How the Ear Tunes In to Sounds: A Physics Approach

Florian Gomez, Victor Saase, Nikolaus Buchheim, and Ruedi Stoop*

Institute of Neuroinformatics, University of Zurich and ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

(Received 16 December 2013; published 27 February 2014)

Listening is a complex sound selection process thought to be located in the auditory cortex. A biophysically motivated Hopf model of the mammalian cochlea reveals that pitch, a main characteristic in the perception of sound, is already materialized at the level of the mammalian hearing sensor. Here, we provide evidence that major elements of listening may similarly be implemented at the auditory periphery by means of efferent connections to the cochlea that tune the hearing sensor towards an auditory object of interest. The cochlea model we use in our investigations is advocated by its performance quality, the simplicity by which efferent control can be implemented, and by the closeness of the control results compared to the biological data. We tune the Hopf parameters to target on a sound, using pitch as the guiding feature. How well we achieve our goal is tested on real-world sounds and measured by a specifically developed tuning-error measure. The results provide a first estimate of how much the peripheral hearing system can assist a listener in focusing on an auditory signal and, thus, what is contributed by the auditory cortex.

DOI: 10.1103/PhysRevApplied.1.014003

I. INTRODUCTION

All the way up from the cochlea to the cortex, neural feedback loops provide efferent input to more peripheral parts of the auditory system. To date, their functional roles are not fully understood. The olivocochlear bundle, the most peripheral loop, originates in the superior olivary complex (SOC) in the brainstem. From the medial part of the SOC, thick myelinated neurons [the so-called medial olivocochlear (MOC) efferents] synapse on the cochlea's outer hair cells. Experiments indicate that MOC neuron stimulation provokes a rapid reduction in cochlear amplification [1–4]. Here we provide a first quantitative test to determine to what extent the MOC system can contribute to the stunning ability of mammals to follow signals of interest within sound mixtures. We use pitch, the overall characterization of a complex sound [5], as the guiding feature to target a sound and extract it from a mixture of sounds by using efferent control.

II. EFFERENT INPUT

Effects of efferent input to the cochlea have been studied at various levels, but so far, they seem not to have been related to an overall performance of the auditory system. At the biochemical level, efferent inhibition in the cochlea was investigated both *in vitro* and *in vivo* (see, e.g., Refs. [6,7]), with the aim of understanding the detailed neurophysiological basis of voltage changes in both inner and outer hair cells. Work at the physiological and biophysical description levels has recently provided a deeper understanding of the mechanics of the cochlear amplifier [8,9]. A full biophysical description of how efferent MOC activity eventually leads to a reduction in cochlear amplification is, however, still missing. Several MOC-related studies have taken reverse-engineered models of the auditory periphery as their starting point to circumvent the problems and the complexity that emerge from a more concise biophysical description [10–16]. In these approaches, a close-to-biology implementation of efferent control of the cochlea is inherently difficult, and in most approaches, the frequency specificity of efferent innervation [17] is not considered.

III. PERIPHERAL HEARING SENSOR

The cochlea model that we take as the starting point for our investigations is based on cochlear hydrodynamics as a passive system [18] combined with the fact that physical (or biological) systems close to a bifurcation can be used as small-signal amplifiers [19,20]. From this, a mesoscopic description of the problem emerges that, in contrast to the reverse-engineering approaches, retains a high level of verifiable biological detail. In the context of hearing, it is suggested that the prominent role of the outer hair cells in providing active amplification to the hearing sensor can, mesoscopically, be modeled by systems close to a Hopf bifurcation [21,22]. Indeed, such systems will generate the correct nonlinearities, compression rates, and sharper tuning for low-intensity sounds as observed in the mammalian cochlea [23–25]. While generally, the software "Hopf cochleas" developed on this basis required that the Hopf elements should exactly "be poised" at the bifurcation point

*ruedi@ini.phys.ethz.ch

[26,27], our model [28] has its Hopf elements tuned away from the bifurcation. We shall see that exactly this feature offers a simple, variable, and individual tuning scheme. We present here for the first time results of such an analysis, where we work with real-world sound examples to highlight the relevance of our approach to application.

In our Hopf cochlea, the active amplification part is described by an ω_{ch}-rescaled Hopf equation [29] $\dot{z} = (\mu + i)\omega_{\mathrm{ch}} z - \omega_{\mathrm{ch}}|z|^2 z - \omega_{\mathrm{ch}} F(t)$, with $z \in \mathbb{C}$. Here, $F(t)$ denotes the external forcing, ω_{ch} is the characteristic frequency (CF) of the oscillator, and μ is the parameter governing the distance to the Hopf bifurcation point (at which $\mu = 0$). Assuming a 1 : 1 locking between the system and forcing, and denoting with R and F the magnitudes of $z(t)$ and $F(t)$, respectively, we obtain at the bifurcation point a gain of $G = R/F \propto F^{-2/3}$ for ω close to resonance ω_{ch} [21]. This forces the gain to increase towards infinity if the stimulus size F approaches zero. Moving away from the bifurcation ($\mu < 0$) but maintaining $\omega \approx \omega_{\mathrm{ch}}$, a response $R \approx -F/\mu$ emerges for weak stimuli. Upon increasing the stimulus F, the compressive regime is entered and the differential gain decreases. Away from resonance, a linear response $R \approx F/(\omega - \omega_{\mathrm{ch}})$ emerges. A cascade of such Hopf oscillators with logarithmically spaced characteristic frequencies then composes the Hopf cochlea. Each oscillator is followed by a sixth-order Butterworth low-pass filter to model the viscous damping of the cochlear fluid. While a schematic overview on the cochlea's design is given in Fig. 2(b), the reader may consult Refs. [29–31] for more details, where detailed comparisons with the available biophysical data of mammalian cochlea are provided (most notably, in Ref. [31], Supplemental Material). Our Hopf amplifier concept comprises outer hair cell response as well as their embedding into the basilar membrane (BM). From this, the amplifier acquires tuning properties different from isolated biological outer hair cells. Our report is based on the software-implemented version of the original electronic hardware sensor. Compared to the hardware version, the software version yields identical results, except for the absence of noise, which permits the use of virtually as many cochlea sections as desired (note that 30 sections are sufficient to compete with the human sensor regarding amplification, frequency range, and resolution).

While the choice of pitch as the guiding sound feature may seem natural, getting the approach to work is not exactly straightforward. First, it must be shown that efferent stimulation implemented by tuning the μ parameters away from criticality reproduces the biological observations. Second, an efficient and reliable strategy to tune the cochlea towards the desired sound and a measure to assess to what degree this was achieved needs to be developed from scratch (we are unaware of a similar line of work). As a first example for the quality of our setting, we reproduce the measurements of Ref. [32], where pitch is represented as the inverse of the most frequent interspike interval of onset-L-cells of the ventral cochlear nucleus. The corresponding

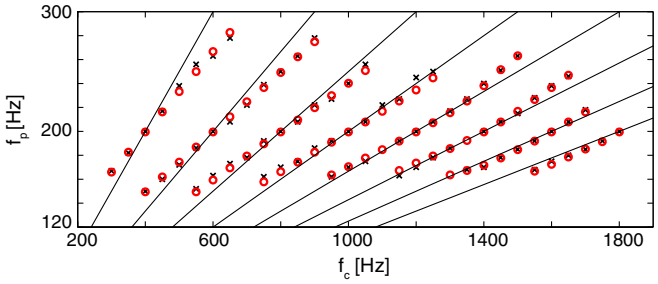

FIG. 1 (color online). "Perceived pitch" f_p in the cat ventral cochlear nucleus. Three-frequency stimulation $(f_c - f_{\mathrm{mod}})$, f_c, $(f_c + f_{\mathrm{mod}})$, $(f_{\mathrm{mod}} = 200$ Hz). Black crosses: Inverse of most frequent interspike intervals, onset-L-cell response [32] (CF = 1100 Hz, 50 dB sound pressure level). Red dots: f_p from section 11 of a full cochlea [[35] CF = 1136 Hz, −60 dB, cochlea settings as in Figs. 6 and 7, "flat tuning" (see text)].

data from our cochlea (where the pitch is calculated from the signal's autocorrelation peak [5], see below for details) provide full agreement (Fig. 1). Similar results are obtained for Smoorenburg's pitch shift experiments [33]. The straight lines in Fig. 1 are the expected results from de Boer's first pitch shift formula [34], a model based on the properties of the incoming complex sound only [33]. The deviations from these lines represent the second pitch shift, which is the consequence of the physical properties of the cochlea [35]. To recover it, a "full" cochlea (Hopf elements, proper coupling, and hydrodynamic properties) but no perceptional brain elements are required.

A. Efferent control

We now focus on how efferent MOC control can assist the mammalian cochlea to enhance desired sounds and suppress unwanted sounds in complex sound environments. In the Hopf cochlea, μ provides the natural parameter for efferent gain control. The correspondence between our model of efferent gain control and biology is exhibited in Fig. 2.

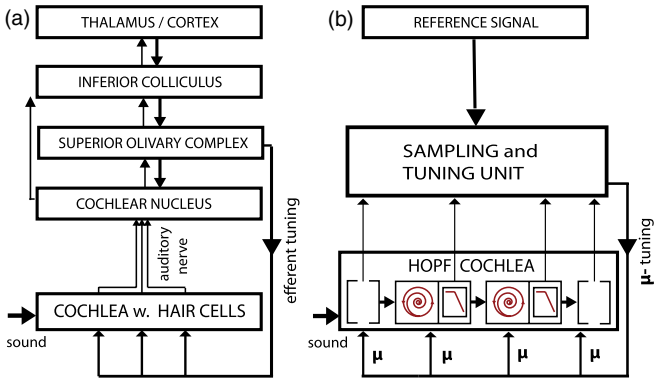

FIG. 2 (color online). Cochlea with efferent feedback. (a) Mammalian example. (b) Our model, where efferent inhibitory input to the cochlea's outer hair cells is realized by tuning the sectional Hopf parameters away from criticality ("μ tuning").

Whereas the amplification of a pure tone can simply be monitored via the change of the Hopf parameter of a section ("local tuning") for complex sounds, due to emergent combination tones, the monitoring of a whole pattern of Hopf parameters is required (cf. Fig. 2). This raises the question whether such control (1) will eventually lead to the desired effect, (2) how close to biology this will be, and (3) whether this will be computationally feasible. Over a wide range of input levels, the response at the intermediate section 5 of a full Hopf cochlea is shown in Fig. 3. The shape of the curves and the difference of 36 dB for two curves 80 dB apart are in full accordance with animal data [36]. To model frequency-specific MOC control, we compare "flat tuning" (i.e., where all μ parameters have the same value, $\mu = -0.1$) to the case where the section's μ is pushed further away to $\mu_5 = -1.0$. A substantial effect is observed, even for frequencies well below a section's CF. It was found in biological efferent stimulations [37] that the cochlear level shifts [38] were largest for stimulus frequencies below CF at low to moderate sound levels, whereas for frequencies above CF, the salient level shifts were obtained at higher levels. Our modeling results (Fig. 4) are fully consistent with these observations. Efferently low-frequency stimulated, high-CF fibers show substantial phase lags if compared to the unstimulated case ([39] auditory nerve measurements; corresponding cochlear evidence is yet missing). The phase changes that we obtain in our model for different μ values and frequencies below and above CF are shown in Fig. 5 (full Hopf cochlea model at −25 dB). Tuning the oscillator from $\mu = -0.1$ away from bifurcation generates phase delays for frequencies below CF and phase leads for frequencies above CF, as observed in biology. The sign change of the relative phase at CF for different sound levels [36] is reproduced as well (inset of Fig. 5).

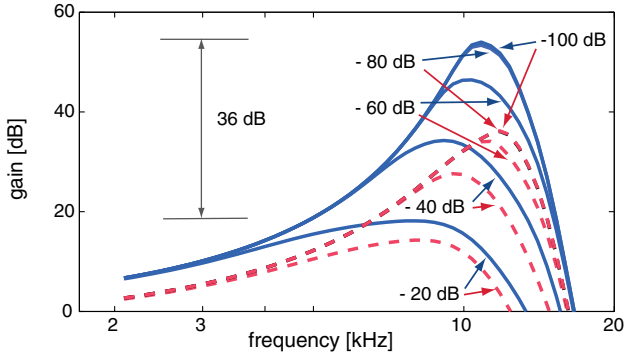

FIG. 3 (color online). Gain isointensity curves at section 5 of a full Hopf cochlea (CF = 10.42 kHz, 18 sections covering a range from 20 to 1.25 kHz) without (solid lines) and with (dashed lines) MOC efferent input. Full curves are fully consistent with animal data [36]; data corresponding to the dashed curves are not available. Flat tuning ($\mu = -0.1$) for the non-MOC case. MOC stimulation is implemented by shifting μ_5 to $\mu_5 = -1.0$ (curves for −80 and −100 dB collapse).

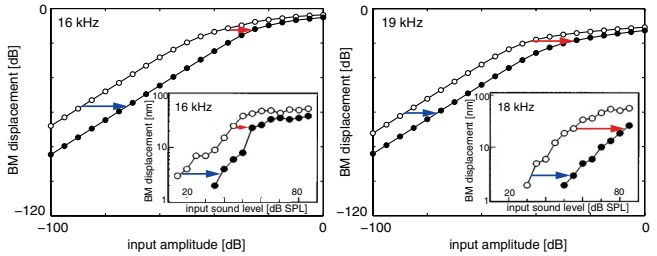

FIG. 4 (color online). BM level shifts (arrows) at section 2, CF = 16.99 kHz (18 sections, 20–1.25 kHz), when stimulated by a 16 and 19 kHz (left and right) pure tone. Open circles: Flat tuning ($\mu = -0.05$). Filled circles: MOC stimulation; μ_2 is shifted to −0.5. Insets: Corresponding animal data [37].

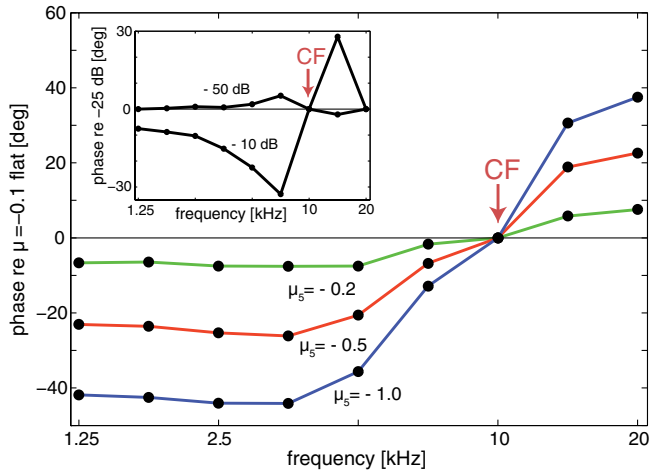

FIG. 5 (color online). Phase shift at section 5 (CF = 10.42 kHz), μ_5 is tuned away from flat tuning $\mu = -0.1$ (18 sections, range 20 to 1.25 kHz, stimulation at −25 dB). Phase delays result for frequencies below CF, phase leads above CF. Inset: Phase level dependence relative to −25 dB (single oscillator, CF = 10 kHz, $\mu = -0.1$). Decreased low-frequency input levels lead to small phase leads; increased input levels lead to phase lags.

B. Tuning in to sounds

After having checked that our efferent tuning replicates the properties known from the biological example, we now measure to what extent this may be helpful for extracting target signals from a mixture of sounds. Taking pitch as the guiding feature for listening implies taking the signal autocorrelation function (ACF) as the target. In our cochlea, this is represented by a summation of the signal autocorrelations over the different sections, which changes the ACF into the so-called summary autocorrelation function (SACF) [40–42]. The first prominent SACF peak again indicates the pitch evoked by the stimulus. However, we maintain the full normalized summary autocorrelation function (NSACF) that accounts for sound characteristics other than pitch as well (e.g., timbre). To measure how strongly a mixture of two input sounds x, y is biased ("tuned") towards signal component x, we use the Euclidean distance from the mixture NSACF to the

normalized autocorrelation function (NACF) of the target signal x divided by the Euclidean distance from the mixture NSACF to the NACF of the undesired signal y. The tuning error (TE)—the measure we develop to assess how close we arrive to the target—has the expression $\mathrm{TE}(x,y) \coloneqq [\|\mathrm{Normalized}\{\sum_i \mathrm{ACF}[f_i(x+y)]\} - \mathrm{NACF}(x)\|_2] / [\|\mathrm{Normalized}\{\sum_i \mathrm{ACF}[f_i(x+y)]\} - \mathrm{NACF}(y)\|_2]$, where f_i denotes the output at section i of the cochlea and where the summation extends over the N sections. The TE values are between 0 and ∞. TE = 0 indicates a perfect focus, and a larger TE is a less perfect focus on the target signal. If one source dominates the mixture, then TE values below unity may be observed even before tuning.

In this setting, listening amounts to finding the patterns of μ values that minimize the target's TE. For our results below, we use a state-of-the-art genetic algorithm [43], where we always start from a flat tuning of $\mu = -0.1$. The sound files are from two church organ pipes: a flute and a cornett having distinctly different sound timbres recorded at a sampling frequency of 44.1 kHz and equalized in loudness with respect to the sum of squared wave (.wav) coefficients. The autocorrelation is performed over 300 time steps. Since the frequencies in our tuning experiments are in the range of 270 Hz to 3 kHz, we choose a cochlea of 20 sections covering a range from 220 Hz to 7.040 kHz (5 octaves). The approach taken here is novel; the closest efforts that we know of are Refs. [14,16].

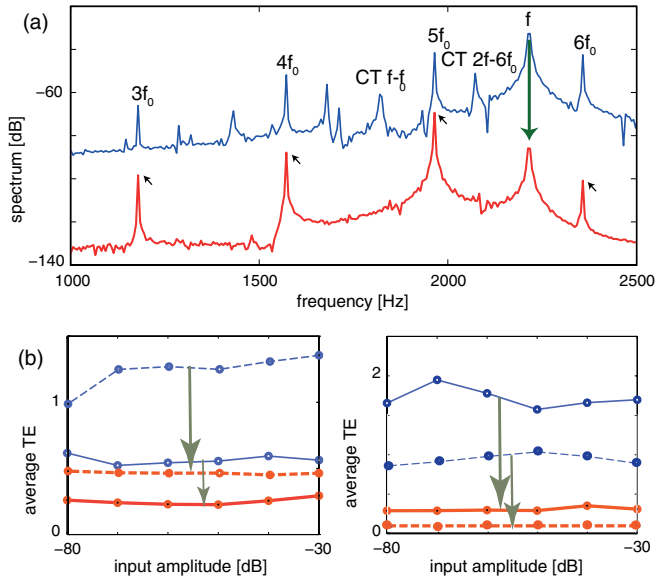

FIG. 6 (color online). TE improvement by μ tuning. (a) Frequency spectrum at section 8 (CF = 1964 Hz). Blue: Flat tuning (−80 dB, target cornett $f_0 = 392$ Hz, disturber flute $f = 2216$ Hz). Cross-combination tones (CT, two explicitly labeled) between the flute fundamental f and higher harmonics of the cornett. Red: Optimized tuning. f (flute) and cross-combination frequencies are suppressed, leaving a harmonic series of the target (small arrows). (b) Averaged TE over 13 different fundamental target frequencies (steps of 1 semitone) demonstrates input amplitude independence. Blue lines: Flat tuning. Red lines: Optimized μ tuning. Left panel: (full lines) target sound cornett (277 to 554 Hz), disturbing sound flute (at 277 Hz); (dashed lines) same target but flute at 2216 Hz. Right panel: Same experiment but target and disturber are interchanged. TE improvements: Arrows in (b).

IV. RESULTS

A. Stationary sounds

From time series containing 8192 measurements (185.8 ms) at each section, we discard the first 4096 elements to avoid, for simplicity, the computationally more demanding transient responses. When we disturb a cornett target signal with a flute at high pitch (interacting with the higher harmonics of the cornett), μ tuning not only suppresses the responses at the frequencies of the flute, it, moreover, cleans the signal from two-signal combination tones [Fig. 6(a)]. We emphasize that the target sound is enhanced via the suppression of undesired frequencies. For all tested pitch combinations of the two instruments, we

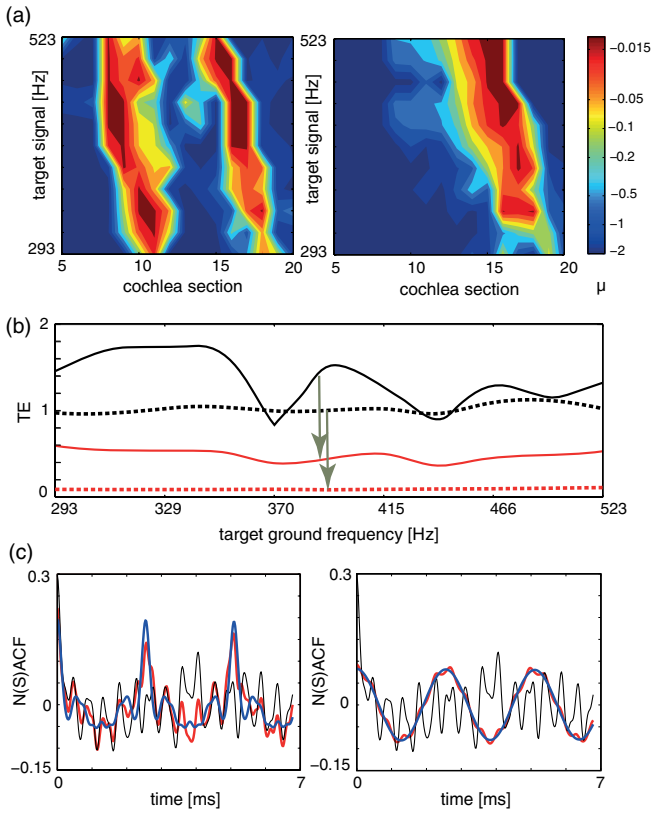

FIG. 7 (color online). (a) Tuning patterns, dynamical case. Colors indicate the Hopf parameter values of the sections. Left: Cornett vs flute (disturber). Right: Flute vs cornett (disturber). (b) TE for the two target signals of (a). Black: Flat tuning. Red: μ tuning. Full: Cornett target. Dashed: Flute target. (c) NSACF, NACF for the two target signals of (a),(b) at a chosen target ground frequency. Black: Flat tuning. Red: μ tuning. Blue: Target signal. Targets at 392 Hz, disturbers at 2216 Hz.

observe a substantial decrease in TE at all input sound levels [Fig. 7(b)]. This is nontrivial, as the strength of the generated combination tones depends in a nonlinear fashion on the stimulation amplitude. Moreover, the results are independent of the initial tuning (μ level).

B. Dynamical sounds

A convincing tuning strategy should be able to follow sounds that change in amplitude and frequency via simple enough tuning patterns. That this is possible is not obvious in light of the combination tones and suppression mechanisms inherent in our setting. We report here on experiments where a frequency-variable sound is to be separated from a stationary disturber. The result is the emergence of remarkably simple tuning patterns [Fig. 7(a)] that provide excellent signal separation results [Figs. 7(b) and 7(c)].

V. DISCUSSION AND CONCLUSIONS

Without particular parameter adjustment, our model of the auditory periphery show excellent agreement with animal BM response. As the main piece of evidence, we present level shifts of similar size as in biology [the largest level shifts for frequencies above CF are not at low sound levels but at high sound levels, and the effect of efferent MOC input is already effective for frequencies well below CF (Fig. 3)]. For efferent stimulation, the cochlear signal exhibits phase lags for tones that are as much as 3 to 4 octaves below CF. We finally reveal that simple tuning patterns provide substantial signal enhancement for frequency-variable sounds.

Regarding biological plausibility, the question remains whether feedback loops are strong enough to achieve the predicted effect and on what temporal scale the tuning would have to change. Whereas Fig. 4 indicates that our μ control parallels the biological experiment closely, the objection may arise that the electrical stimulations generally used in the biological experiments are not too close to the natural conditions. To be able to decisively answer this question, we lack dedicated biological data. The time scale of the tuning we report here (approximately 100 ms, the fast MOC effect) is slightly above the attention escape time, but it appears to be appropriate for tuning in on a quickly changing signal. Realistically, the temporal scale of μ tuning will be determined by the acoustic properties of the monitored signal according to which appropriate feedback loops will be chosen (SOC, other auditory nuclei, auditory cortex, neocortex, or combinations thereof, cf. Fig. 2). The correlation method that we use is independent of the time scale (coarse graining, subsampling, and sliding window smearing will yield a desired time scale); this is a main reason for monitoring the entire autocorrelation of the signal.

Using a mesoscopic approach we focus on a physiological connectivity, the role of which remains unclarified in many aspects. Frequency selectivity of the medial olivocochlear system and tonotopy of the cochlea provide us with an optimally designed machinery for efferent control. While our approach is strongly motivated by biology, we do not claim to implement every single biological detail. The strength of our mesoscopic approach is its ability to reproduce the salient biophysical and also psychoacoustic hearing features (see, also, Refs. [31,44]) and to demonstrate that frequency-specific μ tuning of the cochlea can be an effective means for focusing on desired sounds. A thorough physiological observation by Scharf *et al.* [45] supports our approach. In patients, the main effect of the removal of MOC connections is that the patients detect signals at unexpected frequencies better than before, implying an impaired attentional ability in the frequency domain. In our approach, the cochlea's SACF is guided by the target signal's ACF. In real-world applications, such a guidance may be based on past experience (we know what an instrument and speaker should sound like), may exploit particular information acquired at the beginning of the listening process, or may be a self-enhancing process. We often invest considerable effort "tuning in to" a target sound before we are able to follow it. Even without a biophysical justification of all modeling details, the efferent tuning of the cochlea as we demonstrate in this paper will be of great technological interest. Humanlike sound separation in cocktail-party environments and attending to a selected sound source are key abilities for artificial intelligence and for robotics [46]. By starting the sound separation and selection process already in the cochlea, the effort in dealing with a breadth of combination products among undesired and desired components is drastically reduced.

ACKNOWLEDGMENT

The work was partially supported by a grant (No. 200020-147010/1) of the Swiss National Science Foundation.

[1] R. Galambos, Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea, J. Neurophysiol. **19**, 424 (1956).
[2] D. Mountain, Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alter cochlear mechanics, Science **210**, 71 (1980).
[3] J. Siegel and D. Kim, Efferent neural control of cochlear mechanics? Olivocochlear bundle stimulation affects cochlear biomechanical nonlinearity, Hear. Res. **6**, 171 (1982).
[4] J. J. Guinan, Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans, Ear Hear. **27**, 589 (2006).
[5] J. Licklider, A duplex theory of pitch perception., Cell. Mol. Life Sci. **7**, 128 (1951).

[6] S. F. Maison, S. J. Pyott, A. L. Meredith, and M. C. Liberman, Olivocochlear suppression of outer hair cells in vivo: Evidence for combined action of BK and SK2 channels throughout the cochlea, J. Neurophysiol. **109**, 1525 (2013).
[7] E. Glowatzki and P. A. Fuchs, Cholinergic synaptic inhibition of inner hair cells in the neonatal mammalian cochlea, Science **288**, 2366 (2000).
[8] J. Ashmore, P. Avan, W. E. Brownell, P. Dallos, K. Dierkes, R. Fettiplace, K. Grosh, C. M. Hackney, A. J. Hudspeth, F. Jülicher, B. Lindner, P. Martin, J. Meaud, C. Petit, J. R. Santos Sacchi, and B. Canlon, The remarkable cochlear amplifier, Hear. Res. **266**, 1 (2010).
[9] D. Ó. Maoiléidigh and A. J. Hudspeth, Effects of cochlear loading on the motility of active outer hair cells, Proc. Natl. Acad. Sci. U.S.A. **110**, 5474 (2013).
[10] S. Jennings, M. Heinz, and E. Strickland, Evaluating adaptation and olivocochlear efferent feedback as potential explanations of psychophysical overshoot., J. Assoc. Res. Otolaryngol. **12**, 345 (2011).
[11] A. Chintanpalli, S. Jennings, M. Heinz, and E. Strickland, Modeling the anti-masking effects of the olivocochlear reflex in auditory nerve responses to tones in sustained noise, J. Assoc. Res. Otolaryngol. **13**, 219 (2012).
[12] R. T. Ferry and R. Meddis, A computer model of medial efferent suppression in the mammalian auditory system, J. Acoust. Soc. Am. **122**, 3519 (2007).
[13] G. J. Brown, R. T. Ferry, and R. Meddis, A computer model of auditory efferent suppression: Implications for the recognition of speech in noise, J. Acoust. Soc. Am. **127**, 943 (2010).
[14] N. R. Clark, G. J. Brown, T. Jurgens, and R. Meddis, A frequency-selective feedback model of auditory efferent suppression and its implications for the recognition of speech in noise, J. Acoust. Soc. Am. **132**, 1535 (2012).
[15] O. Ghitza, D. Messing, L. Delhorne, L. Braida, E. Bruckert, and M. M. Sondhi, in *Hearing—From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer, Berlin, 2007), pp. 541–550.
[16] D. P. Messing, L. Delhorne, E. Bruckert, L. D. Braida, and O. Ghitza, A non-linear efferent-inspired model of the auditory system; matching human confusions in stationary noise, Speech Communication **51**, 668 (2009).
[17] M. Liberman and M. Brown, Physiology and anatomy of single olivocochlear neurons in the cat, Hear. Res. **24**, 17 (1986).
[18] A. Kern, Ph.D. thesis, ETH Zurich, 2003.
[19] K. Wiesenfeld and B. McNamara, Period-doubling systems as small-signal amplifiers, Phys. Rev. Lett. **55**, 13 (1985).
[20] B. Derighetti, M. Ravani, R. Stoop, P. F. Meier, E. Brun, and R. Badii, Period-doubling lasers as small-signal detectors, Phys. Rev. Lett. **55**, 1746 (1985).
[21] V. M. Eguíluz, M. Ospeck, Y. Choe, A. J. Hudspeth, and M. O. Magnasco, Essential nonlinearities in hearing, Phys. Rev. Lett. **84**, 5232 (2000).
[22] S. Camalet, T. Duke, F. Jülicher, and J. Prost, Auditory sensitivity provided by self-tuned critical oscillations of hair cells, Proc. Natl. Acad. Sci. U.S.A. **97**, 3183 (2000).
[23] P. Martin and A. J. Hudspeth, Compressive nonlinearity in the hair bundle's active response to mechanical stimulation, Proc. Natl. Acad. Sci. U.S.A. **98**, 14386 (2001).
[24] J. Barral and P. Martin, Phantom tones and suppressive masking by active nonlinear oscillation of the hair-cell bundle, Proc. Natl. Acad. Sci. U.S.A. **109**, E1344 (2012).
[25] E. A. Lopez-Poveda, C. J. Plack, and R. Meddis, Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing, J. Acoust. Soc. Am. **113**, 951 (2003).
[26] M. O. Magnasco, A wave traveling over a Hopf instability shapes the cochlear tuning curve, Phys. Rev. Lett. **90**, 058101 (2003).
[27] T. Duke and F. Jülicher, Active traveling wave in the cochlea, Phys. Rev. Lett. **90**, 158101 (2003).
[28] A. Kern and R. Stoop, Essential role of couplings between hearing nonlinearities, Phys. Rev. Lett. **91**, 128101 (2003).
[29] S. Martignoli, J. -J. van der Vyver, A. Kern, Y. Uwate, and R. Stoop, Analog electronic cochlea with mammalian hearing characteristics, Appl. Phys. Lett. **91**, 064108 (2007).
[30] R. Stoop, T. Jasa, Y. Uwate, and S. Martignoli, From hearing to listening: Design and properties of an actively tunable electronic hearing sensor, Sensors **7**, 3287 (2007).
[31] S. Martignoli and R. Stoop, Local cochlear correlations of perceived pitch, Phys. Rev. Lett. **105**, 048101 (2010).
[32] W. S. Rhode, Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus, J. Acoust. Soc. Am. **97**, 2414 (1995).
[33] G. F. Smoorenburg, Pitch perception of two-frequency stimuli, J. Acoust. Soc. Am. **48**, 924 (1970).
[34] E. De Boer, in *Auditory System* (Springer, New York, 1976), pp. 479–583.
[35] F. Gomez and R. Stoop, Mammalian pitch sensation is essentially shaped by the cochlear fluid (to be published).
[36] M. A. Ruggero, N. C. Rich, A. Recio, S. S. Narayan, and L. Robles, Basilar-membrane responses to tones at the base of the chinchilla cochlea, J. Acoust. Soc. Am. **101**, 2151 (1997).
[37] I. J. Russell and E. Murugasu, Medial efferent inhibition suppresses basilar membrane responses to near characteristic frequency tones of moderate to high intensities, J. Acoust. Soc. Am. **102**, 1734 (1997).
[38] J. J. Guinan and K. M. Stankovic, Medial efferent inhibition produces the largest equivalent attenuations at moderate to high sound levels in cat auditory-nerve fibers, J. Acoust. Soc. Am. **100**, 1680 (1996).
[39] K. M. Stankovic and J. J. Guinan, Medial efferent effects on auditory-nerve responses to tail-frequency tones II: Alteration of phase, J. Acoust. Soc. Am. **108**, 664 (2000).
[40] R. Meddis and M. J. Hewitt, Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification, J. Acoust. Soc. Am. **89**, 2866 (1991).
[41] R. Meddis and L. P. O'Mard, A unitary model of pitch perception, J. Acoust. Soc. Am. **102**, 1811 (1997).
[42] P. A. Cariani and B. Delgutte, Neural correlates of the pitch of complex tones. I. Pitch and pitch salience, J. Neurophysiol. **76**, 1698 (1996).
[43] N. Hansen, in *Towards a New Evolutionary Computation*, edited by J. Lozano, P. Larranaga, I. Inza, and E.

Bengoetxea, Studies in Fuzziness and Soft Computing Vol. 192 (Springer, Berlin, 2006), pp. 75–102.

[44] S. Martignoli, F. Gomez, and R. Stoop, Pitch sensation involves stochastic resonance, Sci. Rep. **3**, 2676 (2013).

[45] B. Scharf, J. Magnan, and A. Chays, On the role of the olivocochlear bundle in hearing: 16 case studies, Hear. Res. **103**, 101 (1997).

[46] S. Haykin and Z. Chen, The cocktail party problem, Neural Comput. **17**, 1875 (2005).