

Dynamical model for the neural activity of singing *Serinus canaria*

Cite as: Chaos **30**, 053134 (2020); <https://doi.org/10.1063/1.5145093>

Submitted: 14 January 2020 . Accepted: 27 April 2020 . Published Online: 18 May 2020

Cecilia T. Herbert , Santiago Boari , Gabriel B. Mindlin , and Ana Amador 

COLLECTIONS

Paper published as part of the special topic on [Instabilities and Nonequilibrium Structures](#)

Note: This article is part of the Focus Issue, instabilities and Nonequilibrium Structures.



View Online



Export Citation



CrossMark

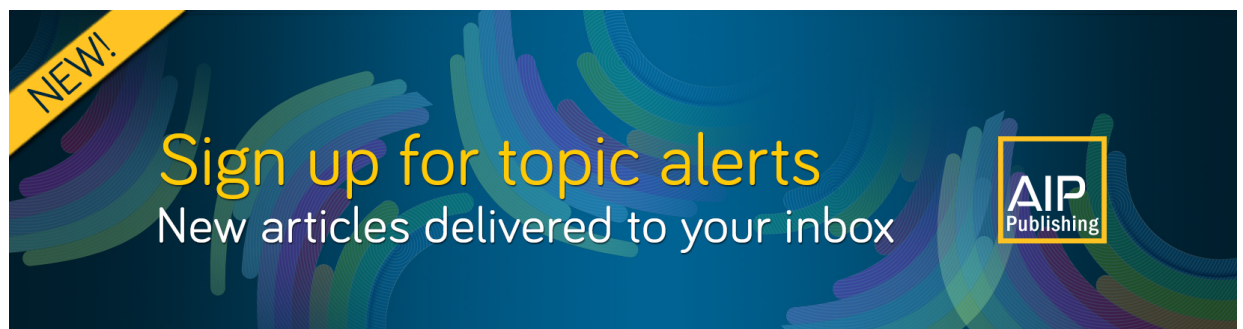
ARTICLES YOU MAY BE INTERESTED IN

[Strong correlations between power-law growth of COVID-19 in four continents and the inefficiency of soft quarantine strategies](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **30**, 041102 (2020); <https://doi.org/10.1063/5.0009454>

[Homoclinic organization in the Hindmarsh–Rose model: A three parameter study](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **30**, 053132 (2020); <https://doi.org/10.1063/1.5138919>



Dynamical model for the neural activity of singing *Serinus canaria*

Cite as: Chaos 30, 053134 (2020); doi: 10.1063/1.5145093

Submitted: 14 January 2020 · Accepted: 27 April 2020 ·

Published Online: 18 May 2020



View Online



Export Citation



CrossMark

Cecilia T. Herbert,  Santiago Boari,  Gabriel B. Mindlin,  and Ana Amador^{a)} 

AFFILIATIONS

Department of Physics, FCEyN, University of Buenos Aires and IFIBA, CONICET, Intendente Güiraldes 2160 (C1428EGA), Pabellón 1, Ciudad Universitaria, Buenos Aires, Argentina

Note: This article is part of the Focus Issue, instabilities and Nonequilibrium Structures.

^{a)}Author to whom correspondence should be addressed: anita@df.uba.ar

ABSTRACT

Vocal production in songbirds is a key topic regarding the motor control of a complex, learned behavior. Birdsong is the result of the interaction between the activity of an intricate set of neural nuclei specifically dedicated to song production and learning (known as the “song system”), the respiratory system and the vocal organ. These systems interact and give rise to precise biomechanical motor gestures which result in song production. Telencephalic neural nuclei play a key role in the production of motor commands that drive the periphery, and while several attempts have been made to understand their coding strategy, difficulties arise when trying to understand neural activity in the frame of the song system as a whole. In this work, we report neural additive models embedded in an architecture compatible with the song system to provide a tool to reduce the dimensionality of the problem by considering the global activity of the units in each neural nucleus. This model is capable of generating outputs compatible with measurements of air sac pressure during song production in canaries (*Serinus canaria*). In this work, we show that the activity in a telencephalic nucleus required by the model to reproduce the observed respiratory gestures is compatible with electrophysiological recordings of single neuron activity in freely behaving animals.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5145093>

Songbirds produce their song by articulating the neural activity of areas in the brain dedicated to song production with the respiratory system and the vocal organ. Functional studies of the song system in the avian brain have been greatly focused on *single units*, i.e., trying to understand how this system works by measuring the activity of many individual neurons. From a technical standpoint, it is not yet possible to measure such large ensembles from multiple brain areas simultaneously, as this would involve voltage recordings of hundreds of thousands of neurons simultaneously. In this work, we present a different approach. Using a model of the neural architecture involved, we found which global neural activity patterns give rise to an output compatible with motor commands used during song production. We then recorded neural activity of individual neurons in singing birds in a particular brain region and found that the neural activity patterns provided by the model were compatible with the experimental results. In this way, models constitute a framework that could guide further experiments in the field to elucidate, from a macroscopic perspective, how different parts of the song system articulate to give rise to a delicate vocal behavior.

I. INTRODUCTION

Birdsong production requires an exquisite combination of motor gestures (Suthers and Margoliash, 2002). In order to generate song, a bird must control the configuration of its vocal organ, the syrinx, as well as the activity of the respiratory system (Gardner *et al.*, 2001). The animal generates large pressure pulses during which energy is transferred to the labia located at the juncture between the bronchi and the trachea, inducing oscillations that modulate the airflow to produce sound (Mindlin and Laje, 2006). Songbirds account for about half of the known bird species. These birds learn their songs after some level of exposure to a tutor. Their brain presents a set of interconnected neural nuclei known as the “song system,” with nuclei specifically dedicated to the learning, production, and maintenance of song (Nottebohm *et al.*, 1990). A set of respiratory nuclei receives inputs both from regions of the brainstem (for all bird species) as well as from cortical regions (in songbirds) (Ashmore *et al.*, 2005). In order to generate song, the bird has to excel in coordinating the activity of several respiratory muscles with syringeal muscles, a delicate action achieved through a sophisticated process of motor learning

which involves specific cortical nuclei (see [Mooney, 2009](#) for a review).

From first principles, it is complex to link the patterns of neural activity in those brain regions to the actual motor activity that occurs in the peripheral organs which ultimately generate the song ([Ott and Antonsen, 2008](#)). For this reason, we have been working over the last years with phenomenological, additive models (Wilson–Cowan equations, [Wilson and Cowan, 1973](#) and [Hoppensteadt and Izhikevich, 2012](#)), which aim to provide a description of the neural activity of the different nuclei of song system in terms of average activities of large neural ensembles ([Alonso et al., 2015, 2016](#)). The resulting average activities in the different nuclei for the production of specific song syllables give us an insight on what the coordinated activity of neurons may be, and is a step forward to discover the motor representation of this behavior in the brain, i.e., its *motor code*.

The experimental strategy is also complex. It is difficult to record single units in behaving individuals and extremely hard to link such overwhelmingly sparse information with the actual instructions controlling the periphery. An alternative that overcomes the need for careful electrode placing and single unit sorting is to record several neurons simultaneously with low impedance electrodes ([Buzsáki et al., 2012](#)), but results from this method are difficult to interpret due to the variety of neuron types involved ([Mooney, 2000](#) and [Rauske et al., 2003](#)) and the amount of global neural activity in the signal.

Furthermore, the physiological instructions that control the biomechanical devices to produce behavior stem from a delicate equilibrium between neural activity and low dimensional dynamics of the periphery ([Tytell et al., 2011](#) and [Takahashi et al., 2015](#)). In the case of song production, physiological instructions arise from the interaction of the nervous system with the syringeal and respiratory motor activity. In canaries, for example, the respiratory activity generated during song production (stereotyped air sac pressure traces) has a structure that can be reproduced with simple dynamical elements ([Trevisan et al., 2006](#) and [Alonso et al., 2010](#)).

In order to find a quantitative description that could link the physiological instructions used during song production with some of what we know about the song system, we recently built a dynamical model that (1) Is consistent with the basic architecture of the song system, (2) Can fit the respiratory output of the song system, and (3) Can make precise predictions on the neural activity of an accessible part of the song system ([Alonso et al., 2015, 2016; Dima et al., 2018a; 2018b](#)). In this work, we report for the first time single-unit neural activity recorded in singing canaries (*Serinus canaria*), in one of the brain areas described by our dynamical model, for which we found a good correspondence between neural recordings and neural modelling.

II. EXPERIMENTAL METHODS

Extracellular activity in neuronal nucleus HVC (used as proper name) and sound were recorded simultaneously in six adult canaries (*S. canaria*) during song production. Additionally, for one of these birds, air sac pressure and audio were recorded during song. Each bird was housed individually but could see and hear other canaries intermittently to promote singing. Birds were recorded for as long as

they presented good physiological signals and high rates of singing. This can be very variable across birds. The recording periods in different birds were between 7 and 47 days, with recording sessions lasting on average 7 h per day. Experiments were performed in accordance with a protocol approved by the University of Buenos Aires (FCEN-UBA) Institutional Animal Care and Use Committee (CICUAL). Neural, audio, and/or air sac pressure were monitored continuously by the recording equipment (see below) and recorded only when triggered by sustained sound (usually, song). For electrophysiological recordings, we used an array of tetrodes made inhouse from polyimide-coated tungsten wires ([Henze et al., 2000](#)), mounted on a manual microdrive. Recordings were performed on Intan Technologies' RHD2000 USB interface board at 30 kS/s, amplified on an RHD2132 16-channel amplifier board. The neural signal was digitally high-pass filtered (300 Hz cutoff frequency, third order Butterworth filter) offline in Matlab (MathWorks, www.mathworks.com).

Audio was acquired using a 20 Hz–20 kHz electret microphone mounted on a Maxim MAX 4466 amplifier board with adjustable gain placed inside the sound attenuating chamber where the bird was. The audio signal was recorded on an analog input of the RHD2000 USB interface board and was sampled at the same rate as the neural signal. The audio signal was digitally high-pass filtered (300 Hz cutoff frequency, third order Butterworth) in Matlab. Spectrograms of the audio signal were computed using Matlab with a 10 ms gaussian window, $\alpha = 3$, and temporal overlap of 95%.

Air sac pressure was recorded using a flexible cannula and a miniature piezoresistive pressure transducer (Fujikura model FPM-02PG). For further details, see, for example, [Amador and Margoliash \(2013a\)](#). Pressure data were amplified with custom hardware and acquired into an analog input of the RHD2000 USB interface board. The pressure signal was digitally low-pass filtered (300 Hz cutoff frequency, third order Butterworth filter) in Matlab.

III. THE MODEL

A. Additive models

The song system is a set of interconnected areas of the brain. Each area is called a neural nucleus and is composed by interconnected neurons highly packed in a defined region. The number of neurons in each region ranges from tens of thousands to hundreds of thousands, depending on the nucleus ([Ward et al., 1998](#)). A way to represent the neural activity is shown in [Fig. 1\(a\)](#) in which each dot indicates the time when an action potential (spike) has occurred. In [Fig. 1\(a\)](#), we represent the activity of ten different units spiking in an interconnected neural pool, and, therefore, the increase of spiking rate occurs in a coordinated way and is represented as a well-defined peak in the average activity [see [Fig. 1\(b\)](#)]. In this way, the variables of our dynamical model will be the average activities of the neurons, represented in [Fig. 1\(b\)](#) ([Alonso et al., 2015](#) and [Dima et al., 2018b](#)).

Additive neural network models are probably the simplest empirical models used in computational neuroscience and are designed to capture very basic features of the average activity of neural populations ([Hoppensteadt and Izhikevich, 2012](#)). They can be used to describe the average dynamics of interconnected excitatory and inhibitory neural populations. If x and y stand for the average activities of the excitatory and inhibitory populations, respectively,

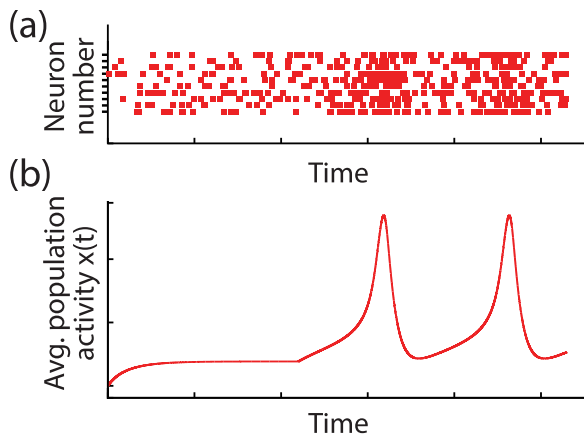


FIG. 1. Neuronal activity and additive model variables. (a) Temporal evolution of the neural activity of different units. In this plot, each point represents a spiking event and the vertical axis separates the activity of ten different neurons from a population. (b) The variables and outputs of rate models are the average activity of a population of neurons $[x(t)]$. In this example, we show that synchronous events of high population wide activation are translated as peaks in $x(t)$, while low-rate and nonsynchronous firing means that $x(t)$ takes a small, near-zero value. Adapted from Dima *et al.* *Papers Phys.* **10**, 1–16 (2018b).

their dynamics will be ruled by

$$\frac{dx}{dt} = -x + S(\rho_x + a_1x + a_2y), \quad (1a)$$

$$\frac{dy}{dt} = -y + S(\rho_y + a_3x + a_4y), \quad (1b)$$

with the vector field defined in terms of a Sigmoidal function,

$$S(u) = \frac{1}{1 + e^{-u}}, \quad (1c)$$

with $\rho_x(t)$, $\rho_y(t)$ being the inputs to the excitatory and inhibitory populations, that may change with time t . The constants a_i ($i = 1, \dots, 4$) describe the architecture of the neural network. A representation of this system is shown in Fig. 2(a). Notice that although it is a rather simple model, it can display very rich dynamics: depending on the values of the input parameters $\rho_x(t)$ and $\rho_y(t)$, the dynamics can be stationary, excitable, or oscillatory [regions I and IV, II, and III in Fig. 2(b), respectively]. In Fig. 2(b), we display the bifurcation diagram of the system in terms of $\rho_x(t)$ and $\rho_y(t)$, and in Fig. 2(c) the dynamics in each region. Since this basic architecture is capable of displaying oscillations (i.e., alternating excitatory and inhibitory activity patterns), it is known as a neural oscillator. This basic neural unit, the neural oscillator, was used in the literature to build motor pattern models, and we have previously reported how to assemble neural oscillators to build a model for the song system (e.g., Trevisan *et al.*, 2006). Moreover, we have also explored that the dynamics of these phenomenological models can be recovered by an extension of the Ott–Antonsen method (Ott and Antonsen, 2008 and Roulet and Mindlin, 2016).

One of the brain areas that we will describe by means of a neural oscillator is the expiratory nucleus of the song system, whose output

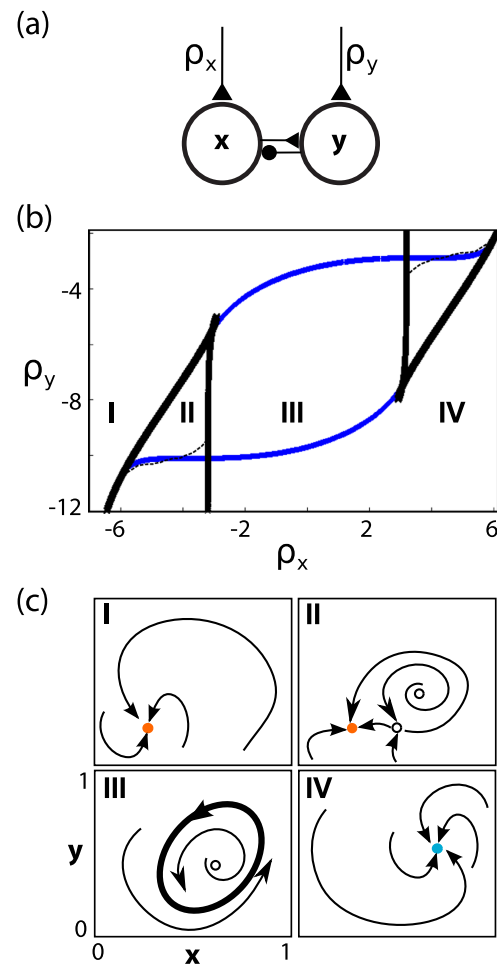


FIG. 2. Bifurcation diagram and dynamical regimes for interconnected excitatory–inhibitory additive models. (a) Diagram representing two interconnected populations of excitatory (x) and inhibitory (y) neurons, with parameters (ρ_x, ρ_y) representing the inputs to x and y , respectively. (b) Bifurcation diagram for the system of Eqs. (1a)–(1c) in terms of the inputs (ρ_x, ρ_y) . Four qualitatively distinct dynamical regimes are achievable: (I) quiescent state, (II) excitable regime, (III) spiking regime, and (IV) constant activity. Regions are delimited by the multiple bifurcation sets. The blue (black) lines are the loci of the multiple Hopf (saddle node in limit cycle or SNILC) bifurcation sets, while the dashed lines correspond to homoclinic (or saddle separatrix loop) bifurcation sets. (c) Phase diagrams for each identified region in (b), showing one stable fixed point in (I), an excitable regime with a stable fixed point, a saddle point and an unstable fixed point in (II), an oscillatory regime due to the presence of a limit cycle in (III) and a stable fixed point in (IV). Adapted from Dima *et al.*, *Papers Phys.* **10**, 1–16 (2018b) and Borisyuk *et al.*, *Biol. Cyber.* **66**(4), 319–325 (1992).

will be responsible for the large expiratory gestures that are necessary in order to push airflow between syringeal labia and generate sound. This nucleus is built out of excitatory and inhibitory neurons, and it has two inputs: one from some region of the brainstem and a second input from the telencephalon. The telencephalon is assumed

to enrich part of the neural activity originated in the brainstem, processing it through an interconnected array of neural nuclei. Since the output of this nucleus is closely related to an observable which is easy to measure (the air sac pressure during song production), over the last years we have built a dynamical system representing the different areas of the song system, to test whether the model was capable of reproducing the respiratory gestures of a singing bird. This model was reported in the literature (e.g., [Alonso et al., 2015, 2016](#) and [Dima et al., 2018b](#)). It has a number of compartments compatible with the known anatomy of the avian brain. Each nucleus has excitatory and inhibitory neural populations as reported, and it was capable of synthesizing realistic air sac pressure patterns and sound for canary song. Its schematic is displayed in [Fig. 3](#).

B. The program

With these elements, we want to explore which neural activity patterns in the telencephalon can reproduce the respiratory outputs observed during birdsong production. We are going to choose two regions to assess physiological activity during singing to experimentally validate our dynamical model. As a proxy for the activity in the respiratory nuclei of the song system, we can record the air sac pressure of a singing bird ([Suthers et al., 1996](#) and [Trevisan et al., 2006](#)). Additionally, HVC (proper name) is a telencephalic nucleus that is close to the scalp, and therefore it is among the easiest regions of the

song system from which to record neuronal activity. HVC is a neural nucleus necessary for birdsong production.

In particular, by measuring in these two extremes of the song system, we plan to test whether the motor gestures (air sac pressure and syringeal muscle tension traces) used in the production of song require temporally heterogeneous (i.e., non-uniform) activity patterns in the telencephalic nucleus HVC. The rationale behind testing this hypothesis is that analysis of the neural activity of a group of singing zebra finches (*Taeniopygia guttata*) selected specifically because of their particularly simple songs, revealed a preference for spiking at particular motor instances ([Amador et al., 2013b](#)). These significant motor instances are specific moments linked to transitions in the motor gestures used during song production. The firing preference at specific instances led to the hypothesis that HVC activity could be non-uniform throughout the song. In this work, we focus on unveiling which non-uniform neural activity patterns give rise to respiratory patterns compatible with experimental recordings of air sac pressure and test this by recording neural activity in HVC.

Preference for spiking at specific instances was also observed in songs of juveniles, and to some degree, in simple syllables of adult zebra finches ([Okubo et al., 2015](#)). These observations, on the other hand, seem to contradict other reports that describe the activity in HVC as uniform during the whole song ([Lynch et al., 2016](#) and [Picardo et al., 2016](#)). One way to reconcile these seemingly opposing views is that uniform activity was recorded in birds singing very

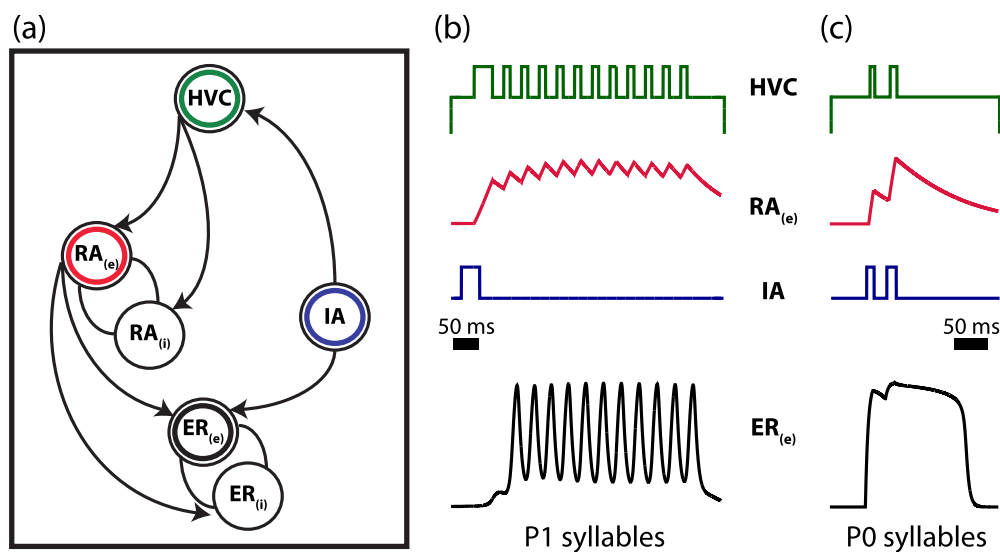


FIG. 3. Architecture of the model and simulations of pressure patterns for two canary song syllable types. (a) Minimal architecture of neuronal populations compatible with the birdsong motor pathway. In this model, an initiating area (IA) in the brainstem is responsible for activating the motor program. IA projects directly to an expiratory-related area (ER) and indirectly through the cortex by projecting the initiating pulse to the neural nucleus HVC (proper name). Activity in HVC is modeled by a function, F_{HVC} , which is compatible with activity in IA and with the hypothesis that HVC activity presents activation peaks at specific moments in the motor program rather than being uniformly active throughout. HVC then projects to nucleus RA (robust nucleus of the arcopallium), responsible for processing and enriching the activity coming from HVC according to Eqs. (2c) and (2d) of our model (see text). Finally, both converging inputs to ER give rise to the model output: a simulated pressure pattern that can be compared to physiological recordings. (b) and (c) Model variables necessary to produce expiratory activity compatible with P1 syllables (b) and P0 syllables (c). P1 syllables correspond to short and periodic expiratory pulses. In the top panel, F_{HVC} required to produce the ER_(e) output is shown. Note the presence of a periodic behavior in HVC activity, which corresponds to the periodic expiratory output. P0 syllables, on the other hand, are longer syllables characterized by a short high-frequency pulse and a long whistle. In this case, F_{HVC} requires only two activity peaks to produce an expiratory output compatible with measured pressure patterns.

complex songs, as it is usually the case with zebra finches. Therefore, the density of significant motor instances is so high, that it is difficult to disambiguate that model from another one postulating a somewhat uniform neural coding. Recently, it was proposed that a suitable species to disambiguate between the two models is the canary (*S. canaria*). The song of canaries is made up of units called syllables that they repeat a variable amount of times to form phrases. Canaries present some syllables that do not require very complex motor patterns (Lassa Ortiz *et al.*, 2019). Actually, any oscine species capable of generating long whistles (i.e., sounds with constant fundamental frequency), could be an adequate candidate for the task. For such species, syllables generated with simple gestures would require very different HVC activity patterns than the ones found in zebra finch song.

C. The computational model

The birdsong production circuit involves the telencephalic area HVC that projects to the robust nucleus of the arcopallium (RA), which projects to nuclei that control respiration. In our model, we are going to describe this as an expiratory related area (ER). Neural nuclei in the brainstem send projections to respiratory centers and to the thalamus that further projects to HVC. We are going to describe these brainstem nuclei as a generic initiating area (IA). A representation of this is shown in Fig. 3(a) and the mathematical implementation of the model reads as follows:

$$\frac{de_{er}}{dt} = 249.5(-e_{er} + S(\rho_{eer} + \alpha_{eer,eer}e_{er} + \alpha_{eer,ier}i_{er} + \alpha_{eer,era}e_{ra} + \alpha_{eer,FIA}F_{IA}(t))), \quad (2a)$$

$$\frac{di_{er}}{dt} = 249.5(-i_{er} + S(\rho_{ier} + \alpha_{ier,eer}e_{er} + \alpha_{ier,ier}i_{er} + \alpha_{ier,era}e_{ra})), \quad (2b)$$

$$\frac{de_{ra}}{dt} = 20(-e_{ra} + S(\rho_{era} + \alpha_{era,era}e_{ra} + \alpha_{era,ira}i_{ra} + \alpha_{era,FHVC}F_{HVC}(t))), \quad (2c)$$

$$\frac{di_{ra}}{dt} = 20(-i_{ra} + S(\rho_{ira} + \alpha_{ira,era}e_{ra} + \alpha_{ira,ira}i_{ra})), \quad (2d)$$

where the variables e_{er} , i_{er} , e_{ra} , i_{ra} stand for the excitatory and inhibitory populations in the expiratory related area (e_{er} and i_{er}) and nucleus RA (e_{ra} and i_{ra}), respectively. The temporal function F_{IA} stands for the activity in the initiating area (IA). As it can be seen from Eq. (2a) in our dynamical system, F_{IA} constitutes a direct input to the expiratory related area. For the two syllables analyzed in this work, F_{IA} is displayed with a blue line in Figs. 3(b) and 3(c). The parameters in our model are chosen in such a way that each area (ER and RA) can display a variety of dynamical regimes (see Fig. 2). These equations were numerically integrated using a fourth order Runge Kutta method (Press *et al.*, 1988). The codes used for the simulations shown in this manuscript are openly available on GitHub (Herbert *et al.*, 2020).

The temporal function F_{HVC} stands for the activity in HVC, which we propose in such a way that it is consistent with both the proposed activity in the initiating area, and the hypothesis that HVC

neural activity codes sparse significant motor instances (i.e., activity is not uniform during the motor program). For the two syllable types considered in this work, F_{HVC} is displayed as a green line in Figs. 3(b) and 3(c). This activity is the superposition of a delayed copy of the activity in the IA and, in the case of P1 syllables, a periodic pattern. In the literature, we have shown that P1 syllables could be achieved with or without this periodic component (see Alonso *et al.*, 2015, 2016, respectively). RA will then process the function that we propose for HVC according to Eqs. (2c) and (2d) of our dynamical system. The output of RA (e_{ra}) will constitute the indirect input to the expiratory related area. It is the purpose of our simulations to explore which HVC activity patterns can generate, with our model, respiratory gestures qualitatively similar to the ones we record in singing canaries. Those patterns will be then compared with our electrophysiological recordings of HVC activity.

In this work, we used conjectured temporal functions F_{HVC} in HVC, but notice that both temporal functions used to generate P1 and P0 syllables could be obtained if HVC is itself modeled as a set of excitatory and inhibitory populations. For example, a neural oscillator in an excitable regime can be forced with pulses from IA, giving rise to delayed pulses in HVC. Alternatively, a neural oscillator can be placed in a region of the parameter space where a stationary attractor coexists with a periodic solution and a pulse from the IA can take HVC from a stationary state to an oscillatory one. In other words, the conjectured HVC patterns could be solutions of HVC populations, if modeled themselves as neural oscillators.

D. Neural activity compatible with respiratory gestures

In Fig. 3, we show the simulations for two different respiratory patterns used during the generation of canary song. In Fig. 3(a), we display the areas (neural nuclei) involved in our model. The black circles at the bottom of the figure represent the excitatory and inhibitory populations of the expiratory related area (ER). The blue circle represents an initiating area of the brainstem (IA); the green circle represents the telencephalic nucleus HVC; and the last two circles on the left stand for the excitatory and inhibitory populations of the telencephalic nucleus RA. Notice that both RA and the IA project to the expiratory related area. In Fig. 3(b), we display the set of activity patterns that lead to the generation of very simple, almost harmonic pressure fluctuations (P1 syllables, see Figs. 1 and 2 of Allende *et al.*, 2010).

Notice that in the literature we have shown that our model of the song system can display solutions similar to the pressure patterns used during song production in different ways. For example, in Alonso *et al.* (2016), we have shown a similar HVC activity as the one presented in Fig. 3(b): periodic HVC activity related to periodic respiratory activity. Conversely, in Alonso *et al.* (2015), we have shown that a constant activity in HVC could be represented as a constant value of ρ_{eer} that would put the respiratory neural oscillator in a region of the parameter space where oscillations occur [see Fig. 2(c), region III], and therefore generate oscillatory activity in the respiratory area. In this work, we propose to measure HVC activity to disambiguate between these different possibilities.

In Fig. 3(c), we display the pressure pattern typically used in the generation of the characteristic canary whistle (P0 syllable). In both

TABLE I. ER-related parameters used in the simulations of P1 syllables (P1) and P0 syllables (P0). These parameters define network properties such as the basal level to which activity is mounted (ρ) and the strength of connection to excitatory and inhibitory areas (parameters α). Subscripts (e) and (i) further indicate whether the parameter corresponds to a connection to excitatory (e) or inhibitory (i) neurons. The remaining letters indicate nuclei (e.g., ER, RA, HVC).

Syll.	ρ_{eer}	$\alpha_{eer,cer}$	$\alpha_{eer,ier}$	$\alpha_{eer,era}$	$\alpha_{eer,F}$	ρ_{ier}	$\alpha_{ier,cer}$	$\alpha_{ier,ier}$	$\alpha_{ier,era}$
P1	-6	10	-10	12	1	-8	10	2	4
P0	-7.5	9	-1	9	1	-11.5	10	2	0

cases, the blue and green time series are the *proposed* temporal functions representing the averaged activities in the IA and HVC areas, respectively, while the red and black time series correspond to the computed activities of the excitatory neurons in RA, and the activity of the excitatory neurons in the expiratory related areas, respectively. These are computed using our dynamical system described in Sec. III C [Eqs. (2a)–(2d)].

For the case in Fig. 3(b), HVC activity consists of a set of periodic fluctuations, mounted on a continuous level of activity. This, in turn, elicits in RA a time series that builds up, which then drives the neural oscillator representing the expiratory related area into an oscillating regime. The parameters that define the neural oscillators of our model to generate this syllable are listed in Tables I and II.

For the case displayed in Fig. 3(c), the two bursts in the initiating area induce two delayed bursts of activity mounted on a continuous level of activity in HVC. The continuous level of activity in HVC accounts for the presence of some non-zero average activity from within the neural nucleus. The resulting F_{HVC} triggers a transient build-up of RA activity, which induces a large and wide peak in the expiratory related area. Dynamically, this is due to the existence of a non-zero fixed point for the activity of the excitatory population in ER. As activity in RA starts to decrease, the fixed point in the expiratory related area disappears in a saddle node bifurcation, and the system returns to its quiescent state. The parameters that define the neural oscillators of our model to generate this syllable are listed in Tables I and II. Here, we are not going to focus in the relative delays between signals in different neural nuclei as in previous work it has been shown that the relative delays can be controlled with different synaptic weights (for example, see Fig. 2 in Dima et al., 2018a). Specifically, it has been shown that a similar neural architecture to the one presented here is able to achieve

TABLE II. RA-related parameters used in the simulations of P1 syllables (P1) and P0 syllables (P0). These parameters define network properties such as the basal level to which activity is mounted (ρ) and the strength of connection to excitatory and inhibitory areas (parameters α). Subscripts (e) and (i) further indicate whether the parameter corresponds to a connection to excitatory (e) or inhibitory (i) neurons. The remaining letters indicate nuclei (e.g., ER, RA, HVC).

Syllable	ρ_{era}	$\alpha_{era,era}$	$\alpha_{era,ira}$	$\alpha_{era,FHVC}$	ρ_{ira}	$\alpha_{ira,era}$	$\alpha_{ira,era}$
P1	-5.25	10	-10	1	-5	10	2
P0	-3	6	-3	1	-6	6	6

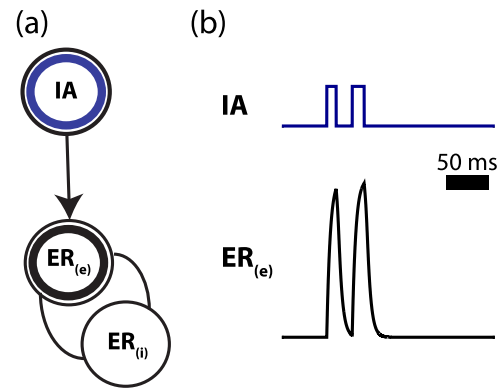


FIG. 4. The cortical loop through HVC and RA is required to simulate complex syllables in canary song. (a) Using the same system parameters as before for P0 syllables but excising the cortical loop through HVC and RA that converges back to ER, the system cannot produce complex syllables such as the P0 syllables shown in Fig. 3(c). (b) In this reduced system, as IA feeds a simple pulse into ER, ER can only produce a short expiratory pulse for each IA input pulse.

delayed synchronization, zero lag synchronization and even anticipated synchronization depending on the synaptic weights used (see Fig. 6 in Dima et al., 2018a).

To illustrate the importance of the telencephalic nuclei in the generation of complex respiratory pulses, in Fig. 4 we show a simulation in which we used the same parameters used to generate a P0 syllables with the neural oscillator representing the expiratory related area, but with no connection from IA to HVC (in this way we modeled the elimination of the rest of the circuit, i.e., RA does not project back into the expiratory related area). The two brief pulses now give rise to two simple expiratory pulses.

Notice that we have generated different syllables with different parameters of the dynamical system representing the song system. Since many of these parameters represent connectivity strengths, we interpret that different neural populations are involved.

IV. PHYSIOLOGICAL RECORDINGS

To put our model to the test, we recorded neural activity extracellularly from nucleus HVC during song production in freely behaving canaries. Figure 5 shows examples of the raw data acquired during singing of two types of syllables. For this work, we selected two syllable types that were sung by all the birds in our study while their neural activity was recorded. These types of syllables are present ubiquitously in a canary repertoire and are generated with simple and consistent motor gestures. The first type is called P1, and they are produced by pressure pulses that display a sinusoidal shape (see Alliende et al., 2010), see bottom of Fig. 3(b) for modeled P1 pressure pattern and Fig. 6(a) for an experimental recording). The second type, called P0 syllables, are characteristic whistles of constant frequency, which have a long expiratory gesture [see bottom of Fig. 3(c) for the modeled P0 pressure pattern and Fig. 6(b) for an experimental recording]. Within and across individuals, syllables of the same types are generated with similar expiratory pressure pulses that cluster in a well-defined group (Alliende et al., 2010 and

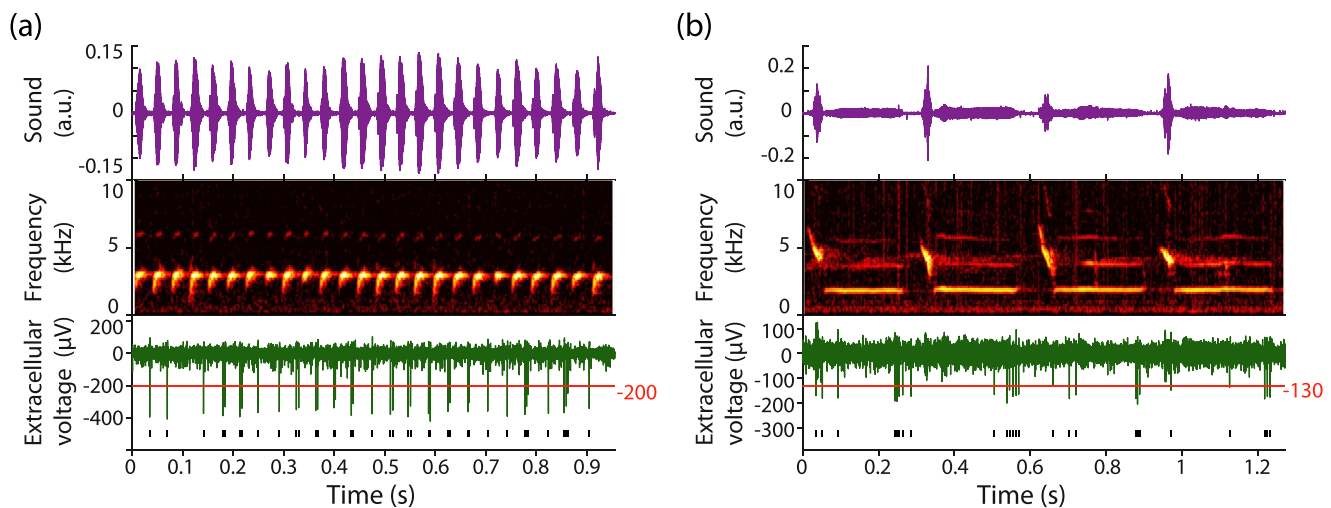


FIG. 5. Song and neuronal recordings from singing canaries. Examples of two syllable types: (a) P1 and (b) P0 syllables. The top panel shows the recorded sound pressure, the middle panel, the spectrogram of that sound, and the bottom panel, the extracellular neural activity recorded during song. The red horizontal line indicates the threshold used to isolate the spiking activity of an individual neuron and the timestamps of its detected spikes using that threshold are shown as black vertical lines below the neural trace.

Dima *et al.*, 2018b). Each type has a specific syllable repetition rate. In this way, just recording and analyzing sound is enough to predict the motor gestures used to generate P1 and P0 syllables. Other syllables not selected for this work may require more diverse and complex motor control and therefore, it is not so clear how to group them together in a well-defined syllable group.

Figure 5(a) shows a phrase of P1 syllables. The top panel shows the recorded audio: the syllables are brief (26.1 ms on average in this example) and very similar across repetitions. The sound spectrogram in the middle panel shows that this particular syllable is an up-sweep, and each syllable spans roughly the same frequency range. The bottom panel shows the high passed filtered neural activity (spikes) recorded in HVC while this phrase was being sung. An individual neuron that can be separated with a threshold of $-200 \mu\text{V}$ fires a single spike or a burst of spikes at each repetition of the P1 syllable in the phrase, at a specific instance within each syllable. We found this rhythmic firing of phasic neurons with syllabic frequency across different P1 phrases in different birds. Figure 5(b) depicts the same information for a phrase of four P0 syllables. Here the audio consists of longer syllables (average 274.7 ms in this example), and the spectrogram shows that they contain a short, high frequency down-sweep note and a long whistle (a tonal sound at constant frequency). These acoustic properties are consistent across all the P0 syllables. The neural data also show a neuron (threshold $-130 \mu\text{V}$) that spikes preferentially at a given moment during the syllable. In this example, it is at the end of the whistle. We also found phasic neurons firing at specific instances within the syllable in other P0 syllables and in different birds. The spike timestamps for each neuron are shown as black vertical lines under the neural traces. The relative latencies of each timestamp with respect to the syllable or phrase onset are shown in the grouped raster plots of Fig. 6.

In order to analyze how phasic neurons spike during P1 and P0 syllables, we grouped the spikes obtained from the raw data after thresholding across phrases sung by different birds. To isolate single units, we considered the amplitude and the shape of the spikes as a characteristic feature of each neuron. In canary song, we can take advantage of the classification of syllables by type to group different syllables from different birds. This is shown in Fig. 6. Each syllable was normalized to have the same length and the relative position of the timestamps of the spikes within the syllable were also normalized to the new uniform length. For P1 syllables, we normalized the duration of each syllable to 70 ms between onsets. In the case of P0 syllables, to account for the variable length of the sound segments within the syllable, we normalized the length of each segment of the syllable to the average length across animals. These were 30.0 ms for the first note, 17.3 ms for the gap between notes, and 222.7 ms for the second note, yielding a uniform length of 270 ms.

Even though canary phrases can contain more syllables [see as an example Fig. 5(a)], we analyzed only the first 12 P1 syllables of each phrase because this is the maximum amount present in all the repetitions these birds sang. The normalized spike timestamps were aligned to the onset of the phrase or syllable and stacked to form a raster plot as shown in Fig. 6. In the top panel of Fig. 6(a) we show a raster plot of the activity of seven neurons in eight unique P1 phrases from four different canaries. Each line corresponds to one repetition of the phrase, extracted from the raw data as shown in the example of Fig. 5(a), and different neurons are indicated by different marker colors. We analyzed up to two phrase repetitions per neuron. The histogram in the middle panel shows the grouped activity of the same neurons in the raster plot. We used a kernel density estimation (ksdensity function in Matlab) with bin size 15 ms, normalized to the maximum bin spike count to smooth the histogram. For P1 syllables, we found one peak of grouped neural activity per syllable,

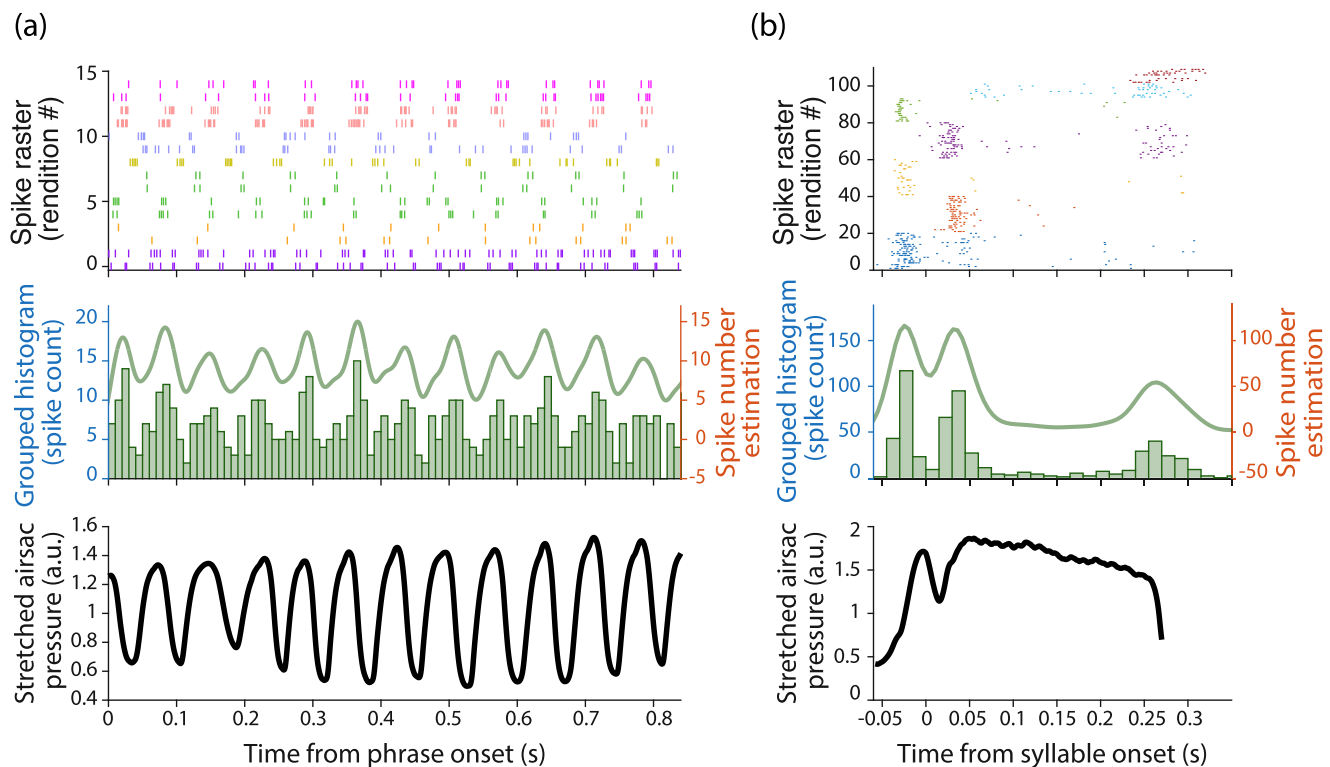


FIG. 6. Grouped neural activity data across different birds and example air sac pressure trace. (a) Top panel: raster plot of spiking activity of seven neurons in eight unique P1 phrases from four different animals. Only the first 12 syllables in each phrase are shown. On the vertical axis, each line corresponds to one repetition of a phrase and each color represents an individual neuron. Middle panel: grouped activity histogram for the neurons shown in the raster plot and smoothed histogram curve. Bin size = 15 ms. Bottom panel: example of the air sac pressure recorded in one animal during the production of a P1 phrase. (b) Same information as in (a), the raster plot shows seven neurons in five different P0 syllables from five animals. The pressure trace in the third panel corresponds to a representative example recorded during production of a P0 syllable in the same animal as in (a).

in a rhythmic fashion across the phrase. This is clearly represented by the smoothed histogram curve and supports the proposed HVC neural activity in the simulations of Fig. 3(b). The bottom panel shows an example of the air sac pressure used to generate this phrase in one animal, in which there is one expiratory pulse per syllable. This trace is normalized in length as was done for neural data, only using linear interpolation. The top panel in Fig. 6(b) shows the raster plot of neural activity per syllable of seven neurons in five different P0 syllables from five canaries. We analyzed up to 20 syllable repetitions for each neuron. The histogram and smoothed histogram curve in the middle panel show well-localized neural activity, in particular, instances during the syllable: before each note, in accordance with the model simulations in Fig. 3(c). Some recorded neurons also showed a more extended activity toward the end of the syllable. This is indicated by the less prominent and wider peak in the smoothed histogram curve. The bottom panel shows an example from one bird of the air sac pressure generated to produce P0 syllables, showing two short pressure pulses within the same expiratory pulse.

The experimental results from all the birds in this study support the results of the simulations of our model, while also point out that

some amount of neuronal activity may be required to terminate the syllable, which was not considered in our simulations.

V. DISCUSSION

Birdsong production requires the precise control of the respiratory system, which is responsible for the generation of large air sac pressure pulses to generate sound. In recent years, the diversity of respiratory gestures canaries use to generate their characteristic repetitive syllables has been interpreted as the sub-harmonic solutions of a periodically forced 2d-normal form (Alonso *et al.*, 2010). This extremely simple model was proposed by first identifying dynamical elements that allow reproducing the shape of the pressure pattern corresponding to one specific syllable type. Then, the same dynamical system, with other parameter values, was shown to be capable of reproducing the respiratory patterns needed for the other syllable types (Alonso *et al.*, 2009). Despite this progress, establishing the connection between the neural substrate responsible for controlling the respiration during song production and the actual respiratory gestures is not trivial. Many factors contribute to making this task difficult: the lack of a comprehensive statistical theory of out

of equilibrium units, the existence of several classes of units (excitatory and inhibitory classes of neurons, and their subpopulations), and the complexity of the neural architecture involved. In other words, even if we have evidence suggesting that the average activity of the neural populations of the song system must obey some precise dynamical system, it is difficult to obtain it from first principles.

Recently, a model for canary song production was proposed whose variables are the average activities of the neural populations within the different nuclei of the song system (Alonso *et al.* 2015, 2016). Since part of this system is a neural population responsible for controlling expiration, the model was tested by its ability to reproduce the respiratory gestures previously recorded in canaries during song production. The model was built consistently with relevant aspects of the architecture of the song system. Therefore, the simple dynamical elements that are necessary to reproduce the observed respiratory gestures can be easily found in the model. In other words, there are different ways, i.e., choice of parameters, in which a given output can be found. The advantage of building a dynamical model is that in the process of reproducing a given output, the predictions the model makes for activities in other parts of the song system can be inspected. In this work, we report for the first time the neural activity in the nucleus HVC during song production in canaries, in order to test the predictions of the dynamical model. We recorded the neural activity in HVC and the air sac pressure in singing canaries. Since our model can reproduce the observed respiratory gestures, we compared the activity it needs in HVC to reproduce them to actual neural recordings. It is important to notice that the solutions are not unique. Yet, they bound the number of possible HVC coding strategies (i.e., *neural activity patterns*) that are consistent with a given respiratory output, under the hypotheses used to build the model.

In our experiments, we were capable of sorting out activity traces of individual neurons. We focused on phasic neurons, since those are putative projection neurons. We concentrated on two syllable types, found in all the birds of our study: brief syllables generated with simple harmonic-like respiratory gestures (P1 syllables), and long whistles (called P0). In previous work, it has been shown that different individuals generate the same syllable types using very stereotyped respiratory gestures (Alliende *et al.*, 2010). We assume in this work that the neural coding for the same syllable types is also stereotyped. Consistent with that hypothesis, we grouped the neural information obtained for different individuals singing similar syllables in a set and studied the statistical nature of the neuronal activity in each set. We found that P1 syllables were generated with a rhythmic neural HVC pattern, and that the P0 syllables presented a sparse neural coding, in which activity was found before each of the two notes in the syllable, as well as at the end. Previous work has been mainly performed in a different species (zebra finches), in which the song is a rapid succession of different and very brief sound elements (i.e., the syllables do not repeat). This led to a description of HVC neural activity that was approximately uniform (Lynch *et al.*, 2016 and Picardo *et al.*, 2016). Given the large number of rapid changes involved in the generation of zebra finch song, it is difficult to disambiguate between a model for which HVC activity is uniform from one in which specific motor instances are preferred. Yet, for juvenile zebra finches that produce simpler sounds, the beginnings of those sounds have been shown to be preferentially coded

(Okubo *et al.*, 2015). Also, for their simplest syllables, the harmonic stacks, neural activity is sparser. Altogether, these results suggest that significant instances such as syllable beginnings might be preferentially coded. This led us to work with canaries, where simpler syllables such as the long whistles studied in this work (P0 syllables) offer an ideal opportunity to disambiguate between these different coding strategies. We found that in fact, the activity of phasic neurons in HVC is highly heterogeneous. This is particularly clear for long whistles in canary song. For these syllables, HVC activity of phasic neurons is preferentially found temporally close to syllabic onsets and offsets. It is interesting to point out that nucleus HVC is not only part of the song system: it receives inputs from the auditory pathway. It is known that parts of this pathway preferentially code for sound onsets (Amin *et al.*, 2004). Therefore, a preferential coding in HVC for syllabic onsets could be related to a need for articulating the production and perception of learned behaviors.

In addition to providing insight on how some cortical areas might code the motor gestures used to generate song, the possibility of reproducing real data by means of a dynamical system of relatively low dimensionality highlights the importance of carrying out a comprehensive theory of coupled, out of equilibrium systems. The interplay between experiments and phenomenological models allows us to test the hypotheses on which the models are built. In this regard, birdsong is a powerful animal model to study the generation of complex and yet reasonably low dimensional motor gestures.

ACKNOWLEDGMENTS

We would like to thank Mariano Belluscio for his help with tetrode and microdrive development. This work describes research partially funded by the National Council of Scientific and Technical Research (CONICET, Argentina—No. 11220130100068), National Agency of Science and Technology (ANPCyT, Argentina—No. PICT2014-1802), University of Buenos Aires (UBA, Argentina—No. 20020170100220BA).

DATA AVAILABILITY

The codes developed for the simulations of the dynamical model are openly available on Github at https://github.com/dynamical-systems-laboratory-ar/model_birdsong_neural [Herbert *et al.* (2020)].

REFERENCES

- Alliende, J. A., Mendez, J. M., Goller, F., and Mindlin, G. B., "Hormonal acceleration of song development illuminates motor control mechanism in canaries," *Dev. Neurobiol.* **70**(14), 943–960 (2010).
- Alonso, L. M., Alliende, J. A., Goller, F., and Mindlin, G. B., "Low-dimensional dynamical model for the diversity of pressure patterns used in canary song," *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **79**(4 Pt 1), 041929 (2009).
- Alonso, L. M., Alliende, J. A., and Mindlin, G. B., "Dynamical origin of complex motor patterns," *Eur. Phys. J. D* **60**(2), 361–367 (2010).
- Alonso, R. G., Amador, A., and Mindlin, G. B., "An integrated model for motor control of song in *Serinus canaria*," *J. Phys.* **110**(3), 127–139 (2016).
- Alonso, R. G., Trevisan, M. A., Amador, A., Goller, F., and Mindlin, G. B., "A circular model for song motor control in *Serinus canaria*," *Front. Comput. Neurosci.* **9**, 41 (2015).

- Amador, A. and Margoliash, D., "A mechanism for frequency modulation in songbirds shared with humans," *J. Neurosci.* **33**(27), 11136–11144 (2013a).
- Amador, A., Perl, Y. S., Mindlin, G. B., and Margoliash, D., "Elemental gesture dynamics are encoded by song premotor cortical neurons," *Nature* **495**(7439), 59–64 (2013b).
- Amin, N., Grace, J. A., and Theunissen, F. E., "Neural response to bird's own song and tutor song in the zebra finch field L and caudal mesopallium," *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* **190**(6), 469–489 (2004).
- Ashmore, R. C., Wild, J. M., and Schmidt, M. F., "Brainstem and forebrain contributions to the generation of learned motor behaviors for song," *J. Neurosci.* **25**(37), 8543–8554 (2005).
- Borisyuk, R. M. and Kirillov, A. B., "Bifurcation analysis of a neural network model," *Biol. Cyber.* **66**(4), 319–325 (1992).
- Buzsáki, G., Anastassiou, C. A., and Koch, C., "The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes," *Nat. Rev. Neurosci.* **13**, 407–420 (2012).
- Dima, G. C., Copelli, M., and Mindlin, G. B., "Anticipated synchronization and zero-Lag phases in population neural models," *Int. J. Bifurc. Chaos* **28**(08), 1830025 (2018a).
- Dima, G. C., Goldin, M., and Mindlin, G. B., "Modeling temperature manipulations in a circular model of birdsong production," *Papers Phys.* **10**, 100002 (2018b).
- Gardner, T., Cecchi, G., Magnasco, M., Laje, R., and Mindlin, G. B., "Simple motor gestures for birdsongs," *Phys. Rev. Lett.* **87**(20), 208101 (2001).
- Henze, D. A., Borhegyi, Z., Csicsvari, J., Mamiya, A., Harris, K. D., and Buzsáki, G., "Intracellular features predicted by extracellular recordings in the hippocampus in vivo," *J. Neurophysiol.* **84**(1), 390–400 (2000).
- Herbert, C. T., Boari, S., Mindlin, G. B., and Amador A (2020), "A neural population model for birdsong production," GitHub, see <https://doi.org/10.5281/zenodo.3774481>.
- Hoppensteadt, F. C. and Izhikevich, E. M., *Weakly Connected Neural Networks* (Springer Science & Business Media, 2012).
- Lassa Ortiz, J. N., Herbert, C. T., Mindlin, G. B., and Amador, A., "Significant instances in motor gestures of different songbird species," *Front. Phys.* **7**, 142 (2019).
- Lynch, G. F., Okubo, T. S., Hanuschkin, A., Hahnloser, R. H., and Fee, M. S., "Rhythmic continuous-time coding in the songbird analog of vocal motor cortex," *Neuron* **90**(4), 877–892 (2016).
- Mindlin, G. B. and Laje, R., *The Physics of Birdsong* (Springer-Verlag, Berlin, 2006).
- Mooney, R., "Different subthreshold mechanisms underlie song selectivity in identified HVc neurons of the zebra finch," *J. Neurosci.* **20**, 5420–5436 (2000).
- Mooney, R., "Neurobiology of song learning," *Curr. Opin. Neurobiol.* **19**, 654–660 (2009).
- Nottebohm, F., Alvarez-Buylla, A., Cynx, J., Kirn, J., Ling, C. Y., Nottebohm, M., and Williams, H., "Song learning in birds: The relation between perception and production," *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **329**(1253), 115–124 (1990).
- Okubo, T. S., Mackevicius, E. L., Payne, H. L., Lynch, G. F., and Fee, M. S., "Growth and splitting of neural sequences in songbird vocal development," *Nature* **528**(7582), 352–357 (2015).
- Ott, E. and Antonsen, T. M., "Low dimensional behavior of large systems of globally coupled oscillators," *Chaos* **18**(3), 037113 (2008).
- Picardo, M. A., Merel, J., Katlowitz, K. A., Vallentin, D., Okobi, D. E., Benezra, S. E., Clary, R. C., Pnevmatikakis, E. A., Paninski, L., and Long, M. A., "Population-level representation of a temporal sequence underlying song production in the zebra finch," *Neuron* **90**(4), 866–876 (2016).
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., "Numerical recipes in C," *Probab. in the Eng. Inf. Sci.* **2**, b1–b3 (1988).
- Rauske, P. L., Shea, S. D., and Margoliash, D., "State and neuronal class-dependent reconfiguration in the avian song system," *J. Neurophysiol.* **89**, 1688–1701 (2003).
- Roulet, J. and Mindlin, G. B., "Average activity of excitatory and inhibitory neural populations," *Chaos* **26**(9), 093104 (2016).
- Suthers, R. A., Goller, F., and Hartley, R. S., "Motor stereotypy and diversity in songs of mimic thrushes," *J. Neurobiol.* **30**(2), 231–245 (1996).
- Suthers, R. A. and Margoliash, D., "Motor control of birdsong," *Curr. Opin. Neurobiol.* **12**(6), 684–690 (2002).
- Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., and Ghazanfar, A. A., "The developmental dynamics of marmoset monkey vocal production," *Science* **349**(6249), 734–738 (2015).
- Trevisan, M. A., Mindlin, G. B., and Goller, F., "Nonlinear model predicts diverse respiratory patterns of birdsong," *Phys. Rev. Lett.* **96**(5), 058103 (2006).
- Tytell, E. D., Holmes, P., and Cohen, A. H., "Spikes alone do not behavior make: Why neuroscience needs biomechanics," *Curr. Opin. Neurobiol.* **21**(5), 816–822 (2011).
- Ward, B. C., Nordeen, E. J., and Nordeen, K. W., "Individual variation in neuron number predicts differences in the propensity for avian vocal imitation," *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1277–1282 (1998).
- Wilson, H. R. and Cowan, J. D., "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Kybernetik* **13**, 55–80 (1973).