\mathbf{G}_2$. We used optimal hyper-parameters $\mathbf{h}_{\mathrm{opt}}\left(r_{\mathrm{FP}} = 5\%\right)$ determined via 4-fold cross-validation (see Sect. "Cross-validation"). The cluster means of the model, computed according to Eq. 22, are shown in Fig. 9.

After training, we employed the CDB method (under hyper-parameters $\mathbf{h}_{\mathrm{opt}}\left(r_{\mathrm{FP}} = 5\%\right)$) to detect chirps in all 1680 min of EOD recordings. A total of 30,734 chirps were detected. We further investigated all detected chirps assigned to the cluster mean with the smallest proportion (6.73%, see Fig. 9). To find sub-clusters, we fitted a new GMM on these chirps according to Sect. "Training" using $N = 4$ and $C = 8$.

This analysis revealed a new chirp type (see Fig. 10) characterized by short, 20–30 ms duration, and two peaks in frequency rise and amplitude drop. These latter characteristics are distinct from all previously identified chirps of similar duration (c.f. Engler et al. 2000). It is important to note that here we focused on the cluster mean with the smallest proportion. The sub-clustering of chirps assigned to other cluster means may reveal further chirp types.

The distinct feature of this novel type, compared to the previously described six chirp types (Engler et al. 2000; Zupanc et al. 2006), is the existence of two frequency peaks (instead of just one peak); and the occurrence of two amplitude drops—the first, rather modest amplitude decrease is followed by a second, more pronounced reduction. Double frequency peaks have also been found in other apteronotid species, most notably in the *A. bonapartii* group (Turner et al. 2007). However, unlike the duplet frequency modulation characterizing the novel type in *A. leptorhynchus*, in *A. bonapartii* the first frequency increase is followed by a second, less pronounced increase.

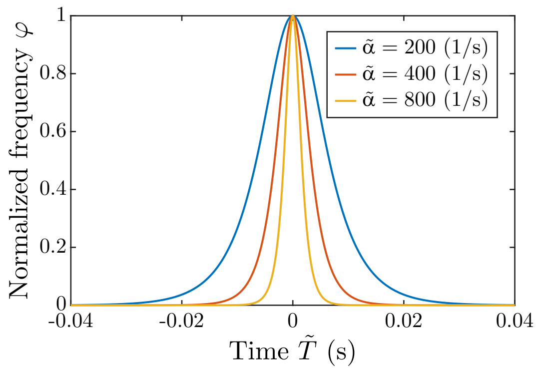


Fig. 6 Time–frequency shape function used for chirp detection in Eske et al. (2023). The time course of normalized frequency (see Eq. 5) during chirping is modeled by a single-parameter function $\varphi(\tilde{T};\tilde{\alpha})$. Different colors correspond to different shape parameter values $\tilde{\alpha}$

Table 2 Fixed hyper-parameters of the time-frequency-shape-based chirp detection algorithm (for definition of parameters, see Eske et al. 2023)

Parameter (unit)	Value
n_{wind} (1)	100
n_{med} (1)	10
n_{α} (1)	200
α_{min} (1/s)	100
α_{max} (1/s)	800

Table 3 Search domains of hyper-parameters for the Mahalanobis-distance-based (MDB), coefficient-of-determination-based (CDB), and time-frequency-shape-based (TFSB) chirp detection algorithms

Hyper-parameter	Search domain		
	MDB	CDB	TFSB
N	$\{2, 3, \ldots, 8\}$	$\{2, 3, \ldots, 8\}$	n/a
C	$\{2, 3, \ldots, 14\}$	$\{2, 3, \ldots, 14\}$	n/a
ε_f	$\{0, 2, \ldots, 20\}$	$\{0, 2, \ldots, 20\}$	$\{0, 2, \ldots, 20\}$
α	$\{0.01, 0.02, \ldots, 0.15\}$	n/a	n/a
ε_{R^2}	n/a	$\{0.3, 0.35, \ldots, 0.8\}$	$\{0.3, 0.35, \ldots, 0.8\}$

Discussion

Advantages of the supervised-learning method

The results presented in this paper demonstrate the superiority of our supervised-learning algorithm over traditional methods for analysis of chirps produced by *A. leptorhynchus*.

The first advantage of our method lies in its versatility, compared to traditional approaches. As shown in Sect. "Principal components and explained variance", the TFSB method performs well for the herein analyzed signal segments because a single time-frequency waveform (associated with type 2 chirps) dominates the collected GT chirp data set, and this waveform matches well the assumed time-frequency shape. If multiple dominant waveforms are present in the GT chirp data set, or if the assumed time-frequency shape does not match the dominant chirp waveform, the performance of the TFSB method would be significantly worse. Furthermore, the design of a shape function representative of the dominant chirp waveform is rather cumbersome and impacted by the researcher's bias. In contrast, the supervised-learning algorithm autonomously trains chirp waveform models by fitting them to the collected GT chirp data. Given that the GT data set is representative of chirps in the analyzed signal, this algorithm provides an unbiased way for the automatic identification of dominant chirp waveforms in the signal.

The second advantage of our supervised-learning method is its ability to identify, in an unbiased way, possible subtypes of a signal. In the case of chirping behavior in *A.*

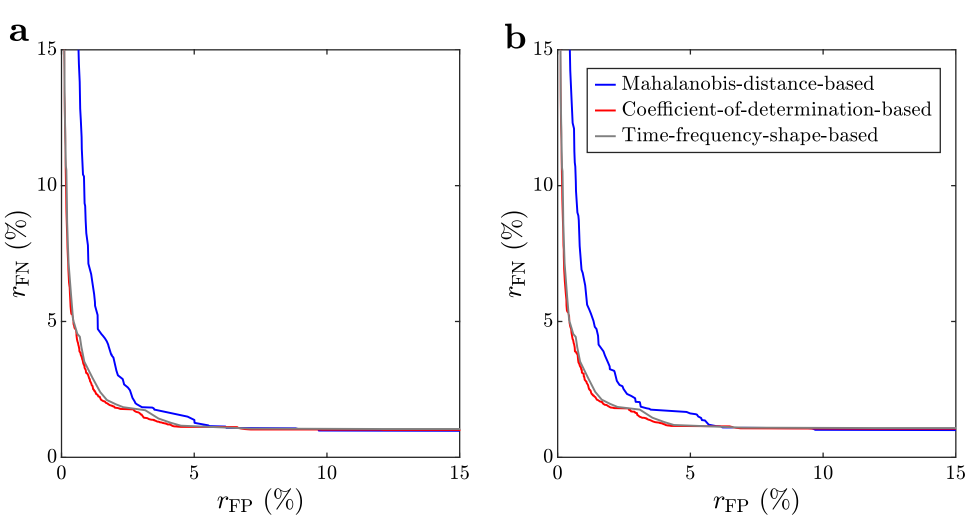


Fig. 7 Lowest achievable average false negative rate (r_{FN}) as a function of maximum allowed average false positive rate (r_{FP}), using k-fold cross-validation with $k = 2$ (**a**) and $k = 4$ (**b**). Curves were calculated for "ground truth" data set $\mathbf{G}_2$ according to Sect. "Cross-validation". Different colors are associated with different methods, as indicated in the top right corner of **b**

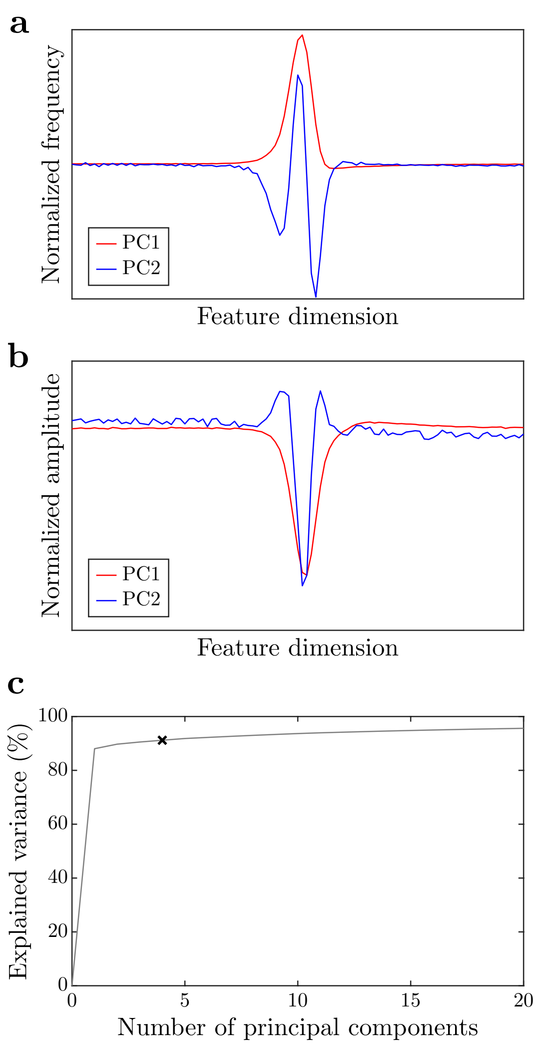


Fig. 8 Normalized frequency (**a**) and amplitude (**b**) in the first two principal components (PC1 and PC2) of the “ground truth” data set $\mathbf{G}_2$. Explained variance as a function of retained principal components (*gray line* in **c**). The *black cross* in **c** corresponds to optimal hyper-parameter settings $\mathbf{h}_{\text{opt}}(r_{\text{FP}} = 5\%)$ determined by 4-fold cross-validation (see Sect. “Cross-validation”)

leptorhynchus, visual inspection of time–frequency plots and time–voltage plots has suggested six subtypes of this signal (Engler et al. 2000; Zupanc et al. 2006). Although, in the present study, the analyzed recordings contained predominantly a single chirp subtype (type 2), our method suggested that further differentiation of this subtype is possible (see Sect. “Chirp detection”).

The third advantage of our method is that, compared to traditional approaches, it extracts more information from the samples used for the validation of the algorithm. Note that only a few traditional approaches validate their algorithm (e.g., Eske et al. 2023) by signals with known chirp types and locations. However, these approaches use the collected set of chirps only to test efficiency, and thus the algorithm itself is not informed by the known chirp content. By contrast, our supervised learning method takes full advantage of known chirps and utilizes them for both training the algorithm and testing its efficiency.

Limitations of the method

Although our algorithm trains itself and identifies chirp clusters automatically, it still relies on the collection of GT samples. Consequently, results are still impacted by the bias of the individual who collects the chirp samples of the GT set. This bias can be reduced if multiple individuals carry out chirp collection using the same signal, and if the GT set

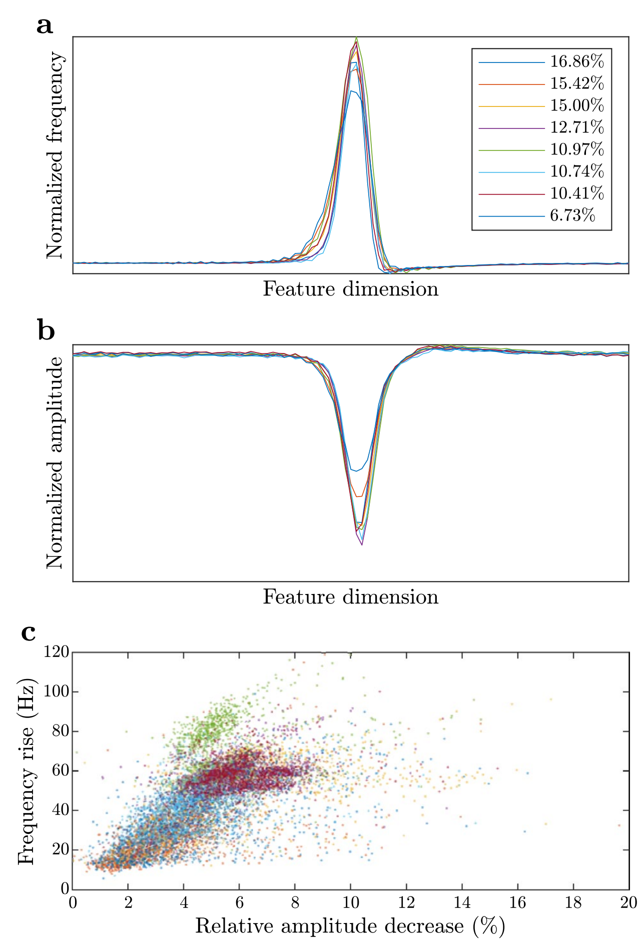


Fig. 9 Normalized frequency (**a**) and amplitude (**b**) components of cluster means for the model trained on the entire “ground truth” data set $\mathbf{G}_2$. Clusters were computed under optimal parameter settings $\mathbf{h}_{\text{opt}}(r_{\text{FP}} = 5\%)$ determined via 4-fold cross-validation (see Sect. “Cross-validation”). Relative voltage-amplitude-decrease/maximum-frequency-rise pairs plotted for each sample (**c**) reveal that $\mathbf{G}_2$ consists entirely of type 2 chirps. Different colors are associated with different clusters, their proportions are indicated in the top-right corner of **a**

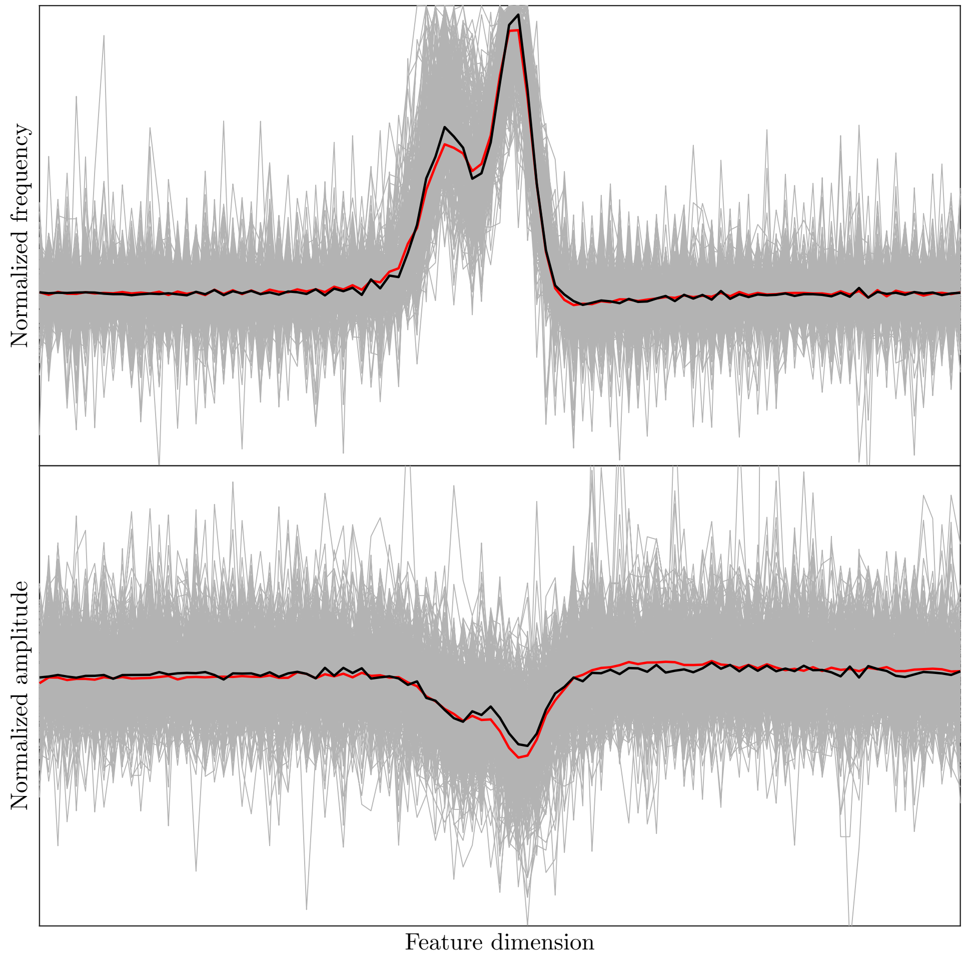


Fig. 10 Cluster mean (*red*) and detected chirp samples (gray) in a sub-cluster (containing 264 samples) related to the cluster with 6.73% proportion in Fig. 9a (see Sect. "Chirp detection"). The black curve displays the median of detected chirp samples across feature dimensions

is assembled based on the overlap across sets collected by different individuals.

When chirps appear in the signal at a low frequency, the time needed for an individual to collect a sufficiently large GT set increases. While the validation of any algorithm requires the collection of all chirps from a test signal, the number of samples needed by our supervised-learning method is higher than the number of samples needed for validation only. Nevertheless, our method can still be advantageous compared to traditional approaches when already detected chirp types are expected in future experiments. In such cases, the cluster shapes from already collected GT sets can be reused. Furthermore, one can even build libraries of cluster shapes which can then be employed to "scan" signals for all formerly identified chirp shapes.

As input, our supervised learning method uses the time–frequency–amplitude signal (T_k, f_k, A_k), $k = 1, 2, \ldots$. While the method for the computation of this signal, described in Sect. "Calculation of EOD frequency and amplitude", works only for time–voltage data that were generated by a single EOD source, for the analysis of multiple (either synthetic or recorded from fish) simultaneously recorded EOD signals, one can employ a different method (e.g., Raab et al. 2022) to extract individual time–frequency–amplitude signals.

Perspectives

The presented supervised learning algorithm provides a valuable tool for further examining the function of chirps. In the present study, it has not only enabled us to validate the previous classification of chirps into different subtypes, but also suggested that further differentiation of these subtypes is possible. Whether these sub-subtypes of chirps subserve any behavioral function remains to be examined.

It is likely that other algorithms based on supervised machine learning will exhibit advantages similar to our approach. Thus, the present study might serve as proof-of-principle of the suitability of the supervised-machine-learning paradigm for a broad range of signals analyzed in neuroethology. It is likely that, in future investigations, algorithms based on machine learning paradigms like the one implemented in the present study will increasingly become standard tools for signal analysis in neuroethological research.

Appendix A: Computation of time-series frequency and amplitude data

Using linear interpolation, we first computed all time instances where the filtered time-series signal $\left(t_i, V_i\right)$, $i = 1, \ldots, M_{\mathrm{v}}$, crosses the time axis toward positive voltage values:

$$\mathbf{t}^{+} = \left(t_{\mathbf{j}(k)} - \frac{V_{\mathbf{j}(k)}}{V_{\mathbf{j}(k)+1} - V_{\mathbf{j}(k)}} \left(t_{\mathbf{j}(k)+1} - t_{\mathbf{j}(k)} \right) \, : \, k = 1, \ldots, M+1 \right). \tag{A1}$$

Here the tuple of all upward crossings $\mathbf{t}^{+}$ contains $|\mathbf{j}| = M + 1$ number of elements, with

$$\mathbf{j} = \left(i \, : \, \mathrm{sgn}\left(V_{i+1}\right) - \mathrm{sgn}\left(V_i\right) > 0, i = 1, \ldots, M_{\mathrm{v}} \right) \tag{A2}$$

being a tuple containing all indices of the filtered time-series data where the voltage changed sign to a positive value. Finally, for each oscillation interval $\left[\mathbf{t}^{+}(k), \mathbf{t}^{+}(k+1)\right]$, $k = 1, \ldots, M$, we computed time instance T_k, and associated frequency f_k and peak-to-peak amplitude A_k values as

$$T_k = \frac{\mathbf{t}^{+}(k+1) + \mathbf{t}^{+}(k)}{2}, \tag{A3}$$

$$f_k = \frac{1}{\mathbf{t}^{+}(k+1) - \mathbf{t}^{+}(k)}, \tag{A4}$$

$$A_k = \max_{\mathbf{j}(k) \le i \le \mathbf{j}(k+1)} \left(v_i\right) - \min_{\mathbf{j}(k) \le i \le \mathbf{j}(k+1)} \left(v_i\right). \tag{A5}$$

Author Contributions Study concept and design: DL and GKHZ. Design of supervised-learning algorithm: DL. Data analysis: DL. Writing of manuscript: DL, GKHZ. Preparation of figures: DL. Review and editing of manuscript: DL, GKHZ. Both authors read and approved the manuscript.

Funding Open access funding provided by Northeastern University Library. This study was supported by Grant 1946910 from the National Science Foundation (GKHZ).

Availability of data and materials The code file implementing the supervised-learning algorithm is available in a public GitHub repository at the following link: https://github.com/LaboratoryOfNeurobiology/supervised_learning_of_chirp_patterns. Data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Allen KM, Marsat G (2019) Neural processing of communication signals: the extent of sender-receiver matching varies across species of *Apteronotus*. eNeuro. https://doi.org/10.1523/eneuro.0392-18.2019

Aumentado-Armstrong T, Metzen MG, Sproule MKJ et al (2015) Electrosensory midbrain neurons display feature invariant responses to natural communication stimuli. PLoS Comput Biol 11(10):e1004430. https://doi.org/10.1371/journal.pcbi.1004430

Bastian J, Schniederjan S, Nguyenkim J (2001) Arginine vasotocin modulates a sexually dimorphic communication behavior in the weakly electric fish *Apteronotus leptorhynchus*. J Exp Biol 204(11):1909–1923. https://doi.org/10.1242/jeb.204.11.1909

Bishop CM (2006) Pattern recognition and machine learning. Springer Science+Business Media, New York

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland

Dulka JG, Maler L (1994) Testosterone modulates female chirping behavior in the weakly electric fish, *Apteronotus leptorhynchus*. J Comp Physiol A 174:331–343. https://doi.org/10.1007/BF00240215

Dunlap KD (2002) Hormonal and body size correlates of electrocommunication behavior during dyadic interactions in weakly electric fish, *Apteronotus leptorhynchus*. Horm Behav 41:187–194. https://doi.org/10.1006/hbeh.2001.1744

Dunlap KD, Larkins-Ford J (2003) Diversity in the structure of electrocommunication signals within a genus of electric fish *Apteronotus*. J Comp Physiol A 189(2):153–161. https://doi.org/10.1007/s00359-003-0393-3

Dunlap KD, Thomas P, Zakon HH (1998) Diversity of sexual dimorphism in electrocommunication signals and its androgen regulation in a genus of electric fish, *Apteronotus*. J Comp Physiol A 183(1):77–86. https://doi.org/10.1007/s003590050236

Dunlap KD, Jashari D, Pappas KM (2011) Glucocorticoid receptor blockade inhibits brain cell addition and aggressive signaling in electric fish, *Apteronotus leptorhynchus*. Horm Behav 60(3):275–283. https://doi.org/10.1016/j.yhbeh.2011.06.001

Engler G, Zupanc GKH (2001) Differential production of chirping behavior evoked by electrical stimulation of the weakly electric fish, *Apteronotus leptorhynchus*. J Comp Physiol A 187:747–756. https://doi.org/10.1007/s00359-001-0248-8

Engler G, Fogarty CM, Banks JR et al (2000) Spontaneous modulations of the electric organ discharge in the weakly electric fish,

Apteronotus leptorhynchus: a biophysical and behavioral analysis. J Comp Physiol A 186:645–660. https://doi.org/10.1007/s003590000118
Eske AI, Lehotzky D, Ahmed M et al (2023) The effect of urethane and MS-222 anesthesia on the electric organ discharge of the weakly electric fish *Apteronotus leptorhynchus*. J Comp Physiol A. https://doi.org/10.1007/s00359-022-01606-6
Feng AS, Riede T, Arch VS et al (2009) Diversity of the vocal signals of concave-eared torrent frogs (*Odorrana tormota*): evidence for individual signatures. Ethology 115(11):1015–1028. https://doi.org/10.1111/j.1439-0310.2009.01692.x
Field CE, Petersen TA, Alves-Gomes JA et al (2019) A JAR of chirps: the gymnotiform chirp can function as both a communication signal and a jamming avoidance response. Front Integr Neurosci 13:55. https://doi.org/10.3389/fnint.2019.00055
Gama Salgado JA, Zupanc GKH (2011) Echo response to chirping in the weakly electric brown ghost knifefish (*Apteronotus leptorhynchus*): role of frequency and amplitude modulations. Can J Zool 89:498–508. https://doi.org/10.1139/Z11-014
Hechavarría JC, Beetz JM, García-Rosales F et al (2020) Bats distress vocalizations carry fast amplitude modulations that could represent an acoustic correlate of roughness. Sci Rep 10(1):7332. https://doi.org/10.1038/s41598-020-64323-7
Henninger J, Krahe R, Kirschbaum F et al (2018) Statistics of natural communication signals observed in the wild identify important yet neglected stimulus regimes in weakly electric fish. J Neurosci 38(24):5456–5465. https://doi.org/10.1523/jneurosci.0350-18.2018
Ho WW, Turner CR, Formby KJ et al (2013a) Sex differences in the electrocommunication signals of *Sternarchogiton nattereri* (Gymnotiformes: Apteronotidae). J Ethol 31(3):335–340. https://doi.org/10.1007/s10164-013-0382-0
Ho WW, Turner CR, Smith GT (2013b) Transition across polymorphic phenotypes observed in a male *Sternarchogiton nattereri* (Gymnotiformes: Apteronotidae). J Fish Biol 83(3):667–670. https://doi.org/10.1111/jfb.12188
Hupé GJ, Lewis JE (2008) Electrocommunication signals in free swimming brown ghost knifefish, *Apteronotus leptorhynchus*. J Exp Biol 211(Pt 10):1657–1667. https://doi.org/10.1242/jeb.013516
Ilieş I, Traniello IM, Sîrbulescu RF et al (2014) Determination of relative age using growth increments of scales as a minimally invasive method in the tropical freshwater *Apteronotus leptorhynchus*. J Fish Biol 84(5):1312–1325. https://doi.org/10.1111/jfb.12354
Kolodziejski JA, Sanford SE, Smith GT (2007) Stimulus frequency differentially affects chirping in two species of weakly electric fish: implications for the evolution of signal structure and function. J Exp Biol 210(14):2501–2509. https://doi.org/10.1242/jeb.005272
Meyer JH, Leong M, Keller CH (1987) Hormone-induced and maturational changes in electric organ discharges and electroreceptor tuning in the weakly electric fish *Apteronotus*. J Comp Physiol A 160(3):385–394. https://doi.org/10.1007/BF00613028
Neeley B, Overholt T, Artz E et al (2018) Selective and context-dependent social and behavioral effects of Δ^9-tetrahydrocannabinol in weakly electric fish. Brain Behav Evol 91(4):214–227. https://doi.org/10.1159/000490171
Oboti L, Pedraja F, Ritter M et al (2023) Why the brown ghost chirps at night. bioRxiv. https://doi.org/10.1101/2022.12.29.522225
Raab T, Madhav MS, Jayakumar RP et al (2022) Advances in non-invasive tracking of wave-type electric fish in natural and laboratory settings. Front Integr Neurosci. https://doi.org/10.3389/fnint.2022.965211
Schwartz C, Tressler J, Keller H et al (2007) The tiny difference between foraging and communication buzzes uttered by the mexican free-tailed bat, *Tadarida brasiliensis*. J Comp Physiol A 193(8):853–863. https://doi.org/10.1007/s00359-007-0237-7
Smith GT, Combs N (2008) Serotonergic activation of 5 HT_{1A} and 5 HT_2 receptors modulates sexually dimorphic communication signals in the weakly electric fish, *Apteronotus leptorhynchus*. Horm Behav 54(1):69–82. https://doi.org/10.1016/j.yhbeh.2008.01.009
Triefenbach F, Zakon H (2003) Effects of sex, sensitivity and status on cue recognition in the weakly electric fish, *Apteronotus leptorhynchus*. Anim Behav 65:19–28. https://doi.org/10.1006/anbe.2002.2019
Turner CR, Derylo M, de Santana CD et al (2007) Phylogenetic comparative analysis of electric communication signals in ghost knifefishes (Gymnotiformes: Apteronotidae). J Exp Biol 210(23):4104–4122. https://doi.org/10.1242/jeb.007930
Zupanc GKH, Bullock TH (2005) From electrogenesis to electroreception: an overview. In: Bullock TH, Hopkins CD, Popper AN et al (eds) Electroreception. Springer Science + Business Media, New York, pp 5–46. https://doi.org/10.1007/0-387-28275-0_2
Zupanc GKH, Maler L (1993) Evoked chirping in the weakly electric fish *Apteronotus leptorhynchus*: a quantitative biophysical analysis. Can J Zool 71:2301–2310. https://doi.org/10.1139/z93-323
Zupanc MM, Engler G, Midson A et al (2001) Light-dark-controlled changes in modulations of the electric organ discharge in the teleost *Apteronotus leptorhynchus*. Anim Behav 62(6):1119–1128. https://doi.org/10.1006/anbe.2001.1867
Zupanc GKH, Sîrbulescu RF, Nichols A et al (2006) Electric interactions through chirping behavior in the weakly electric fish, *Apteronotus leptorhynchus*. J Comp Physiol A 192:159–173. https://doi.org/10.1007/s00359-005-0058-5
Zupanc GKH, Ilieş I, Sîrbulescu RF et al (2014) Large-scale identification of proteins involved in the development of a sexually dimorphic behavior. J Neurophysiol 111:1646–1654. https://doi.org/10.1152/jn.00750.2013