

Spike train statistics for consonant and dissonant musical accords in a simple auditory sensory model

Yuriy V. Ushakov* and Alexander A. Dubkov†

Radiophysics Department, N.I. Lobachevsky State University, 23 Gagarin Avenue, 603950 Nizhny Novgorod, Russia

Bernardo Spagnolo‡

Dipartimento di Fisica e Tecnologie Relative, Group of Interdisciplinary Physics, Università di Palermo and CNISM-INFM, Viale delle Scienze, I-90128 Palermo, Italy

(Received 7 August 2009; revised manuscript received 17 February 2010; published 13 April 2010)

The phenomena of *dissonance* and *consonance* in a simple auditory sensory model composed of three neurons are considered. Two of them, here so-called *sensory neurons*, are driven by noise and subthreshold periodic signals with different ratio of frequencies, and its outputs plus noise are applied synaptically to a third neuron, so-called *interneuron*. We present a theoretical analysis with a probabilistic approach to investigate the interspike intervals statistics of the spike train generated by the interneuron. We find that tones with frequency ratios that are considered *consonant* by musicians produce at the third neuron inter-firing intervals statistics densities that are very distinctive from densities obtained using tones with ratios that are known to be *dissonant*. In other words, at the output of the interneuron, inharmonious signals give rise to blurry spike trains, while the harmonious signals produce more regular, less noisy, spike trains. Theoretical results are compared with numerical simulations.

DOI: [10.1103/PhysRevE.81.041911](https://doi.org/10.1103/PhysRevE.81.041911)

PACS number(s): 87.19.lc, 87.10.Ca, 87.19.lt

I. INTRODUCTION

Since 1980-th it has been well known that noise in physical systems does not always play a negative role. Noise usually means something that causes unwanted disturbance and blurs the signal processing. However, noise in many cases can be a message by itself or a highly desirable part of the message, important for signal processing. In different branches of science, indeed, the interplay of nonlinearity, environmental noise, and periodic forces gives rise to many noise-induced effects such as stochastic resonance [1,2], coherence resonance [3,4], noise enhanced stability [5,6], etc., which show the constructive role of the noise. A typical field of investigation is the wide class of neural systems, which are naturally noisy. In neurons, in fact, noise arises from many different sources, such as the quasirandom release of neurotransmitter by the synapses, the random switching of ion channels (channel noise), and most importantly random synaptic input from other neurons (each neuron receives on the average 10^4 inputs from its neighbors [7]). Noise affects all aspects of nervous-system function. Both the central and peripheral nervous system are subjected to the noise [8].

A natural question arises: how does a signal survive in such a noisy environment? Looking for an approach to this problem we focused our attention on sensory systems [9,10]. Typically, in sensory systems there is a set of neurons, referred to as *sensory neurons* or *sensors*, receiving signals directly from the environment. Sensory neurons are responsible for converting external stimuli from the environment into internal stimuli. A classical approach to modeling the

auditory system is to consider the periphery as a Fourier transform followed by a number of bandpass filters and to view the function of entire lower auditory system as being a spectrum estimator [11,12]. Therefore, the sensory neurons attached directly to the basilar membrane receive different sinusoidal components (depending on the coordinates of connection along the membrane) of the input sound. These periodical signals, together with noise, act as input signals to sensory neurons, resulting in trains of very short pulses, “spikes,” which are transmitted to other neurons (*interneurons*) along the neural fibers.

The perception and processing of environmental complex signals resulting from the combination of two or more input periodical signals are still an open problem for physicists and physiologists. In particular, the precise neural and physiological bases for our perception of musical *consonance* and *dissonance* are still largely unknown [10,13,14]. Although there is no single musical definition, *consonance* is usually referred to as the pleasant stable sound sensation produced by certain combinations of two tones played simultaneously. Conversely, *dissonance* is the unpleasant unstable sound heard with other sound combinations [15]. The dominant and the oldest theory of *consonance* and *dissonance* is that of Pythagoras (around 500 BC). He observed that the simpler the frequency ratio between two tones, the more *consonant* they will be perceived. For example, the *consonant* octave is characterized by a 1/2 frequency ratio between two tones, while the *dissonant* semitone is characterized by a 15/16 ratio. In 1877, Helmholtz analyzed the phenomenon of *consonance* and *dissonance* in the more general context of complex tones and proposed the “beat theory.” When two complex tones are played together (as an interval), the harmonics of each tone are present in the stimulus arriving at the ear of the listener. For some combinations (simple ratio n/m) the harmonic frequencies match, for others (complicated ratio

*nomah@list.ru

†dubkov@rf.unn.ru

‡spagnolo@unipa.it

n/m) they do not. As the frequency ratio n/m becomes more “complicated,” the two tones share fewer common harmonics and there is an increase in harmonics pair slightly mismatched in frequency which give unpleasant beating sensation. In other words, the *dissonance* is proportional to the number of frequency components present in the two complex tones that produce beats [11].

Here, after shortly reviewing two recent theoretical approaches to the perception of complex signals we present our approach to the statistics of *consonant* and *dissonant* musical accords in the presented sensory model.

A. Pitch perception and ghost stochastic resonance

Pitch is a subjective sensation in which a listener assigns perceived tones to relative positions on a musical scale based primarily on the frequency of vibration. In other words, pitch represents the perceived frequency of a sound. How the brain estimates the pitch of complex sounds, formed by a combination of pure tones, remains a controversial issue [16–20]. For harmonic complex sound signals, whose constituent frequencies are multiple integers of a fundamental frequency, the perceived pitch is the fundamental, even if that frequency is not spectrally present in the input signal. This is known as “missing fundamental illusion.” Recently, a mechanism for the perception of pitch has been proposed on the basis of the so-called *ghost stochastic resonance* (GSR) [21,22]. The proposed mechanism shows that a neuron responds optimally to the missing fundamental of a harmonic complex signal for an appropriate level of noise. The main ingredients are: (i) a linear interference between the individual tones, producing peaks of constructive interference at the fundamental frequency (missing fundamental or ghost frequency), whose amplitude is not suitable to trigger the neuron, and (ii) a nonlinear threshold that detects those peaks with the help of a suitable amount of noise. GSR has been observed experimentally in semiconductor lasers [23] and in electronic circuits [24]. The GSR mechanism was extended later to describe a higher level of perception processing, which is the binaural pitch perception [25,26]. Two different neurons, each one representing detection at a different auditory channel, receive one single component of the complex signal each, and their output spike trains drive a third neuron that processes the information. This processing neuron responds preferentially at the ghost frequency and the response is optimized by synaptic noise.

B. Nonlinear synchronization theory of musical consonance

A synchronization theory of *consonance* that goes beyond the linear beating theory of Helmholtz was recently proposed in Ref. [13]. By using a simple scheme of two mutually coupled neural oscillators, the authors showed that the mode-locked states ordering gives precisely the standard ordering of *consonance*. They analyzed the dynamics of two coupled leaky integrate-and-fire neuron models, with mutual excitatory coupling, by finding that the mode locking ratios n/m (with n and m integers) are ordered according to the “Farey sequence,” which orders all rational fractions n/m in the interval $[0,1]$ according to their increasing denominators m

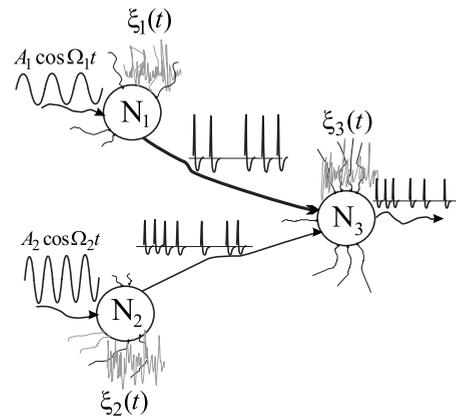


FIG. 1. The investigated model. Sensory neurons N_1 and N_2 are driven by subthreshold sinusoidal signals with different frequencies. Spike trains of sensors are received by the interneuron N_3 . All three neurons are perturbed by independent noise sources $\xi_1(t)$, $\xi_2(t)$ and $\xi_3(t)$.

[27]. By plotting the ratio of actual firing frequencies as a function of the ratio of natural intrinsic frequencies of the two coupled oscillators, they reproduce the so-called “Devil’s Staircase,” with flat steps corresponding to different mode-locked states. This is a universal feature of driven coupled oscillators [28]. The width of each step, which is of the mode-locked interval, is an indicator of the structural stability of the synchronization. It is therefore possible to order the mode-locked states by their stability index (the width of the step), by finding a correspondence with the theoretical ordering of musical intervals according to their *consonance* evaluation. Heffernan and Longtin in Ref. [14] analyzed in detail the same model of Ref. [13] by considering different values of coupling between the oscillators. They found that the ordering of mode-locked states is not universal but depends on the coupling strength. The authors also explored generic aspects of a possible synchronization theory by driving the model neurons with sinusoidal forcing or random forcing.

C. Probabilistic approach

In this paper we investigate the phenomena of *consonance* and *dissonance* in a simple perception model system by a probabilistic approach. The system, shown in Fig. 1, is composed of three neural-like noisy elements. Two of them represent *sensory neurons* and are driven by noise and subthreshold periodic signals with different ratio of frequencies. The outputs of these neurons, in the presence of noisy environment, are synaptically connected (only spike train transmission) with the third element which represents an *interneuron*, which is an internal neuron which connects sensory neuron to other neurons within the same region of the brain. We analyze the probability distribution of interspike intervals (ISIDs) of the output signal of the *interneuron* by assuming to know the ISIDs of the output signals of the two sensory neurons $\rho_1(t)$ and $\rho_2(t)$. In principle, there are many events, each one characterized by different probability, for which the third neuron can fire due to the spikes arriving from the two

sensors. We reduce the number of events for which the interneuron can fire to four main scenarios because all other events have a very negligible probability to happen in comparison with the previous four. In this way we are able to calculate the first passage time distribution at the output of the interneuron $\rho_3(t)$, using conditional probabilities and first passage time distributions at the output of sensory neurons. Moreover, for periodical input signal at the sensors with frequency ratio m/n we obtain $(m+n-1)$ different patterns of input spike trains for the interneuron, with different ISIDs at its output. The final interspike interval density of the interneuron $\rho_{out}(T)$ is obtained by averaging the first passage time density $\rho_3(t)$ over all different states $(m+n-1)$ of the interneuron.

We show how a complex input composed of two harmonic signals is transformed by the proposed simple sensory system into different types of spike trains, depending on the ratio of input frequencies. Looking for the differences in the statistical sense, we find out that the output ISIDs for some combinations of frequencies, corresponding to *consonant* accords, have more regular pattern, while inharmonious signals, corresponding to *dissonant* accords, show less regular spike trains and blurry ISIDs. This difference indicates that *consonant* accords are higher stable, with respect to the noise environment, in comparison with the *dissonant* accords in the processing of information throughout the auditory system.

We note that our approach differs from the previous one (*nonlinear synchronization theory*) because we use three neurons in our sensory system model and there is no mutual connection that gives raise to synchronization between neurons. Moreover, we focus on ISI statistics to analyze *consonant* and *dissonant* accords. Concerning the *pitch perception* and *ghost stochastic resonance*, our investigation has close connections with this phenomenon. In fact, the output ISIDs of the interneuron comprise peaks which are not present in the ISIDs of the sensory neurons and peaks corresponding to ghost frequencies, which are not present in the input complex tone of the sensory system. But our analysis is mainly focused on finding the analytical expression of the output ISID of the interneuron, and we do not look for the optimum noise level. In addition, our fixed noise intensity value is out of optimum range of the ghost stochastic resonance.

In what follows, after describing the model in Sec. II, we present the theoretical analysis based on a probabilistic approach in Sec. III and a flowchart for calculation algorithm in Sec. IV. The results of numerical simulations on the proposed model and comparison with the theoretical ones are reported in Sec. V. A discussion on the qualitative agreement of our theoretical results with those obtained under the hypotheses about *consonance*, and *dissonance* in music proposed by Helmholtz (1877) and Boomsliiter and Creel (1961) is reported in Sec. VI. In the final section we draw our conclusions.

II. MODEL

As a neuron model for our sensory system (see Fig. 1) we consider the leaky integrate-and-fire (LIF) model. The LIF

model is one of the most widely used spiking neuron models [29]. While the LIF model does not provide complete descriptions of real neurons, it has successfully been applied to explain the high temporal precision achieved in the auditory [30] and visual systems [31]. The sensory neurons (N_1 and N_2) are driven by the external sinusoidal signals, and the interneuron (N_3) receives the weighted spikes of the sensors through synaptic connections. Therefore, the set of stochastic differential equations describing our system is

$$\begin{cases} \dot{v}_1 = -\mu_1 v_1 + A_1 \cos(\Omega_1 t) + \sqrt{D_1} \xi_1(t), \\ \dot{v}_2 = -\mu_2 v_2 + A_2 \cos(\Omega_2 t) + \sqrt{D_2} \xi_2(t), \\ \dot{v}_3 = -\mu_3 v_3 + k_1 s_1(t) + k_2 s_2(t) + \sqrt{D_3} \xi_3(t), \end{cases} \quad (1)$$

where $v_i(t)$ and μ_i stand for the membrane potential and the relaxation parameter, respectively, and subscript i labels the different neurons, with $i=1,2$ representing the two input sensory neurons (N_1 and N_2) and $i=3$ (N_3) denoting the processing interneuron. A_i and Ω_i (with $i=1,2$) are the amplitude and the frequency of the corresponding harmonic input of the sensors. We consider that the three neurons have different synaptic connections, they are not subject to the same background noise, and the three noise sources $\xi_i(t)$ of Fig. 1 are independent of each other [7]. Therefore, in Eq. (1), the three white Gaussian noise terms $\xi_i(t)$ ($i=1,2,3$) are uncorrelated and with the usual statistical properties $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t) \xi_j(t') \rangle = \delta(t-t') \delta_{ij}$. D_i is the noise intensity in each neuron. In Eq. (1) $s_i(t) = \sum_{j=0}^{N_i(t)} \delta(t-t_{ij})$, $i=1,2$ are the spike trains generated by the sensors and received by the interneuron as input, k_i ($i=1,2$) are the coupling coefficients. Spikes are modeled by Dirac δ functions. The LIF model does not comprise any mechanism of spike generation. When the membrane potential v_i reaches the threshold value v_{th} , the neuron is said to fire a spike, and v_i is reset to its initial value v_i^0 . In particular, the input spikes at the interneuron, coming from the sensory neurons, can produce spikes or jumps in the membrane potential of the interneuron, depending on whether or not they are suitable to fire the interneuron.

The firing times t_{ij} are the times in which the membrane potentials of the sensory neurons cross the threshold v_{th} , which we consider equal for all the neurons. $N_i(t)$ is the number of spikes generated by the i^{th} sensor since the initial time. For the output interneuron the *refractory period* (T_{ref}) is introduced explicitly. This neuron does not respond to any external signal after reset at time, t_{res} , until the membrane potential $v_3^0 e^{-\mu_3(t-t_{res})}$ reaches the level $v_3 = -0.1$. Hence, the refractory period can be written as

$$T_{ref} = \frac{1}{\mu_3} \ln(-10v_3^0). \quad (2)$$

All simulation and theoretical results presented in the paper are obtained using the following set of values of system parameters, namely $\mu_1 = \mu_2 = 1$, $\mu_3 = 0.3665$, $D_1 = D_2 = D_3 = 1.6 \times 10^{-3}$, $k_1 = k_2 = 0.98$, $v_1^0 = v_2^0 = 0$, $v_3^0 = -1$, and $v_{th} = 1$, unless stated otherwise. For the chosen parameter values we have $T_{ref} = 6.28$.

III. THEORETICAL APPROACH

The first two equations of system (1) describe the Ornstein-Uhlenbeck processes with harmonic driving forces. For the Ornstein-Uhlenbeck neuronal model, the ISID was obtained analytically with different approaches in Refs. [32,33]. This distribution, which coincides with the first passage time probability distribution related to the firing event of sensory neurons, is our starting point to obtain the ISID at the output of the interneuron.

It is important to note here that the ISIDs at the output of two sensors are non-Poissonian [see next Fig. 2(b)]. Indeed, non-Poissonian distribution of interspike intervals was very

recently observed in Primate parietal cortical neurons [34]. These spike trains are the input of the third neuron, and as a consequence the dynamics of the membrane potential of the interneuron is non-Markovian [35]. Hence, because we cannot use the Markovian theoretical apparatus, we are compelled to investigate the ISI statistics of the interneuron using another analytical approach.

A. Assumptions

The analytical solutions for the membrane potential of the sensors $v_i(t)$ ($i=1,2$) and for the interneuron $v_3(t)$ are [36]

$$\begin{cases} v_i(t) = \left[v_i^0 - \frac{A_i}{\sqrt{\Omega_i^2 + \mu_i^2}} \cos(\Omega_i t_i^0 - \phi_i) \right] e^{-\mu_i(t-t_i^0)} + \frac{A_i}{\sqrt{\Omega_i^2 + \mu_i^2}} \cos(\Omega_i t - \phi_i) + \sqrt{D_i} \zeta_i(t), & (i=1,2) \\ v_3(t) = v_3^0 e^{-\mu_3(t-t_3^0)} + \sum_{i=1}^2 k_i S_i(t) + \sqrt{D_3} \zeta_3(t) \end{cases} \quad (3)$$

These equations are valid from the initial time t_i^0 ($i=1,2,3$) until the next spike at each neuron occurs. Here $S_i(t) = \sum_{j=0}^{N_i(t)} e^{-\mu_3(t-t_{ij})}$ ($i=1,2$) is a sum of decaying impulses evoked by spikes of the i^{th} sensory neuron; $\zeta_i(t) = \int_{t_i^0}^t e^{-\mu_i(t-t')} \xi_i(t') dt'$ ($i=1,2,3$) is the Ornstein-Uhlenbeck process with zero mean, variance

$$\sigma_i^2(t) = \frac{1}{2\mu_i} (1 - e^{-2\mu_i(t-t_i^0)}) \quad (i=1,2,3) \quad (4)$$

and probability distribution

$$w_{\zeta_i}(s) = \frac{1}{\sqrt{2\pi\sigma_i(t)}} \exp\left(-\frac{s^2}{2\sigma_i^2(t)}\right), \quad (i=1,2,3). \quad (5)$$

In Eq. (3), t_i^0 are the reset times (spike generation) of the sensors ($i=1,2$) and the interneuron ($i=3$); and $\phi_i = \arctan(\Omega_i/\mu_i)$, with $i=1,2$.

The temporal realizations of membrane potentials of neurons allow us to understand the conditions of spike generation by the interneuron and to establish connections between these events and the input signals.

In order to perform this analysis we use three main assumptions:

(1) the input harmonic signals are *subthreshold* for the sensors, which is the values of the amplitude A_i and the frequency Ω_i are such that the signal $A_i \cos(\Omega_i t)$ is not able to bring the membrane potential of the i^{th} sensor above the threshold in the absence of noise ($D_i=0$). This means absence of spikes at the output of the sensors. From the first equation of system (3) we obtain

$$\frac{A_i}{\sqrt{\Omega_i^2 + \mu_i^2}} < v_{th}. \quad (6)$$

(2) Only one spike can be generated at each period of the harmonic driving force, and, at the same time, the spiking on each period is the most probable situation [see Fig. 2(a)]. This means that the relaxation times of sensors are smaller than the periods of the sinusoidal signals

$$\frac{1}{\mu_i} \leq \frac{2\pi}{\Omega_i}. \quad (7)$$

(3) Each of coupling coefficients k_i is less than the threshold value of the membrane potential v_{th} . It means that any separate incoming spike (that is spike coming after the relaxation time or the refractory time [see Fig. 2(c)]) evokes a subthreshold impulse of the membrane potential of the interneuron $v(t)$; i.e., spike generation is impossible without noise. At the same time, the sum of the two coupling coefficients is greater than v_{th} ,

$$\begin{cases} k_{1,2} < v_{th}, \\ k_1 + k_2 > v_{th}. \end{cases} \quad (8)$$

B. First passage time probability distribution of the interneuron

1. Probability of spike generation at the interneuron

Initially all three neurons of system (1) are reset, i.e., $v_1(0)=v_1^0$, $v_2(0)=v_2^0$, and $v_3(0)=v_3^0$. Since the starting time is $t=0$, we measure the first interspike period of the output neuron as the first time to cross the threshold v_{th} , that is the first passage time problem. The first passage time probability distributions (FPTPDs) are considered to be known for the input neurons: $\rho_1(t)$ and $\rho_2(t)$, respectively. We can use their theoretical expressions [32,33] or we can calculate them numerically from Eq. (1). All the theoretical results of this pa-

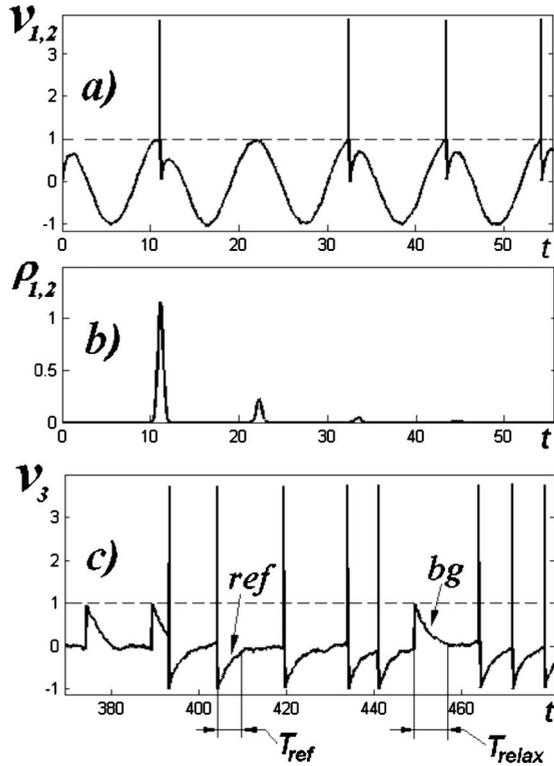


FIG. 2. (a) Typical behavior of the membrane potential $v_i(t)$ of sensory neurons versus time for a noise realization. (b) ISI distribution of the sensory neurons ($\Omega=0.6$). The highest probability of a spike after $t=0$ is near one period of external force ($t=10.47$). The probability of firing after two, three, etc. periods decreases exponentially. (c) Typical behavior of the membrane potential $v_3(t)$ of the interneuron versus time for the same noise realization. Here are well visible the refractory state (*ref*), characterized by the refractory time T_{ref} , and the noisy background (*bg*) during the relaxation time T_{relax} . The values of the system parameters are $\mu_1=\mu_2=1$, $\mu_3=0.3665$ (relaxation parameters), $D_1=D_2=D_3=1.6 \times 10^{-3}$ (noise intensities), $k_1=k_2=0.98$ (coupling coefficients), $v_1^0=v_2^0=0$, $v_3^0=-1$, (reset values of the membrane potentials), and $v_{th}=1$ (threshold value of the membrane potential); $A_1=1.084$, $A_2=1.115$, $\Omega_1=0.4$, and $\Omega_2=0.42$ (amplitudes and frequencies of the periodical driving forces); and $T_{max}=10^5$ (maximum output time length of the time series generated from the system). From Eqs. (2) and (14), we have $T_{ref}=6.28$ and $T_{relax}=8.73$.

per are obtained using the FPTPDs $\rho_1(t)$ and $\rho_2(t)$, numerically evaluated from the first two equations of system (1). Therefore, we know the characteristics of the output spike sequences from the sensory neurons N_1 and N_2 , which go, through the synaptic connection, to the interneuron N_3 . Because of the harmonic input signals and the noise, the sensory neurons fire with the highest probability at times in which the harmonic driving forces have their maxima. The sensors have a narrow probability distribution of spike intervals near each of these maxima, and the probability of skipping one, two, etc. periods decays exponentially [see Figs. 2(a) and 2(b)].

Before to analyze the FPTPD of the interneuron, we describe its typical output behavior [see Fig. 2(c)]. The instant of time in which the interneuron receives a spike coming

from the sensors, plus the noise signal, the interneuron produces a jump or a spike, depending on whether the membrane potential is below or above the threshold v_{th} . After jumping the neuron relaxes exponentially toward the zero value of the membrane potential ($v_3=0$) with time scale equal to $\tau_3=1/\mu_3$. Within this time the interneuron can fire because the incoming spike (from the other sensor) adds to the noisy “background” level of the membrane potential [indicated by the symbol *bg* in Fig. 2(c)], and the membrane potential crosses the threshold. In other words, one sensor causes the jump and the other one the spike. After releasing a spike, the interneuron is in a refractory state, and the membrane potential relaxes from its reset value $v_3=-1$ toward the level $v_3=-0.1$ in the time T_{ref} given by Eq. (2). Now, we define “separate spike,” a spike coming at the interneuron N_3 at the time in which its membrane potential fluctuates around the zero level, after a relaxation or a refractory period of time [Fig. 2(c)]. Therefore, the dynamics of the interneuron after spiking or jumping is characterized by refractory time and relaxation time, respectively.

The next spiking of the interneuron depends on these time scales, the connectivity coefficients, the frequencies of the input periodical signals driving the sensors, and the noise intensity acting on neuron N_3 , giving rise to different firing events for the interneuron. Due to the two typical membrane potential time behavior of the interneuron [Fig. 2(c)] and the chosen range of variability of system parameters, we have only four main different scenarios to fire the interneuron.

Therefore, we can evaluate the probability $\Delta P_3(t) = \rho_3(t)\Delta t$ that the interneuron N_3 fires in the short time interval $(t, t+\Delta t)$ by considering the occurrence of the following events:

- (1) receiving a *separate firing spike* from the sensory neuron N_1 ;
- (2) receiving a *separate firing spike* from the sensory neuron N_2 ;
- (3) receiving a firing spike from the neuron N_1 on the *background* of the membrane potential relaxing, after the jump due to the spike from the N_2 neuron, toward the zero value; in other words, sensor N_2 causes the jump and then sensor N_1 the spike; and
- (4) receiving a firing spike from the neuron N_2 on the *background* of the membrane potential relaxing, after the jump due to the spike from the N_1 neuron, toward the zero value; in other words sensor N_1 causes the jump and then sensor N_2 the spike.

We neglect the contribution of multiple jump events to fire the interneuron and the noise-induced spike events occurring during the relaxation of the membrane potential after a jump because they have very negligible probability of happening in comparison with the previous four, with the chosen range of system parameters. In fact, the amplitudes of jumps, produced in the membrane potential of the interneuron and due to the incoming spikes from the sensors, are equal to the connecting coefficients k_1 and k_2 . This means that after a jump of the neuron N_3 , the next incoming spike from one of the two sensors will produce a spike at the output of the interneuron with very high probability because of relation $(k_1+k_2) > v_{th}$ of Eq. (8).

The four described scenarios exclude each other, so they are mutually exclusive events. As a result, according to the

formula of total probability we have to add up all probabilities of the above mentioned events. These depend on the coupling coefficients $k_{1,2}$, the noise intensity D_3 , the membrane threshold v_{th} and the ISIDs $\rho_1(t)$, $\rho_2(t)$.

The first probability is the product of the probability to have a spike from the sensory neuron N_1 within the time interval $(t, t+\Delta t)$ and the conditional probability to fire the interneuron N_3 , given the occurrence of the spike generation at the neuron N_1 ; that is,

$$\rho_1(t_1)\Delta t \cdot \text{Prob}\{k_1 + \sqrt{D_3}\zeta_3(t_1) \geq v_{th}\}, \quad (9)$$

where t_1 is the time when spike from neuron N_1 appears and fires the interneuron ($t < t_1 < t+\Delta t$). To obtain the probability of the second scenario we should replace subscript 1 with 2 in Eq. (9).

Particular attention should be paid to the evaluation of the third contribution to the probability of the spike generation of the neuron N_3 . Suppose that the spike of the neuron N_2 , plus the noise signal, arrive at the neuron N_3 within the interval $(t', t'+\Delta t')$ and are not able to fire the neuron. In this case the membrane potential $v_3(t)$ produces a ‘‘jump’’ with height k_2 , but after this decays in average exponentially toward its zero value. During this relaxation time the interneuron can fire because the incoming spike from the sensor N_1 adds to the noisy ‘‘background’’ level of the membrane potential, and the membrane potential crosses the threshold at time t_1 [see the first spike in Fig. 2(c)], with $t' < t$ and $t < t_1 < t+\Delta t$. In other words, the sensor N_2 causes the jump within the time interval $(t', t'+\Delta t')$ and the sensor N_1 fires the neuron N_3 at time t_1 . Therefore, the joint probability of this complicated event is expressed as the product of the probabilities of these two independent output events at the interneuron, which is jumping and firing. The probability of jumping of the interneuron, due to the spike of the neuron N_2 , is the product of probability to have a spike from sensory neuron N_2 at time t'_1 within the interval $(t', t'+\Delta t')$ with the conditional probability of the jump at the interneuron, namely,

$$\rho_2(t'_1)\Delta t' \cdot \text{Prob}\{k_2 + \sqrt{D_3}\zeta_3(t'_1) < v_{th}\} = \rho_2(t'_1)\Delta t' P_2(t'_1). \quad (10)$$

The probability to fire the interneuron is the product of probability to have a spike from the sensory neuron N_1 within the time interval $(t, t+\Delta t)$ and the conditional probability of the spike generation of neuron N_3 , given the occurrence of the jumping within the time interval $(t'+\Delta t')$ (caused by a spike from neuron N_2) and the firing at time t_1 due to the arrival of a spike from neuron N_1 in the interval $(t, t+\Delta t)$. We have

$$\begin{aligned} & \rho_1(t_1)\Delta t \cdot \text{Prob}\{k_1 + k_2 e^{-\mu_3(t_1-t'_1)} + \sqrt{D_3}\zeta_3(t_1) \geq v_{th}\} \\ & = \rho_1(t_1)\Delta t P_{12}(t_1 - t'_1). \end{aligned} \quad (11)$$

Now, since the jumping of neuron N_3 can occur at any time t'_i , within the relaxation period T_{relax_i} ($i=1,2$), we have to integrate all these contributions within a time interval less or equal to this time, which is

$$\begin{aligned} & \lim_{\Delta t'_i \rightarrow 0} \sum_i \rho_2(t'_i)\Delta t'_i P_2(t'_i)\rho_1(t_1)\Delta t P_{12}(t_1 - t'_i) \\ & = \rho_1(t_1)\Delta t \int_{t_1-T_{relax_2}}^{t_1} \rho_2(t')P_2(t')P_{12}(t_1 - t')dt'. \end{aligned} \quad (12)$$

When $\Delta t'_i \rightarrow 0$, then $t'_i \rightarrow t'_i$ and because of integration the discrete time variable t'_i is converted into the continuous one t' . Before integration, we evaluate the relaxation time T_{relax_i} as the time in which the value of the membrane potential $v_3(t)$ because of the relaxation toward the zero value, attains the value of the mean-square deviation of the noise

$$k_i e^{-\mu_3 T_{relax_i}} = \sqrt{D_3}\sigma_3 = \sqrt{\frac{D_3}{2\mu_3}}, \quad (i=1,2) \quad (13)$$

then

$$T_{relax_i} = \frac{1}{\mu_3} \ln \left(k_i \sqrt{\frac{2\mu_3}{D_3}} \right). \quad (i=1,2) \quad (14)$$

In Eq. (14) we report the two relaxation times related to jumping events due to neuron N_1 and neuron N_2 , respectively. Because of our choice of parameter values $k_1=k_2=0.98$, $\mu_3=0.3665$, and $D=1.6 \times 10^{-3}$, we have $T_{relax_1}=T_{relax_2}=T_{relax}$ and $T_{relax}=8.73$.

Therefore, the total joint probability of the third scenario is

$$\begin{aligned} & \rho_1(t_1)\Delta t \int_{t_1-T_{relax_2}}^{t_1} \text{Prob}\{k_2 + \sqrt{D_3}\zeta_3(t') < v_{th}\}\rho_2(t')\text{Prob}\{k_1 \\ & + k_2 e^{-\mu(t_1-t')} + \sqrt{D_3}\zeta_3(t_1) \geq v_{th}\}dt'. \end{aligned} \quad (15)$$

The fourth contribution, due to the fourth scenario, is obtained from the third one by changing the indices $2 \leftrightarrow 1$. Therefore, dividing by Δt and making the limit $\Delta t \rightarrow 0$ ($t_1 \rightarrow t$), we get the ISID at the output of the interneuron

$$\begin{aligned} \rho_3(t) & = \rho_1(t)\text{Prob}\{k_1 + \sqrt{D_3}\zeta_3(t) \geq v_{th}\} + \rho_2(t)\text{Prob}\{k_2 \\ & + \sqrt{D_3}\zeta_3(t) \geq v_{th}\} + \rho_1(t) \int_{t-T_{relax_2}}^t \text{Prob}\{k_2 + \sqrt{D_3}\zeta_3(t') \\ & < v_{th}\} \cdot \rho_2(t')\text{Prob}\{k_1 + k_2 e^{-\mu(t-t')} + \sqrt{D_3}\zeta_3(t) \geq v_{th}\}dt' \\ & + \rho_2(t) \int_{t-T_{relax_1}}^t \text{Prob}\{k_1 + \sqrt{D_3}\zeta_3(t') \\ & < v_{th}\} \cdot \rho_1(t')\text{Prob}\{k_1 e^{-\mu(t-t')} + k_2 + \sqrt{D_3}\zeta_3(t) \geq v_{th}\}dt'. \end{aligned} \quad (16)$$

Of course, all the expressions in Eq. (16) are valid only for $t > T_{ref}$; that is, $\rho_3(t)=0$ for $t < T_{ref}$. We also note that, because of the choice of the values of the system parameters, the scenario of jumping and firing of neuron N_3 , due to the consecutive arrival of two spikes from the same neuron sensor within the relaxation period, does not occur. Actually we will analyze the perception auditory process for periodic signal such that their periods are greater than the relaxation time, that is $T_i > T_{relax}$, with $T_i=2\pi/\Omega_i$ ($i=1,2$).

2. Conditional probabilities of spike generation

To obtain an explicit expression of Eq. (16) we need to evaluate the terms

$$\text{Prob}\{v_3(t) \geq v_{th}\}, \quad (17)$$

which are the conditional probabilities of the spike generation of neuron N_3 given the occurrence of different events related to the four scenarios of the previous paragraph. After the refractory period and before any incoming spike, the interneuron membrane potential is equal to the Ornstein-Uhlenbeck process $v_3(t) = \sqrt{D_3}\zeta_3(t) = \sqrt{D_3}\int_{t_3}^t e^{-\mu_3(t-t')} \xi_3(t') dt'$, which has a Gaussian distribution and it is asymptotically stationary. Now because the refractory period T_{ref} is long enough with respect to the time scale of the Ornstein-Uhlenbeck process $\tau_3 = 1/\mu_3$, the process $v_3(t)$ can be considered a stationary process. Once an external spike is arrived at the interneuron, for example, from the sensory neuron N_1 , the membrane potential $v_3(t)$ performs a jump to the value $k_1 + \sqrt{D_3}\zeta_3(t)$, and the process $v_3(t)$ remains stationary after the jump. In other words, after the refractory period, the process $v_3(t)$ can be considered a stationary process, whether or not a jump occurs, until another spike is produced at the interneuron N_3 .

Therefore, by setting

$$(v_{th} - k_i - k_j e^{-\mu_3(t-t')}) = f(k_i, k_j), \quad (18)$$

we obtain the expression of the conditional probabilities for all scenarios

$$\begin{aligned} \text{Prob}\left\{\zeta_3(t) \geq \left(\frac{f(k_i, k_j)}{\sqrt{D_3}}\right)\right\} \\ = \int_{f(k_i, k_j)/\sqrt{D_3}}^{\infty} w_{\zeta_3}^{st}(s) ds \\ = \frac{1}{2} \text{erfc}\left\{\sqrt{\frac{\mu_3}{D_3}}(v_{th} - k_i - k_j e^{-\mu_3(t-t')})\right\}, \end{aligned} \quad (19)$$

where $i \neq j$, and $i, j = 1, 2$. Here

$$w_{\zeta_3}^{st}(s) = \sqrt{\frac{\mu_3}{\pi}} \exp(-\mu_3 s^2) \quad (20)$$

is the stationary probability distribution of the noise amplitude [36], and $\text{erfc}(x)$ is the complementary error function. We have

- (i) for scenario 1, $k_i = k_1$, $k_j = k_2 = 0$;
- (ii) for scenario 2, $k_i = k_2$, $k_j = k_1 = 0$;
- (iii) for scenario 3, $k_i = k_1$, $k_j = k_2$; and
- (iv) for scenario 4, $k_i = k_2$, $k_j = k_1$.

We note that

$$\begin{aligned} \text{Prob}\{\zeta_3(t) \geq (v_{th} - k_i)/\sqrt{D_3}\} \\ = 1 - \text{Prob}\{\zeta_3(t) < (v_{th} - k_i)/\sqrt{D_3}\}. \end{aligned} \quad (21)$$

$(i = 1, 2).$

Moreover, because $v_3(t)$ becomes a stationary process [see

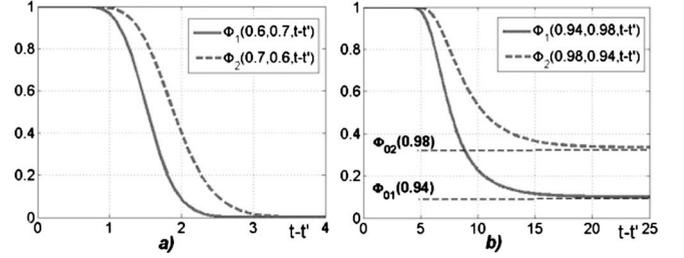


FIG. 3. Probabilities $\Phi_i(k_i, k_j, t-t')$ as functions of the time difference $(t-t')$ for weak and strong connectivity. (a) *Weak connectivity*. The values of the coupling coefficients are: $k_1=0.6$, $k_2=0.7$. The probabilities relax quite rapidly to the zero value. Only two incoming spikes very close in time should be able to fire the interneuron. (b) *Strong connectivity*. The values of the coupling coefficients are: $k_1=0.94$, $k_2=0.98$. The time dependent probabilities relax slowly toward the values of conditional probabilities of “separate” spikes. Two consecutive incoming spikes from sensors have more time and greater probability to fire the interneuron. The values of the other parameters are: $D=1.6 \times 10^{-3}$, $\mu=0.3665$, and $v_{th}=1$.

Eqs. (19) and (20)] we can factorize the probabilities of Eq. (21) out the integral in Eq. (16). Therefore, the probability distribution of spiking of the interneuron $\rho_3(t)$ [Eq. (16)] becomes

$$\begin{aligned} \rho_3(t) = \rho_1(t)\Phi_{01}(k_1) + \rho_2(t)\Phi_{02}(k_2) + \rho_1(t)[1 - \Phi_{02}(k_2)] \\ \times \int_{t-T_{relax_2}}^t \rho_2(t')\Phi_1(k_1, k_2, t-t') dt' + \rho_2(t)[1 \\ - \Phi_{01}(k_1)] \int_{t-T_{relax_1}}^t \rho_1(t')\Phi_2(k_2, k_1, t-t') dt', \end{aligned} \quad (22)$$

where

$$\Phi_{0i}(k_i) = \text{Prob}\{\zeta_3(t) \geq (v_{th} - k_i)/\sqrt{D_3}\}, \quad (i = 1, 2) \quad (23)$$

and

$$\Phi_i(k_i, k_j, t-t') = \text{Prob}\left\{\zeta_3(t) \geq \frac{(v_{th} - k_i - k_j e^{-\mu_3(t-t')})}{\sqrt{D_3}}\right\}, \quad (24)$$

with $i \neq j$, and $i, j = 1, 2$.

The time behavior of the probabilities $\Phi_i(k_i, k_j, t-t')$ as a function of the time difference $(t-t')$ is shown in Fig. 3, for weak and strong connectivities. The values of the coupling coefficients influence the relaxation times of the membrane potential of the neuron N_3 after a jump and at the same time the height of the jump. For weak connectivity (case a) in Fig. 3, the probabilities relax quite rapidly to the zero value. After a jump at the neuron N_3 , due to an incoming spike from one sensor, there is a very short interval $(t-t')$ for which a second incoming spike can fire the interneuron. Only two incoming spikes very close in time should be able to fire the interneuron. For strong connectivity (case b) in Fig. 3, the conditional probabilities relax slowly toward the conditional probabilities of “separate” spikes. Two consecutive incoming spikes from sensors have more time and greater probab-

ity to fire the interneuron. Moreover, any *separate* spike is able to fire the interneuron.

3. Probability distribution of first passage time

Because we considered only four scenarios to derive the probability distribution of spike generation (PDSG) at the interneuron, and we neglect events with low probabilities, we expect that the probability distribution calculated with the Eq. (22) is not normalized. We call the normalized PDSG as $\tilde{\rho}_3(t) = \rho_3(t) / (\int_0^\infty \rho_3(t') dt')$. To obtain the first passage time probability distribution of the interneuron, we should evaluate the joint probability that the interneuron fires at time t and no firing occurs in the previous time interval $[0, t]$. We then obtain

$$\hat{\rho}_3(t) = \tilde{\rho}_3(t) \left(1 - \int_0^t \tilde{\rho}_3(t') dt' \right). \quad (25)$$

Now we should recall that all previous calculations are valid until the first spike is generated by the third neuron N_3 . The question is: what happens after? Every time the interneuron fires because of the incoming spike from one of two sensors, the membrane potential $v_3(t)$ of the neuron N_3 is reset to the initial value v_3^0 , and the interneuron “forgets” all previous history. Suppose that the spike at the neuron N_3 is due to a spike of the neuron N_1 exactly at the same moment. Consequently, after reset, the FPTPD $\rho_1(t)$ has the same shape as it was previously. The other sensor N_2 is not reset synchronously with $v_1(t)$ and $v_3(t)$. Therefore the FPTPD $\rho_2(t)$ is shifted, in comparison with the initial pattern, because its starting time coincides with the reset of the interneuron. As a consequence from Eq. (22) we obtain a different FPTPD of the interneuron, which represents some *state* s of the interneuron and we call it $\rho_3^{(s)}(t)$. We use the superscript $(\cdot)^{(0)}$ to denote the initial situation (before the first spike was generated at the neuron N_3), which is the *state* 0. Every time the interneuron fires we have a time reset and one of the FPTPD of the two sensors is shifted, giving rise to a different FPTPD $\rho_3^{(s)}(t)$. In other words, after each reset of the interneuron, one of the two $\rho_i(t)$ ($i=1,2$) is similar to its initial form (corresponding to the state 0), while the other one is shifted to the left or to the right in time. As a consequence we can obtain from Eq. (22) all different states of $\rho_3^{(s)}(t)$ corresponding to different patterns in input to the interneuron. We note that the areas of the peaks of $\rho_3^{(s)}(t)$ give the probabilities of switching between states.

For sinusoidal inputs and a finite number of sensors we have a finite number of these states. Specifically let us consider two sensors with input sinusoidal signals of different frequencies $\Omega_1 \neq \Omega_2$, and we suppose that these frequencies are in a ratio of integer numbers m and n ; i.e., $\Omega_1/\Omega_2 = m/n$. Then $T_2/T_1 = m/n$, and after time $T_0 = nT_2 = mT_1$ we have coincidence of the m^{th} peak of $\rho_1(t)$ with the n^{th} peak of $\rho_2(t)$. This coincidence of peaks is the same as in the initial time $t=0$. In other words the first $(m-1)$ peaks of $\rho_1(t)$ do not coincide with the first $(n-1)$ peaks of $\rho_2(t)$. The patterns of $\rho_1(t)$ and $\rho_2(t)$ are repeated periodically in time with period T_0 . Each of these distinct peaks gives a firing probability of the interneuron and therefore a reset with a shift of one of

the two $\rho_i(t)$, which is a *state* of the interneuron. Consequently, the interneuron has $M = (m-1) + (n-1) + 1 = m+n-1$ different states corresponding to the peaks of $\rho_1(t)$ and $\rho_2(t)$. We note that the sequence of the states repeats periodically in time with the overall period T_0 , which is the period of phase coincidence of $\cos \Omega_1 t$ and $\cos \Omega_2 t$.

Hence, we obtain the interspike intervals distribution at the output of the interneuron by averaging the FPTPD $\hat{\rho}_3(t)$ of Eq. (25) over all these different M states

$$\rho_{out}(t) = \langle \hat{\rho}_3(t) \rangle = a_0 \hat{\rho}_3^{(0)}(t) + a_1 \hat{\rho}_3^{(1)}(t) + a_2 \hat{\rho}_3^{(2)}(t) + \dots + a_{M-1} \hat{\rho}_3^{(M-1)}(t). \quad (26)$$

The coefficients a_s denote the relative frequencies of the s states. In the case of musical accords and strong connections between neurons, the system visits all $m+n-1$ states almost uniformly; that is, all states give almost equal contributions into the expression of $\rho_{out}(t)$. So, as a first good approximation we consider all these coefficient equal, so $a_s = 1/M$, $\forall s$ and the Eq. (26) becomes

$$\rho_{out}(t) = \frac{1}{M} \sum_{s=0}^{M-1} \hat{\rho}_3^{(s)}(t). \quad (27)$$

Now, as an example, we consider a particular accord, namely, the “perfect fourth,” which consists of two sinusoids of frequencies related by the ratio $\Omega_1/\Omega_2 = 4/3$. The FPTPDs $\rho_1(t)$ and $\rho_2(t)$ are obtained from numerical simulations of the stochastic differential equations of the two sensors of Eq. (1) and are reported in Fig. 4. Here in Fig. 4(a) we have the representation of the state 0 of the interneuron, in which the fourth peak of $\rho_1(t)$ coincides with the third peak of $\rho_2(t)$ at time $T_0 = 4T_1 = 3T_2 = 41.89$ (with $T_1 = 10.47$, $\Omega_1 = 0.6$ and $T_2 = 13.96$, $\Omega_2 = 0.45$). All the peaks before T_0 do not coincide, so we have $M = 4 + 3 - 1 = 6$ distinct peaks in both $\rho_1(t)$ and $\rho_2(t)$ and therefore the same number of different states of the interneuron. In Fig. 4 these peaks are marked by numbers, representing the corresponding states of the interneuron. The area under each peak gives the probability finding an incoming spike at the defined short period of time. If this spike is able to fire the interneuron then the system is switched into a new state. In Fig. 4(a) the most probable and close in time spike comes from the sensor N_1 . If this spike fires the interneuron, the FPTPD will be shifted, and then the system is switched from state 0 to the state 1 [see Fig. 4(b)], where the most probable and close spike comes from the sensor N_2 . However, this spike, which will be generated with a certain probability, arrives during the refractory period of neuron N_3 , then the closest spike in state 1 is still by the sensor N_1 , and if it is able to fire the neuron N_3 , then it switches the system into state 3, and so on.

We note that the peak of $\rho_1(t)$ or $\rho_2(t)$, which appears during the refractory period, is one of the main reasons why $\rho_3(t)$ is not normalized in Eq. (16).

IV. FLOWCHART OF THE THEORETICAL APPROACH

Summarizing the previous sections, let us present the described theoretical approach in the form of a flowchart for

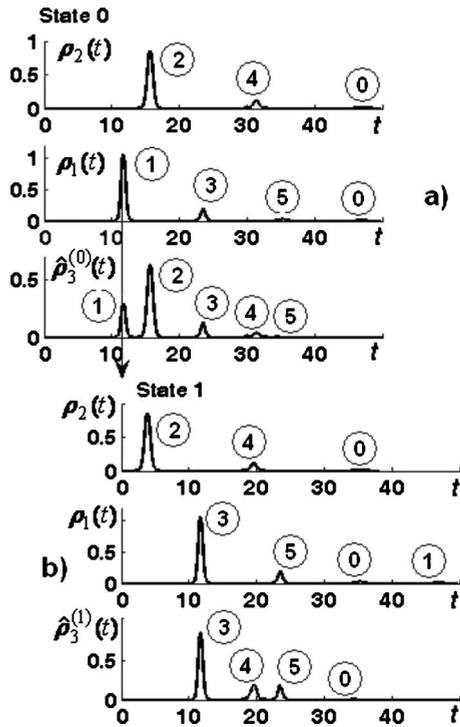


FIG. 4. FPTPDs of three neurons in states 0 and 1. This is the example of the probable transition between two states for the case of input sinusoids with frequencies related by ratio $4/3$ (the perfect fourth accord). $\rho_2(t)$ is shifted in state 1 in comparison with $\rho_2(t)$ in state 0. As a result, $\hat{\rho}_3^{(0)}(t)$ and $\hat{\rho}_3^{(1)}(t)$ are different. All possible states of the interneuron are: state 0, state 1, ..., and state 5. The peaks of $\rho_1(t)$ and $\rho_2(t)$ are marked by numbers in circles in order to establish the correspondence between them and the interneuron states.

calculation algorithm. Thus, in order to obtain the interneuron's ISID curve under chosen parameters of system (1) we should perform the following steps;

(1) to obtain the sensors' FPTPDs $\rho_1(t)$ and $\rho_2(t)$ using the direct numerical simulation of system (1) without the interneuron or the theoretical approaches described in Refs. [32,33];

(2) to find all possible *states* of the interneuron from the peaks of $\rho_1(t)$ and $\rho_2(t)$ (see Fig. 4, for example);

(3) to calculate $\hat{\rho}_3^{(s)}(t)$ in each *state* using Eqs. (22) and (25); and

(4) finally, to sum all the calculated $\hat{\rho}_3^{(s)}(t)$ and to normalize according to Eq. (27).

Despite the relative complexity of the algorithm, its usage decreases consumption of computer resources necessary to obtain good interneuron ISIDs and it is faster, in comparison with direct simulation of system (1). Moreover, the presented approach gives a consistent theoretical description of the noisy nonlinear system (see Fig. 1) modeled by Eq. (1).

V. NUMERICAL EXPERIMENTS

In this subsection we calculate numerically the interspike interval distributions of the interneuron for two groups of *consonant* and *dissonant* accords by numerical simulations

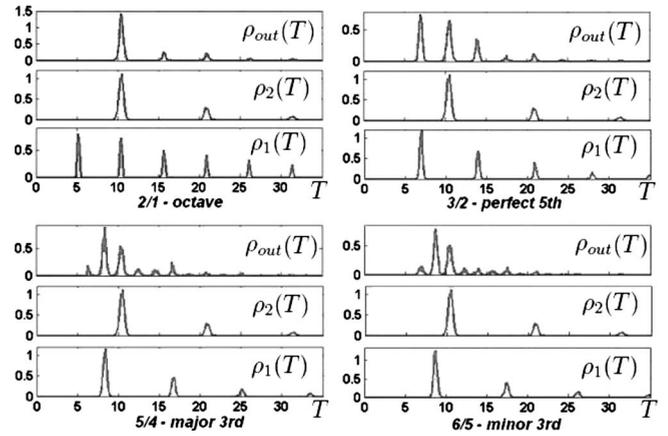


FIG. 5. Interspike interval distributions of the *consonant* accords: octave (2/1), perfect fifth (3/2), major third (5/4), and minor third (6/5). Under each picture there is the ratio of frequencies (m/n) and the name of the accord used in the common musical terminology. All curves are obtained through the direct numerical simulation of the Eq. (1) with $\Omega_2=0.6$, $A_2=1.165$, $\Omega_1=(m/n)\Omega_2$, and A_1 chosen according to the subthreshold input sinusoidal signal condition of Eq. (6). Specifically $A_1=1.52$ for octave (2/1), $A_1=1.325$ for perfect fifth (3/2), $A_1=1.243$ for major third (5/4), and $A_1=1.222$ for minor third (6/5). The other values of parameters are $\mu_1=\mu_2=1$, $\mu_3=0.3665$, $D_1=D_2=D_3=1.6 \times 10^{-3}$, $k_1=k_2=0.98$, $v_1^0=v_2^0=0$, $v_3^0=-1$, and $v_{th}=1$.

of Eq. (1) and compare them with the theoretical results expressed by Eqs. (22) and (27).

Particularly we consider the *consonant* accords: octave (2/1), perfect fifth (3/2), major third (5/4), and minor third (6/5), and the ISIDs are shown in Fig. 5. It is interesting to note the presence of new peaks in the distribution of $\rho_{out}(T)$, which are not present in the patterns of $\rho_1(T)$ and $\rho_2(T)$. Moreover, the ghost frequency $f_r=f_1-f_2$, characteristic of the ghost stochastic resonance (GSR) phenomenon, is present in all the *consonant* accords considered. Specifically in the octave and perfect fifth accords, in correspondence with the peaks at the ghost periods $T_r=10.52$ and $T_r=20.73$ ($T_r=1/f_r$), we have the ghost frequencies $f_r=0.095$ and $f_r=0.04823$, respectively. For the major third and minor third the ghost periods are $T_r=42.105$ and $T_r=52.63$, respectively, the relative peaks belong to the next time interval ($35 < T < 70$) and are very small. However, we should note that the noise intensity considered in all our simulations ($D=1.6 \times 10^{-3}$) is smaller than the suitable noise intensity to trigger the interneuron with an input signal coming from the sensory neuron. In other words we are out of the optimum range of noise intensities for GSR phenomenon.

Moreover we note the very regular behavior of the patterns of $\rho_{out}(T)$ in all the *consonant* accords considered, and particularly the very rich pattern with many peaks in the major third (5/4) and minor third (6/5) accords.

In Fig. 6 the ISIDs are shown for the following *dissonant* accords: major second (9/8), minor seventh (16/9), minor second (16/15), and augmented fourth (45/32). We see that in all *dissonant* accords considered, the interspike interval distribution has many new peaks not present in the patterns of $\rho_1(T)$ and $\rho_2(T)$. The ghost frequency is also present here,

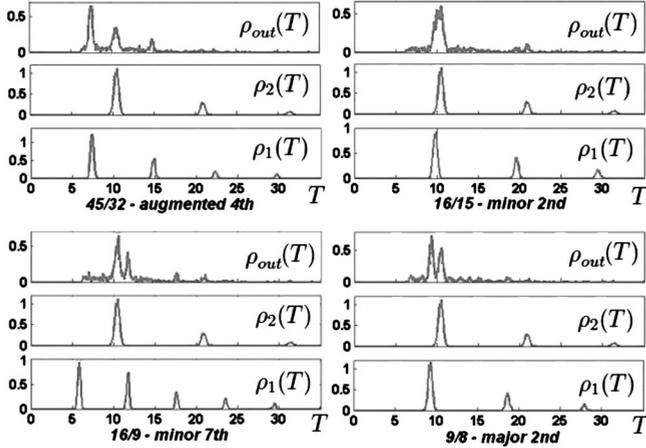


FIG. 6. Interspike interval distributions of the *dissonant* accords: major second (9/8), minor seventh (16/9), minor second (16/15), and augmented fourth (45/32). Under each picture there is the ratio of frequencies (m/n) and the name of the accord used in the common musical terminology. All curves are obtained through the direct numerical simulation of the Eq. (1) with $\Omega_2=0.6$, $A_2=1.165$, $\Omega_1=(m/n)\Omega_2$, and A_1 chosen according to the subthreshold input sinusoidal signal condition Eq. (6). Specifically $A_1=1.2$ for major second (9/8), $A_1=1.436$ for minor seventh (16/9), $A_1=1.17$ for minor second (16/15), $A_1=1.305$ for augmented fourth (45/32). The other values of parameters are the same as in Fig. 5.

but the heights of the corresponding peaks are less than those of the *consonant* accords of Fig. 5. Specifically the ghost frequency is visible in the augmented fourth and minor seventh accords with peaks at the ghost periods $T_r=25.91$ and $T_r=13.53$, respectively. For minor second and major second the ghost periods are out of the time interval considered in Fig. 6.

The ISIDs of *dissonant* accords are blurry with respect to the ISIDs of the *consonant* accords. This means that we can consider the ISID as an investigative tool to discriminate between *consonant* and *dissonant* accords. In fact higher are the integers m, n less regular and blurry are ISIDs, while lower are the integers and more regular are the ISIDs.

Finally in Fig. 7 the comparison between the theoretical ISIDs $\rho_{out}(T)$ (solid lines) for two *consonant* and one *dissonant* accords, given by Eqs. (22) and (27), and the numerical simulations (dashed lines) of Eq. (1) are shown. Particularly, we report from bottom to top in the figure, the *dissonant* major second (9/8), the *consonants* minor third (6/5) and the perfect fourth (4/3). The agreement between theoretical and numerical results is very good.

VI. HYPOTHESES ABOUT CONSONANCE AND DISSONANCE

There are only a few key assumptions that explain why humans and animals perceive the harmony or disharmony while listening to different complex tones.

Helmholtz in 1877 proposed the notion that *dissonance* arises due to beating between adjacent harmonics of complex tones. In effect, *dissonance* arises due to rapid amplitude fluctuations and it is proportional to the number of frequency

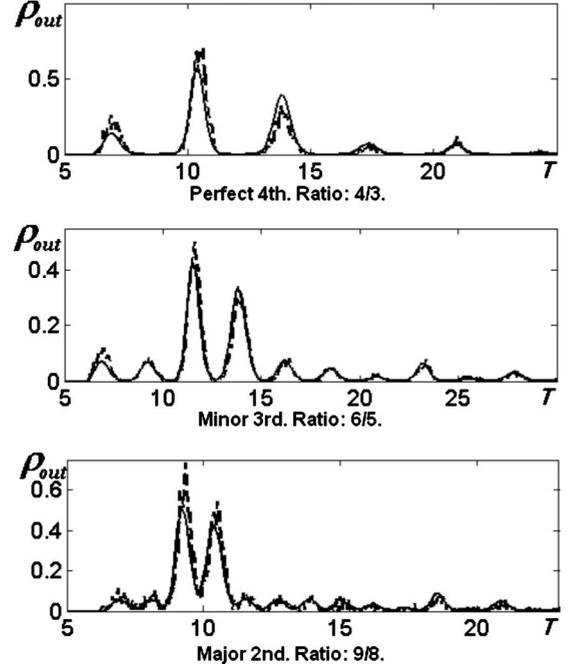


FIG. 7. ISI distributions of the output neuron for different accords. Solid lines are the theoretical results obtained from Eqs. (22), (25), and (27). Dashed lines are the distribution obtained throughout the direct numerical simulation of system (1). The parameters are $\mu_1=\mu_2=1$, $\mu=0.3665$, $k_1=k_2=0.97$, and $D_1=D_2=D=1.6 \times 10^{-3}$. **Perfect fourth:** $A_1=1.165$, $\Omega_1=0.6$, $A_2=1.085$, and $\Omega_2=0.45$. **Minor third:** $A_1=1.125$, $\Omega_1=0.54$, $A_2=1.085$, and $\Omega_2=0.45$. **Major second:** $A_1=1.2$, $\Omega_1=0.675$, $A_2=1.165$, and $\Omega_2=0.6$.

components present in the two complex tones producing beats [11].

Boomsliiter and Creel in 1961 suggested the long pattern hypothesis, which states that a *consonance* is based on the length of the overall period of a stimulus $T_0=nT_2=mT_1$. They showed that *consonant* intervals, based on simple integer ratios of fundamental frequencies, have shorter overall periods than those of *dissonant* intervals [37].

Here we propose that the input signals, which are transformed into spike trains with blurry or regular interspike interval distributions, are perceived as unpleasant (*dissonant*, inharmonious) or pleasant (*consonant*, harmonious) due to the analysis, recognition, and permanence in noisy environment occurring in the auditory system and then in the brain. This hypothesis is based on the simple auditory model of Fig. 1 and described by Eq. (1). A complex subthreshold input signal is transformed with the addition of environmental noise by the proposed sensory system into different types of spike trains, where the minimal distance between peaks in the probability distributions of spiking of the interneuron $\rho_{out}(T)$ is crucial. In fact by considering a complex input signal composed by two sinusoids with frequency ratio $\Omega_1/\Omega_2=m/n$, the minimal distance between peaks of $\rho_3^{(s)}(t)$ is $T_{min}=T_0/(mn)$. See the Appendix for the proof, also peaks 1 and 2 of the $\rho_3^{(0)}(t)$, and peaks 4 and 5 of the $\rho_3^{(1)}(t)$ in Fig. 4. This T_{min} defines the distance between peaks of the final ISID $\rho_{out}(t)$.

That is why the sufficiently high value of m gives raise to many peaks in a short distance and as a consequence pro-

duces blurry ISIDs, typical for *dissonant* accords. This means that even for pure input tones, which is without harmonics, and even if they are not close in frequency, in order to produce beats, we can perceive *dissonance*.

Therefore, the Helmholtz's hypothesis continues to be correct, if we consider the minimal distance among all peaks of $\rho_1(t)$ and $\rho_2(t)$ and not only the distance between the first peaks, which show the difference between frequencies of the input tones.

Moreover, the hypothesis of Boomsliker and Creel is also valid. Indeed, as we obtain from our model, the higher the integers m and n (*dissonant* accords) are, the higher the number of states ($m+n-1$) of the output neuron due to the patterns of $\rho_1(t)$ and $\rho_2(t)$ peaks, and the longer the overall period T_0 . Vice versa, for *consonant* accords we have shorter overall periods.

We note finally that because the output interneuron goes into a new state randomly and because of longer overall periods (*dissonance*), we need more longer realizations to obtain quite good, from statistical point of view, ISID $\rho_{out}(T)$. Conversely, in the case of *consonant* input, we obtain good statistics in short realizations.

VII. CONCLUSIONS

We have investigated by a probabilistic approach the phenomena of *dissonance* and *consonance* in a simple auditory sensory model, composed by two sensory neurons and one interneuron. Our theoretical approach is useful to investigate the interspike interval statistics of the spike train at the output of the interneuron by calculating the probability of spike generation at this output using conditional probabilities and the ISIDs at the output of sensory neurons. The proposed algorithm (see procedure of Sec. IV) allows us to quickly estimate the output distributions of the interneuron because of the steplike shape of the conditional probabilities and narrow peaks of FPTPDs $\rho_1(t)$ and $\rho_2(t)$. Moreover, this procedure is clear enough to be implemented in widely used programming environments. We find very good agreement between the theoretical results [Eq. (27)] and direct numerical simulations of the system of Langevin Eq. (1), which is computer time consuming with respect to the procedure of Sec. IV.

With our simple model of the auditory system, we are able to discriminate between *consonant* and *dissonant* accords by analyzing the first passage time probability distributions at the output of the interneuron. Blurry ISIDs characterize *dissonant* accords, while quite regular ISIDs characterize *consonant* accords (Figs. 5 and 6). By considering an extension of this simple model to a more complex realistic auditory system, composed of many sensory neurons and different layers (surface and deep layers in the neural system), we should be able to know at which extent the *dissonant* accords will "survive," against the *consonant* ones, in the noisy neural environment of the brain. The obtained results may be applied also in the context of such recent studies on stimulus reconstruction from neural spike trains [38], where the information transmission under the noise influence is investigated. Another suitable context of applica-

tion is the continuous investigation of the neuron's behavior under the influence of a constant bombardment of inhibitory and excitatory postsynaptic potentials [39]. This resembles a background noise, which is typical for functioning conditions of the neocortical neurons.

In real life, a human being deals with relatively simple combinations of sinusoidal signals when listening to music. Musical accords, which are combinations of tones, are classified as *consonant* (pleasant, harmonious) or *dissonant* (unpleasant, disharmonious), depending on the ratio between frequencies [15]. Thus, use of musical notations appears to be convenient in the context of our work to classify input signals. However, we should note that our results are obtained using the so-called "just intonation" musical accords, which are not appropriate for modern music.

Consonance and *dissonance* of accords are also recognized by animals, which never deal with music [10]. So, the underlying mechanism seems to be common and fundamental for the auditory neural system of mammals. This is a good reason to use our neural-like model of the auditory system to investigate the propagation of signals, such as musical accords, through a noisy nonlinear environment. Investigation of this process can help to understand which types of input signals are able to survive in the noisy environment of the brain, what is the mechanism of this process, and what does it mean from the perceptual and cognitive point of view.

Finally we note that the noise-induced phenomena such as coherence resonance, ghost stochastic resonance, aperiodic stochastic resonance, suprathreshold stochastic resonance, and noise enhanced stability are appropriate candidates for a solution of signal propagation and signal "survival" problems. But they allow revelation of very particular peculiarities of signal propagation through the nonlinear noisy environment of neural-like systems and do not provide a full statistical picture. However, in this paper we present a statistical analytical description of the transformation process for spike trains propagating from one layer of neurons to another one under the influence of noise. Our hypothesis is based on the positive role of environmental noise in the recognition and possible permanence of information, contained in complex input signals such as *consonant* or *dissonant* accords, occurring in the brain.

ACKNOWLEDGMENTS

Authors acknowledge the financial support by MIUR and by Federal purpose-oriented program "Scientific and Teaching staff of innovative Russia" 2009–2013 (State Contract No. П457).

APPENDIX: MINIMAL DISTANCE BETWEEN PEAKS

The question about the minimal distance between peaks of $\rho_3^{(i)}(t)$ can be considered as follows. Let us take the line segment of length L (Fig. 8). White dots divide it into n equal parts. Black dots divide it also into m equal parts. As a result, it is divided into $m+n-1$ different parts by both types of dots, and our aim is to find the minimal part. In our case $n < m < 2n$ and m/n is the simple ratio, i.e., $m = n+k$, where

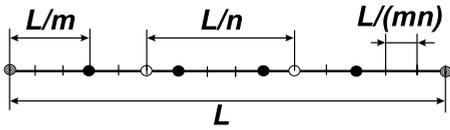


FIG. 8. Illustration for the procedure of minimal peak distance obtaining.

k is integer number and $k < n$. Let $i \in [0, n]$ is the number of white dot and $j \in [0, m]$ the number of black dot. Then, the distance between any two black and white dots is

$$|iL/n - jL/m| = |im - jn|L/(mn). \quad (\text{A1})$$

Now, i , j , m , and n are integers, therefore the minimal length of a segment part cannot be less than $L/(mn)$. Let us divide the whole segment into smallest parts $L/(mn)$, as it is shown in Fig. 8, and find the number l_i of such parts between the i th white dot and a previous black dot. Obviously, $l_0 = 0$ and $l_1 = k$. By trivial calculations we get the recurrent expression

$$l_{i+1} = k + l_i - \left\lfloor \frac{k + l_i}{n} \right\rfloor n, \quad (\text{A2})$$

where $\lfloor x \rfloor$ means the integer part of x . Equation (A2) can be also written as

$$l_{i+1} = (k + l_i) \bmod n. \quad (\text{A3})$$

It is obvious (see Fig. 8) that l_i takes values from 0 to $n - 1$. On the other hand, Eq. (A3) can be easily reduced to

$$l_i = (ik) \bmod n, \quad (\text{A4})$$

and this allows to find the period of l_i repetition by solving the simple equation $l_{i+p} = l_i$, where p is the period

$$[(i + p)k] \bmod n = (ik) \bmod n. \quad (\text{A5})$$

This gives $(pk) \bmod n = 0$, and we find that the period of l_i is $p = n$. Eventually, there are n possible values repeating with the period n , i.e., some i^* is always present, for which $l_{i^*} = 1$. Recalling Eq. (A1), this means that the minimal part of the L segment between white and black dots is $L/(mn)$. For the peaks of $\rho_3^{(i)}(t)$ we substitute $L = T_0 = nT_1 = mT_2$.

-
- [1] D. R. Chialvo, A. Longtin, and J. Muller-Gerking, *Phys. Rev. E* **55**, 1798 (1997).
- [2] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, *Rev. Mod. Phys.* **70**, 223 (1998).
- [3] A. S. Pikovsky and J. Kurths, *Phys. Rev. Lett.* **78**, 775 (1997).
- [4] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier, *Phys. Rep.* **392**, 321 (2004).
- [5] R. N. Mantegna and B. Spagnolo, *Phys. Rev. Lett.* **76**, 563 (1996); A. A. Dubkov, N. V. Agudov, and B. Spagnolo, *Phys. Rev. E* **69**, 061103 (2004).
- [6] E. V. Pankratova, A. V. Polovinkin, and B. Spagnolo, *Phys. Lett. A* **344**, 43 (2005).
- [7] W. Gerstner and W. Kistler, *Spiking Neuron Models. Single Neurons, Populations, Plasticity* (Cambridge University Press, Cambridge, 2002).
- [8] A. Aldo Faisal, Luc P. J. Selen, and Daniel M. Wolpert, *Nat. Rev. Neurosci.* **9**, 292 (2008).
- [9] A. Longtin, J. W. Middleton, J. Cieniak, and L. Maler, *Math. Biosci.* **214**, 87 (2008).
- [10] Y. I. Fishman, I. O. Volkov, M. D. Noh, P. C. Garell, H. Bakken, J. C. Arezzo, M. A. Howard, and M. Steinschneider, *J. Neurophysiol.* **86**, 2761 (2001).
- [11] H. L. F. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Dover, New York, 1954).
- [12] P. Cariani, *J. New Music Res.* **30**, 107 (2001).
- [13] I. S. Lots and L. Stone, *J. R. Soc., Interface* **5**, 1429 (2008).
- [14] B. Heffernan and A. Longtin, *J. Neurosci. Methods* **183**, 95 (2009).
- [15] R. Plomp and W. J. M. Levelt, *J. Acoust. Soc. Am.* **38**, 548 (1965).
- [16] P. A. Cariani, *J. Neural Transplant Plast.* **6**, 147 (1999).
- [17] M. J. Tramo, P. A. Cariani, B. Delgutte, and L. D. Braida, *Ann. N.Y. Acad. Sci.* **930**, 92 (2001).
- [18] C. J. Plack and A. J. Oxenham, in *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, R. R. Fay, and A. N. Popper (Springer-Verlag, New York, 2005), pp. 7–55.
- [19] A. de Cheveigné, in *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, R. R. Fay, and A. N. Popper (Springer-Verlag, New York, 2005), pp. 169–233.
- [20] D. Wile and E. Balaban, *PLoS ONE* **2**, e369 (2007).
- [21] D. R. Chialvo, O. Calvo, D. L. Gonzalez, O. Piro, and G. V. Savino, *Phys. Rev. E* **65**, 050902(R) (2002).
- [22] D. R. Chialvo, *Chaos* **13**, 1226 (2003).
- [23] J. M. Buldú, D. R. Chialvo, C. R. Mirasso, M. C. Torrent, and J. García-Ojalvo, *Europhys. Lett.* **64**, 178 (2003); G. Van der Sande, G. Verschaffelt, J. Danckaert, and C. R. Mirasso, *Phys. Rev. E* **72**, 016113 (2005).
- [24] O. Calvo and D. R. Chialvo, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **16**, 731 (2006).
- [25] P. Balenzuela and J. García-Ojalvo, *Chaos* **15**, 023903 (2005).
- [26] A. Lopera, J. M. Buldu, M. C. Torrent, D. R. Chialvo, and J. Garcia-Ojalvo, *Phys. Rev. E* **73**, 021101 (2006).
- [27] L. Glass and M. C. Mackey, *From Clocks to Chaos* (Princeton University Press, Princeton, 1988).
- [28] H. G. Schuster, *Deterministic Chaos* (VCH-Verlag, Weinheim, 1994).
- [29] H. C. Tuckwell, *Stochastic Processes in the Neurosciences* (SIAM, Philadelphia, PA, 1989).
- [30] W. Gerstner, R. Kempter, J. L. van Hemmen, and H. Wagner, *Nature (London)* **383**, 76 (1996).
- [31] P. Maršálek, C. Koch, and J. Maunsell, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 735 (1997).
- [32] H. E. Plesser and S. Tanaka, *Phys. Lett. A* **225**, 228 (1997); H. E. Plesser and T. Geisel, *Phys. Rev. E* **63**, 031916 (2001).
- [33] C. F. Lo and T. K. Chung, in *First Passage Time Problem for the Ornstein-Uhlenbeck Neuronal Model*, Lecture Notes in Computer Science, edited by King *et al.* (Springer-Verlag, Ber-

- lin Heidelberg, 2006) Vol. 4232, pp. 324–331, ICONIP 2006, Part I.
- [34] G. Maimon and J. A. Assad, *Neuron* **62**, 426 (2009).
- [35] D. Cox, *Renewal Theory* (Chapman and Hall, London, 1967).
- [36] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1985).
- [37] P. Boomsliter and W. Creel, *J. Music Ther.* **5**, 2 (1961).
- [38] A. Nikitin, N. G. Stocks, and R. P. Morse, *Phys. Rev. E* **75**, 021121 (2007).
- [39] S. Luccioli, T. Kreuz, and A. Torcini, *Phys. Rev. E* **73**, 041902 (2006).