Response of wild songbirds to songs synthesized with a low-dimensional model

Roberto Bistel ®,[1,2] Ana Amador ®,[1,2] and Gabriel B. Mindlin ®[1,2,3]

[1]*Facultad de Ciencias Exactas y Naturales, Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*Instituto de Física Interdisciplinaria y Aplicada (INFINA), CONICET–Universidad de Buenos Aires, Buenos Aires, Argentina*
[3]*Departamento de Matemática Aplicada, Universidad Rey Juan Carlos, Madrid, Spain*

In this work, we used a dynamical system derived from an avian vocal production model to generate synthetic songs that mimic the *Zonotrichia capensis* songs. We confirmed that these synthetic renditions elicited behavioral responses similar to those evoked by real songs in wild songbirds of the same species. Specifically, we observed an increase in the singing rate of individual birds when a playback device was introduced into their territories. The success of our approach instills confidence in the hypotheses underpinning the model and provides a valuable tool for investigating a wide range of biological questions.

I. INTRODUCTION

Among the approximately 10 000 bird species existing in the world, around 4000 belong to a group known as oscine birds, often referred to as songbirds. In these species song is a behavior that involves learning in a stellar way [1]; some aspects of song are genetically conditioned, but others are learned through interaction between the juveniles and adults of a given species.

The mathematical model used in this work is based on the physics of avian phonation and is capable of generating a time series data emulating song [2–5]. Specifically, it describes the dynamics of the syringeal labia, which are responsible for modulating airflow and, consequently, the production of sound. The muscle activity during birdsong production enters the model as time-dependent parameters, and it is possible to choose them in such a way that the sound presents the temporal and acoustic properties of the song to be modeled. In this model, many of the timbral properties of the sound result from the nonlinear properties of the model [6]. Despite its low dimensionality, the synthetic songs generated by this model evoke neural responses, similar to those evoked by the bird's own song [7,8]. As the neural responses are highly selective to the bird's own song (i.e., the neural response is severely reduced when presenting other auditory stimuli) these experiments serve as a biological accuracy test of the synthetic copy generated by the model. Positive results were also found when recording electromyographic activity in syringeal muscles [9], which receive direct input from the neural network that generates and processes birdsong.

The success of a low-dimensional model in evoking neural or muscular activity in a highly controlled experiment performed in a laboratory, does not guarantee that wild birds, displaying their natural, active and alert behaviors during the mating and breeding season, will react in a similar way to real and synthetic songs. For this reason, to further test the dynamical model of song production, we devised an experiment to compare the response of wild birds to both types of stimuli. We took advantage of the highly territorial behavior of the *Zonotrichia capensis* (Rufous-collared sparrow) whose typical response to a conspecific bird singing in its territory during the breeding season involves an investigation of the invasive sound source, and an increase in the rate of singing [10,11]. In this work, we quantified the behavioral responses of territorial male adult birds to synthetic and real songs, in order to estimate the degree of realism of the synthetic acoustic stimuli. Playing back both real and synthetic songs in the field increases the rate of individual singing, while songs from other species fail to evoke such a response. Furthermore, in this work we present that there are no significant differences between responses to real and synthetic songs. This validates our approach and highlights its potential as a versatile tool for understanding various facets of avian behavior.

Zonotrichia capensis presents an advantageous attribute. Within a population of subjects capable of interacting in a common territory, different individuals sing different songs (typically, each individual sings a single song during his life). In this way, it is possible to develop an experimental protocol that avoids "pseudo-replication" [12]. This is a potential failure in experimental designs with auditory playback, in which a single stimulus is used on a number of subjects, and the repeated samples are analyzed statistically as if they were replicates [12]. A careful experimental design requires that both the subject population and the stimulus population be adequately sampled, and there are many ways to avoid pseudo-replication (see, e.g., [13–15]). Given the diversity of different songs in our experimental region, this point could be carefully designed; the different individuals were subjected to stimuli corresponding to various songs [16–18]. To avoid pseudo-replication, we used the flexibility of the dynamic model to create 11 synthetic songs of *Zonotrichia capensis*, which have the acoustic parameters of a sample of individuals sharing the same theme. These synthetic songs, along with the real songs and the control species songs, were randomly reproduced to individuals of *Zonotrichia capensis*, with a spatial separation of at least 100 m.

The article is structured as follows: Section II provides a description of the materials and methods used in our study.
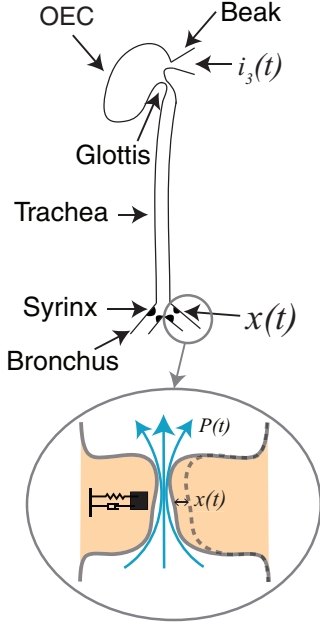
FIG. 1. The components within the low-dimensional model for the oscine syrinx include two pairs of labia situated at the junction of the bronchi and the trachea, which regulate airflow. Subsequently, the resulting sound waves undergo filtration as they pass through the trachea, the oro-esophageal cavity (OEC), and the beak.

This section covers the models utilized for generating synthetic songs, the hardware designed for playing stimuli and recording responses, as well as the protocols performed for our experiment. In Sec. III, we present our findings, and in Sec. IV, we discuss our conclusions.

II. MATERIALS AND METHODS

A. Generation of synthetic songs

Songbirds possess a well-preserved vocal organ across species. This vocal apparatus includes two pairs of labia located at the junctures between the bronchi and the trachea (see Fig. 1), as described in studies by [19,20]. When these labia are set in a specific "prephonatory position" and subjected to a sufficiently high airflow, they can produce self-sustained oscillations, resulting in flow fluctuations that produce sound waves. This phenomenon has been discussed by [2,4,21]. These sound waves subsequently pass through the trachea and the oro-esophageal cavity (OEC; see Fig. 1), where they are filtered, before being emitted into the environment, as explained in studies by [22,23]. These oscine birds have been the subject of extensive research, primarily due to the crucial role of learning in song acquisition. The species examined in this study is the *Zonotrichia capensis*, commonly known as the Rufous-collared sparrow, and it falls within the category of oscine birds.

Considering these mechanisms, we produce synthetic songs through a structured three-step process. Firstly, we begin by solving the differential equations derived from the application of Newton's equations to the syringeal labia [24]. If the variable x describes the departure from the resting point

of the midpoint position of a labium, its dynamics can be described using the following dynamical system:

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = -\gamma^2 \kappa x - \gamma \, x^2 y + \beta \gamma y,$$

where γ is a timescale factor, κ is proportional to the tension in the syringeal muscles, and β is proportional to the air sac pressure. The flow fluctuations at the entrance of the trachea are due to the variations of the effective area between the labia due to the dynamics of x. If the trachea is described as a tube of length L, the pressure at its entrance can be calculated as

$$P(t) \propto y + P_{\text{back}}(t - L/c),$$
$$P_{\text{back}}(t) = -r P(t - L/c),$$

where L represents the length of the tube, c the speed of sound, and r the reflection coefficient at the upper end of the trachea. Finally, the OEC filters the sound that comes out of the trachea, whose dynamics obey the following equations [22,23] corresponding to a Helmholtz resonator coupled to the beak:

$$\frac{di_1}{dt} = i_2,$$
$$\frac{di_2}{dt} = -\frac{1}{C_h L_g} i_1 - \left(\frac{r_{\text{dis}}}{L_b} + \frac{r_{\text{dis}}}{L_g}\right) i_2$$
$$+ \left(\frac{1}{C_h L_g} - \frac{r_b \, r_{\text{dis}}}{L_b L_g}\right) i_3 + \frac{1}{L_g}\frac{dP_{\text{back}}}{dt} + \frac{r_{\text{dis}}}{L_b L_g} P_{\text{back}},$$
$$\frac{di_3}{dt} = -\frac{L_g}{L_b} i_1 - \frac{r_b}{L_b} i_3 + \frac{1}{L_b} P_{\text{back}}.$$

In these equations, i_3 corresponds to the output sound, the resonator dissipation and beak are represented by r_{dis} and r_b, the inertances of the beak and the resonator by L_b and L_H, and their compliances, by C_b and C_H. The values of these parameters were $(\frac{1}{C_h}, \frac{1}{L_b}, \frac{1}{L_g}, r_b, r_{\text{dis}}) = (7 \times 10^9, 10^{-4}, 5^{-2}, 10^7, 1.2^7)$. It is worth pointing out that a previous model of *Zonotrichia capensis* song [4] did not include detailed models of the OEC and beak, which was proved to be a key factor for eliciting neural responses in playback experiments with synthetic songs [7].

To emulate a given song, the motor gestures represented by the parameters (β, κ) were estimated as follows. The parameter β (proportional to the pressure of the bird's air sacs during the expiration) was set as a sinusoidal function with its argument spanning half a period in the time interval that the syllable lasts. This means that for time $t \in [t_i, t_f]$, with t_i, t_f standing for the initial and final times of the syllables, respectively, β reads

$$\beta = \sin[\pi(t - t_i)/(t_f - t_i)].$$

The values of κ were obtained following two steps. First, the song was separated into syllables. For each of these, the fundamental frequency was measured at $N \sim 10$ times, obtaining pairs (t_j, ω_j), $j = 1, \ldots, N$. These values were used to perform an interpolation that gave rise to frequency values ω_i at all the times t_i of the numerical integration between the initial and final times of each syllable, i.e.,

$t_i = i dt$, $i = 1, 2, \ldots, (t_{\text{final}} - t_{\text{initial}})/dt$. This interpolation was implemented using the function `interp1d` (cubic interpolation), from the Scipy library (PYTHON). Once the frequency time series for each syllable was obtained, the second step consisted of using the numerically fitted function

$$\kappa(\omega) = 6.56 \times 10^{-8} \omega^2 + 4.23 \times 10^{-5} \omega + 2.67 \times 10^{-2},$$

in order to obtain the values of κ that allow the model to synthesize a sound of fundamental frequency ω. In this way we obtain the time-dependent parameters $[\beta(t_i), \kappa(t_i)]$, for $t_i = i\, dt, i = 1, 2, \ldots, (t_{\text{final}} - t_{\text{initial}})/dt$. The numerical integration of the model makes it possible to obtain sounds of the desired fundamental frequencies, with a spectral content resulting from the nonlinear dynamics underlying the labial movement (see [6]) and the posterior filtering performed by the trachea, the OEC, and the beak [23].

B. Hardware for programmable sound playback and sound recording

The field experiments conducted on Rufous-collared sparrows involved playing auditory stimuli (previously recorded songs) and recording the birds' responses, which consisted of the songs produced by the tested bird. To perform these experiments, we designed and built a device capable of playing back sounds stored on a memory card and recording the bird's vocal responses elicited by the stimulus. For building this device we used the Arduino hardware platform [25].

The recording block comprises an Arduino Nano (5 V/16 MHz), a microphone module with an automatic gain amplifier MAX9814, a real-time clock (RTC) DS3231 and a microSD card module to store the recorded sounds. The AtMega328P microcontroller within the Arduino Nano incorporates an analog-digital converter that digitizes the signal from the microphone [26]. We use the real-time clock DS3231 to configure cyclical alarms that allow us to start and stop the recording and playback functions, allowing an automatic operation of the system.

The playback module includes the Arduino microcontroller, a DF Player Mini module and a speaker. The DF Player Mini module includes a card to store the sound files to be played as auditory stimuli. The system was programmed using the Arduino IDE (v1.8.12) platform [25]. The entire system is powered by a 12 V/7 Ah gel battery, connected to a regulated power source based on the XL4005 IC, which reduces the battery voltage to 5 V of direct current. In Fig. 2, we show the block diagram of our autonomous audio recording and playback device for field use. The circuits of our system, the Arduino firmware, and construction instructions are accessible in [27].

C. Protocol for comparing behavioral responses to synthetic and real songs

The protocol for comparing behavioral responses to auditory stimuli of real and synthetic songs was carried out using programmable recording-playback systems (see the section Hardware for programmable sound playback and sound recording). Briefly, the device would play back an auditory stimulus (song) and the behavioral response would be the
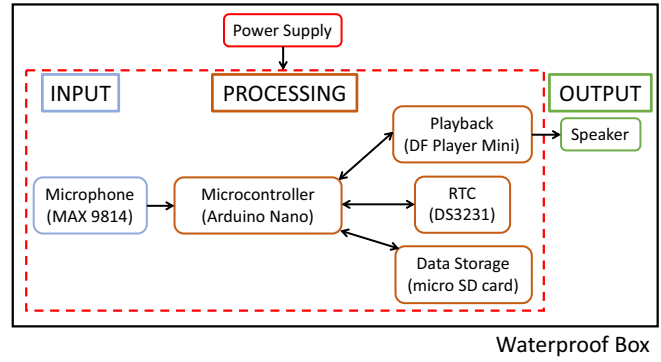

FIG. 2. Block diagram of our autonomous audio recording and playback device for field use. The system's core is a processor from the Arduino family, which controls the modules required to perform device tasks. The recording block includes a microphone module with a gain amplifier MAX9814, a microcontroller (Arduino Nano), a RTC DS3231, and a microSD card module to store the recorded sounds. The playback module includes the Arduino microcontroller, a DF Player Mini module, and a speaker.

amount of singing (measured as the number of songs) that the auditory stimulus elicited. The songs used as auditory stimuli (real, synthetic, and heterospecific songs) were saved in "wav" format, 8-bits resolution in mono mode, at 32 kHz sampling frequency. The maximum song duration was 2.1 s. The recorded songs (serving as quantifiers of the behavioral response) were saved in wavformat, 8-bits resolution in mono mode and 22 kHz sampling frequency. This is the maximum recording sampling rate that guarantees no data loss. The down-sampling of the recorded songs was not a problem as the quantifier was the number of songs and no detailed analysis was performed in the recorded songs.

Our study area covers approximately 0.4km^2 and is located within the Parque Pereyra Iraola (biosphere reserve by UNESCO), Province of Buenos Aires, Argentina (GPS coordinates $[-34.861\,552, -58.116\,316), (-34.857\,520, -58.126\,272), (-34.860\,566, -58.128\,160), (-34.864\,017, -58.120\,951)]$. Since 2020, our group has been systematically recording Rufous-collared sparrow songs in that area. The protocols for comparing the responses to real song and synthetic songs were carried out during the breeding season, between August 4th, 2022 and November 8th, 2022. The approximate start time of the protocols was 8:30 a.m., and the end time around 1:00 p.m. In each protocol, an individual is exposed to the playback of a real song or its synthetic copy. For this, we used Rufous-collared sparrow songs that were recorded in the period 2020–2022 [28,29]. Protocols were performed on each individual with songs both similar and qualitatively different to their own song. As there are no statistically significant differences in these responses, we grouped these stimuli as "real songs", i.e., recorded Rufous-collared sparrow songs. In addition, each individual was exposed to heterospecific songs from three species present in the study area: Great Kiskadee (*Pitangus Sulphuratus*), Rufous hornero (*Furnarius Rufus*), and Spot-winged pigeon (*Patagioenas maculosa*). We used three different songs for each of the three species.

The protocols were 13 min long. For the first 2 min, the sound box recorded the natural sound. Next, the sound box
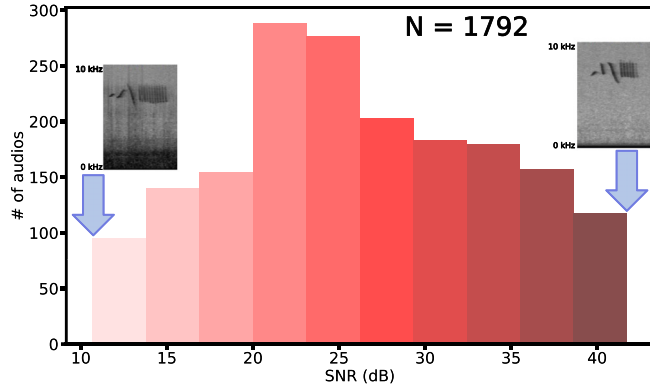
FIG. 3. Signal-to-noise ratio (SNR) of field recordings of Rufous-collared sparrow songs. The insets are spectrograms of song recordings with low and high SNR to show that in both cases the recording quality is good enough to process the songs.

repeatedly played a given song for 1 min, with 15-s intervals between songs. Afterwards, the sound box continued recording for 10 min. The protocol was repeated at regular 35-min intervals, controlled by the RTC alarm signal. In 1 day, an individual was exposed on average to three series of real songs and three series of synthetic songs.

Before beginning the first protocol of each day, the study area was explored. The locations of individuals with moderate singing activity (maximum four songs per minute) were identified. Adults in this species hardly vary the positions from where they sing throughout the breeding season [11]. The recording-reproduction system was located as close as possible to the place where the chosen individual sang. The minimum distance between two sites where simultaneous protocols were performed was 100 m. The playback volume allowed listening to each system within a radius of approximately 30 m. In total, 26 individuals were recorded in 15 different sites. These individuals were exposed to 256 protocols, of which 90 were carried out with real songs, 81 with synthetic songs, and 85 with heterospecific songs. We conducted a total of 76 protocols employing the same song as that produced by an individual (40 real and 36 synthetic songs), whereas 95 protocols involved distinct songs (50 real and 45 synthetic songs).

The signal-to-noise ratio of the recorded audios ($N = 1792$) presents a minimum of 10.68 dB, a maximum of 41.74 dB, and an average of 26.30 dB (see Fig. 3). The recorded audio files were processed using the AUDACITY software, obtaining the corresponding spectrograms (Hamming Window, 2048-sample window length and 50% overlap). We computed each spectrogram after processing the audio recordings with a noise reduction filter and a bandpass filter with cutoff frequencies between 1.5 and 8 kHz. In this way, we only focus on the frequency range of the Rufous-collared sparrow and minimize the influence of background noise.

The response to each stimulus was quantified by visual inspection of spectrograms. Two independent observers counted the number of songs per minute of each individual during the protocol. Each bird's unique song allowed us to confirm its identity throughout the entire procedure. Rufous-collared sparrows exhibit song stereotypy
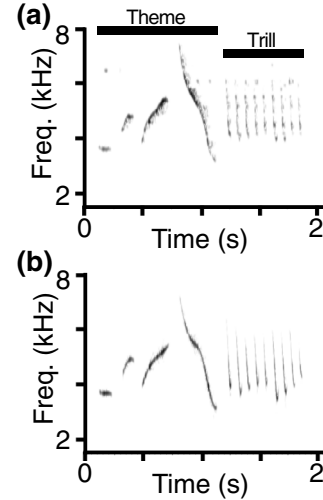

FIG. 4. Recorded Rufous-collared sparrow song (a), and its synthetic copy (b). Each song is composed by a theme (specific to each individual) and a trill. The synthetic song is generated by a dynamical model of the vocal organ, with frequency modulations similar to the song.

within individuals across breeding seasons [30]. Our focus was on quantifying the presence of songs linked to each individual.

III. RESULTS

A. *Zonotrichia capensis* songs and its synthetic copies generated with a low-dimensional model

The song of the *Zonotrichia capensis* consists of a stereotyped, brief theme composed of two to four notes, with each male typically singing a unique combination of them, followed by a trill [see Fig. 4(a)]. The entire song usually lasts approximately 2 s, and during the breeding season, the male will sing it approximately three times in 1 min. The identity of the individual is reflected in the theme and not in the trill. In fact, trill rate, and the frequency range of the syllables in the trill, are acoustic features shared by the individuals in geographical regions much larger than the area of our study [11].

Using the low-dimensional model for vocal production presented in Materials and Methods and following the procedure described in Sec. II A, we generated 11 synthetic themes, corresponding to 11 different themes recorded between 2020 and 2022 in the Parque Pereyra Iraola (Buenos Aires, Argentina). We obtained the range of variation of the parameters from the statistical analysis of song samples of each theme identified in the field recordings. We analyzed the spectrograms of ten songs per theme. The parameters that characterize each song (the initial and final values of the fundamental frequency of each note, the duration of each note, and the time difference between the notes) have variation in a range of less than 7% between the different songs. To obtain the variations of the values in the parameters, we calculated a Gaussian distribution with mean and standard deviation from the song examples for each of the themes. For further details, see [28,29].

In Fig. 4(a), we show the spectrogram of a recorded song of an adult male *Zonotrichia capensis*, while in Fig. 4(b), we
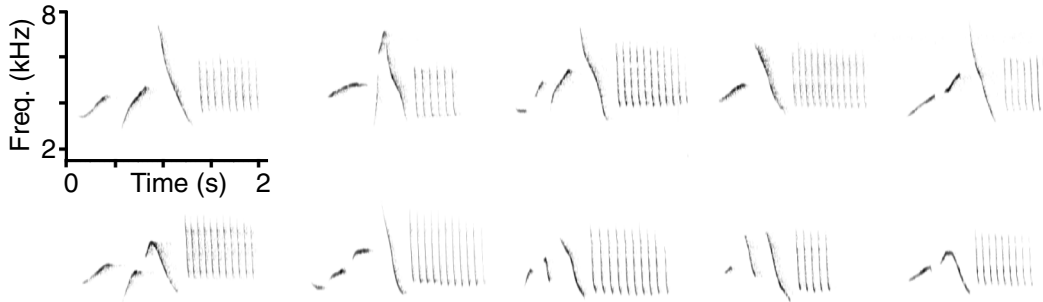
FIG. 5. Spectrograms of the other ten songs used in our study. As in Fig. 4, each of these songs was used as a template to integrate our model in order to produce a synthetic song.

show the spectrogram of its synthetic copy generated with the dynamical system model for vocal production. Figure 5 displays the spectrograms of the other ten songs used as models for our simulated songs. It is worth noting that, even though learning is a part of the process, in this species each bird usually acquires a single song, which, once learned, is consistently repeated with a high degree of stereotypy. In Fig. 6 we show the spectrograms of the ten remaining synthetic songs.

B. Response of territorial *Zonotrichia capensis* adults to real and synthetic acoustic stimuli

As described in Sec. II C, we implemented a protocol for testing synthetic and real songs of *Zonotrichia capensis*. In Fig. 7(a) we show the temporal evolution of the song rate execution by the individuals exposed to the protocol. Prior to the start of the protocol, individuals sang three to four songs per minute. During the period of time in which a recording is played [between minute 2 and 3; black box in Fig. 7(a)], the singing performance increased significantly. Upon completion of playing back the auditory stimuli, the original singing rate was gradually recovered. There is no significant difference between the responses to real and synthetic songs [green dashed line and blue solid line, respectively in Fig. 7(a)], which are significantly higher from the responses obtained using recordings of heterospecific subjects [red dash-dotted line in Fig. 7(a)]. These results are in accordance with previous results that showed that Rufous-collared sparrow subjects responded with a significantly higher rate of song execution to the conspecific playback than to heterospecific playback [10,31,32].

We employed the Welch test to assess the statistical significance of our results. For each protocol (which included the use of real songs, synthetic songs, and heterospecific songs, respectively), and for each minute of the protocol, we created a distribution with the singing rates of the subjects exposed to it. Subsequently, we conducted pairwise comparisons of the distributions, employing the Welch test to assess the null hypothesis of equal means among them. The results are depicted in Fig. 7(b). As expected from Fig. 7(a), between minute three and minute seven of the protocol, the p values for the synthetic-heterospecific and real-heterospecific pairs dropped below a significance level α of 0.05 [olive solid line in Fig. 7(b)], indicating rejection of the null hypothesis. However, throughout all the minutes in the protocols, p values for the synthetic-real pair consistently remained above $\alpha = 0.05$, demonstrating the similarity between the curves describing the evolution of production rates in response to synthetic and real stimuli.

IV. DISCUSSION

A physical model for the avian vocal apparatus that is capable of generating synthetic copies of birdsong allows us to test the hypotheses implicit in its construction. Syntheses generated by low-dimensional models [2] were used to train birds in captivity, isolated from all other auditory stimuli [33]. In sleeping birds, songs generated by low-dimensional models were good enough to elicit highly selective responses in cortical neurons to the bird's own song [5,7] and even activate syringeal muscles [9]. In this work, we show that these low-dimensional models are able to evoke in wild songbirds' behavioral responses similar to those obtained with

FIG. 6. Spectrograms of the synthetic copies generated with the dynamical model of the vocal organ. Each synthetic song copies one recorded song in Fig. 5.
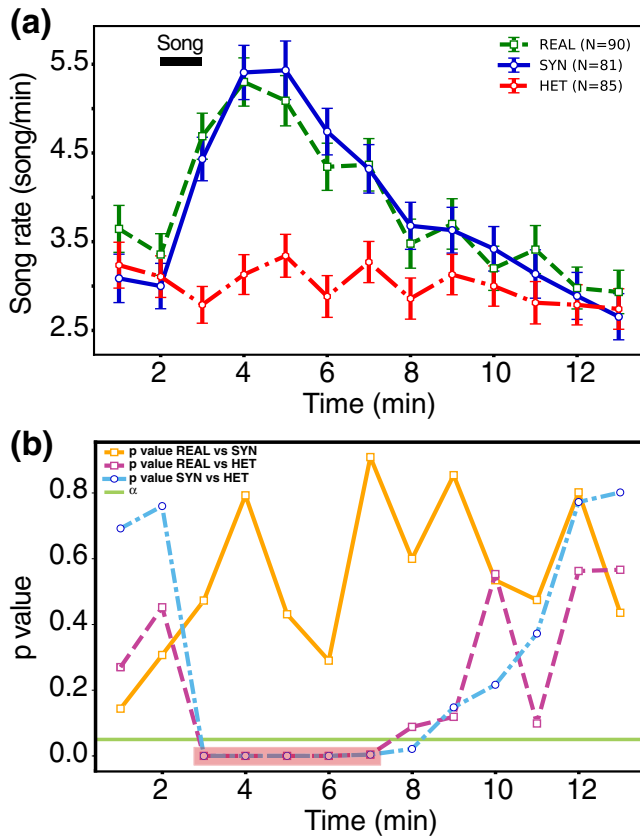
FIG. 7. Temporal evolution of song rates obtained when the birds are exposed to real songs (REAL, green dashed line), synthetic songs of similar frequency modulations (SYN, blue solid line), and heterospecific songs, used as control (HET, red dash-dotted line) (a). These songs were presented between minutes 2 and 3 of the protocol (see black bar "Song"). (b) The p values obtained from a Welch test, for the distribution of singing rates for each of the 13 minutes of the protocols. Each curve corresponds to the comparison of distributions obtained when the protocols were (synthetic, real: yellow solid line), (synthetic, heterospecific: blue light dash-dotted line), and (real, heterospecific: purple dashed line). The results indicate that there was no statistically significant difference in responses between the third and seventh minutes for both synthetic and real stimuli.

recordings of real songs. Measuring the rate of birdsong production in times of territorial defense, we found no significant differences in the responses to real and synthetic songs. In both cases, the response consisted of an increase in the song execution rate, significantly higher than the execution rate measured when the stimuli were heterospecific songs.

Creating low-dimensional models for the physical processes underpinning complex behaviors presents a genuine challenge. Biological systems emerge through evolution, not deliberate design, making it nontrivial to identify a relevant low-dimensional model with interpretable parameters. In this context, a straightforward low-dimensional model founded on minimal assumptions highlights the biologically essential elements responsible for generating complex vocal behaviors. It also provides insights into the extent to which the intricacies of the behavior stem from the nervous system's time-dependent parameters versus the biomechanical aspects.

Furthermore, this work paves the way for manipulating auditory stimuli in an interpretable way, allowing one to address a series of questions that can greatly benefit from the flexibility in the generation of acoustic stimuli permitted by our physical model. Among them, we can mention the testing of the performance hypothesis, which proposes that two specific attributes of singing (the frequency of syllable production and the spectral range of syllables), constitute reliable indicators of the quality of the singer [34]. This influential idea requires stimuli that are not structurally abnormal. The type of stimuli that can be developed with our models allows smooth and continuous variation of the parameters of interest, compatible with the biophysics of the problem. This work provides the appropriate standard for null parametric variation. Another interesting application opens from the result reported in [35], which showed that wild birds can learn via playback, using automatic audio players. Our stimuli can become an invaluable tool in the study of learning in a wild avian population. These examples constitute only two within a wide range of possibilities that open up from the validation that the synthesis generated by interpretable physical models is capable of eliciting responses similar to those obtained through the execution of real singing.

In recent years, different species have been shown to display acoustic features compatible with different nonlinear phenomena: biphonation, nonlinear source-source interaction, and subharmonicities, among them [36–38]. It will be interesting to explore the responses of the individuals of those species to synthetic stimuli failing to reproduce them; an ultimate test of the importance of these subtle yet widely found effects.

[1] P. R. Marler and H. Slabbekoorn, *Nature's Music: The Science of Birdsong* (Elsevier, New York, 2004).

[2] T. Gardner, G. Cecchi, M. Magnasco, R. Laje, and G. B. Mindlin, Simple motor gestures for birdsongs, Phys. Rev. Lett. **87**, 208101 (2001).

[3] R. Laje, T. J. Gardner, and G. B. Mindlin, Neuromuscular control of vocalizations in birdsong: A model, Phys. Rev. E **65**, 051921 (2002).

[4] G. B. Mindlin and R. Laje, *The Physics of Birdsong* (Springer Science & Business Media, Berlin, 2005).

[5] A. Amador and G. B. Mindlin, Synthetic birdsongs as a tool to induce, and listen to, replay activity in sleeping birds, Front. Neurosci. **15**, 647978 (2021).

[6] J. D. Sitt, A. Amador, F. Goller, and G. B. Mindlin, Dynamical origin of spectrally rich vocalizations in birdsong, Phys. Rev. E **78**, 011905 (2008).

[7] A. Amador, Y. S. Perl, G. B. Mindlin, and D. Margoliash, Elemental gesture dynamics are encoded by song premotor cortical neurons, Nature (London) **495**, 59 (2013).

[8] S. Boari, Y. S. Perl, A. Amador, D. Margoliash, and G. B. Mindlin, Automatic reconstruction of physiological gestures used in a model of birdsong production, J. Neurophysiol. **114**, 2912 (2015).

[9] A. Bush, J. F. Döppler, F. Goller, and G. B. Mindlin, Syringeal EMGs and synthetic stimuli reveal a switch-like activation of the songbird's vocal motor program, Proc. Natl. Acad. Sci. USA **115**, 8436 (2018).

[10] D. Shallin Busch, J. C. Wingfield, and I. T. Moore, Territorial aggression of a tropical passerine, Zonotrichia capensis, in response to a variety of conspecific intruders, Behaviour **141**, 1173 (2004).

[11] P. L. Tubaro, Aspectos causales y funcionales de los patrones de variación del canto del chingolo (Zonotrichia capensis), Ph.D. thesis, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1990.

[12] D. E. Kroodsma, Suggested experimental designs for song playbacks, Anim. Behav. **37**, 600 (1989).

[13] R. J. Brooks and J. B. Falls, Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbors and strangers, Can. J. Zool. **53**, 879 (1975).

[14] J. B. Falls and R. J. Brooks, Individual recognition by song in white-throated sparrows. II. Effects of location, Can. J. Zool. **53**, 1412 (1975).

[15] R. J. Brooks and J. B. Falls, Individual recognition by song in white-throated sparrows. III. Song features used in individual recognition, Can. J. Zool. **53**, 1749 (1975).

[16] S. H. Hurlbert, Pseudoreplication and the design of ecological field experiments, Ecol. Monogr. **54**, 187 (1984).

[17] P. K. Stoddard, M. D. Beecher, C. L. Horning, and S. E. Campbell, Recognition of individual neighbors by song in the song sparrow, a species with song repertoires, Behav. Ecol. Sociobiol. **29**, 211 (1991).

[18] P. K. McGregor *et al.*, Design of playback experiments: The thornbridge Hall NATO ARW consensus, in *Playback and Studies of Animal Communication*, edited by P. K. McGregor (Springer US, Boston, MA, 1992), pp. 1–9.

[19] J. L. van Leeuwen, R. Suthers, F. Goller, and C. Pytte, The neuromuscular control of birdsong, Philos. Trans. R. Soc. B **354**, 927 (1999).

[20] F. Goller and O. Larsen, New perspectives on mechanisms of sound generation in songbirds, J. Comp. Physiol. A **188**, 841 (2002).

[21] F. Goller and O. N. Larsen, A new mechanism of sound generation in songbirds, Proc. Natl. Acad. Sci. USA **94**, 14787 (1997).

[22] N. H. Fletcher, T. Riede, and R. A. Suthers, Model for vocalization by a bird with distensible vocal cavity and open beak, J. Acoust. Soc. Am. **119**, 1005 (2006).

[23] Y. S. Perl, E. M. Arneodo, A. Amador, F. Goller, and G. B. Mindlin, Reconstruction of physiological instructions from zebra finch song, Phys. Rev. E **84**, 051909 (2011).

[24] G. B. Mindlin, Nonlinear dynamics in the study of birdsong, Chaos **27**, 092101 (2017).

[25] Arduino, Arduino LLC, 2023.

[26] Atmel 2023, Atmega328p Datasheet, 2023.

[27] https://github.com/LaboratorioSistemasDinamicos/UnifiedSystem_2023-204.

[28] R. A. Bistel, A. Martinez, and G. B. Mindlin, Neural networks that locate and identify birds through their songs, Eur. Phys. J.: Spec. Top. **231**, 185 (2022).

[29] R. Bistel, A. Martinez, and G. B. Mindlin, An analysis of the persistence of Zonotrichia capensis themes using dynamical systems and machine learning tools, Chaos, Solitons Fractals **165**, 112803 (2022).

[30] J. R. Fotheringham, Differences in singing behavior between Rufous-collared sparrows in Costa Rica and Northwestern Argentina, Condor **97**, 821 (1995).

[31] I. T. Moore, H. Wada, N. Perfito, D. S. Busch, T. P. Hahn, and J. C. Wingfield, Territoriality and testosterone in an equatorial population of Rufous-collared sparrows, *Zonotrichia capensis*, Anim. Behav. **67**, 411 (2004).

[32] J. E. Danner, R. M. Danner, F. Bonier, P. R. Martin, T. W. Small, and I. T. Moore, Female, but not male, tropical sparrows respond more strongly to the local song dialect: Implications for population divergence, Am. Nat. **178**, 53 (2011).

[33] T. J. Gardner, F. Naef, and F. Nottebohm, Freedom and rules: The acquisition and reprogramming of a bird's learned song, Science **308**, 1046 (2005).

[34] J. Podos, Birdsong performance studies: Reports of their death have been greatly exaggerated, Anim. Behav. **125**, e17 (2017).

[35] D. J. Mennill, S. M. Doucet, A. E. M. Newman, H. Williams, I. G. Moran, I. P. Thomas, B. K. Woodworth, and D. R. Norris, Wild birds learn songs from experimental vocal tutors, Curr. Biol. **28**, 3273 (2018).

[36] N. H. Fletcher, A class of chaotic bird calls? J. Acoust. Soc. Am. **108**, 821 (2000).

[37] J. F. Döppler, A. Amador, F. Goller, and G. B. Mindlin, Dynamics behind rough sounds in the song of the *Pitangus sulphuratus*, Phys. Rev. E **102**, 062415 (2020).

[38] S. A. Zollinger, T. Riede, and R. A. Suthers, Two-voice complexity from a single side of the syrinx in northern mockingbird *Mimus polyglottos* vocalizations, J. Exp. Biol. **211**, 1978 (2008).