nature electronics

Article https://doi.org/10.1038/s41928-023-00957-5

Neuromorphic acoustic sensing using an adaptive microelectromechanical cochlea with integrated feedback

In the format provided by the authors and unedited

Contents

1 Sensor response to extrinsic noise

Figure 1 shows the influence of extrinsic noise on the sensing signal. The experiments were done using the sound-optimized design detailed below in section 2. The input signal was a single tone (sine wave at the resonance frequency of the sensor 2.2 kHz) with added white noise (bandwidth 100 kHz). Due to the filtering of the sensor, the signal-to-noise ratio (SNR) of the output signal is constant over a large range of input SNR values. Increasing the feedback strength improves the SNR of the sensing signal, since it increases the Q factor. Above the bifurcation the sensing signal becomes independent of the input signal. Thus, a constant SNR is observed. SNR values were obtained applying the function `snr` (MATLAB) to the time series of input or output signals, respectively.

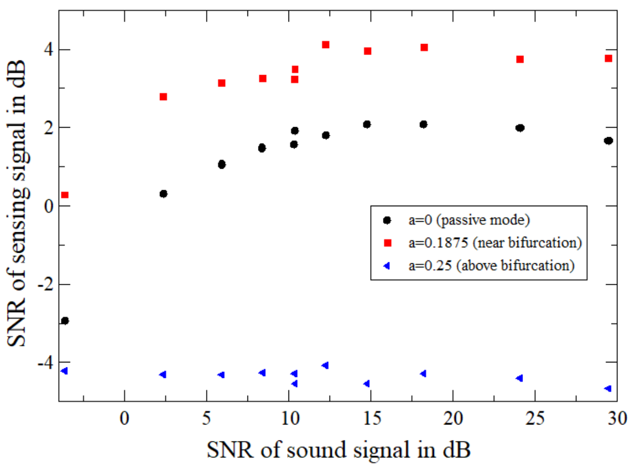

FIG. 1: Signal-to-noise ratio of sensor output in dependence of the signal-to-noise ratio of the sound signal for different feedback strengths a.

2 Sound-optimized sensor design

A new sensor design was developed to improve the sound sensing properties. Its physical dimensions are shown in Fig. 2, and the corresponding model parameters used in the simulation are given in Tab. 1.

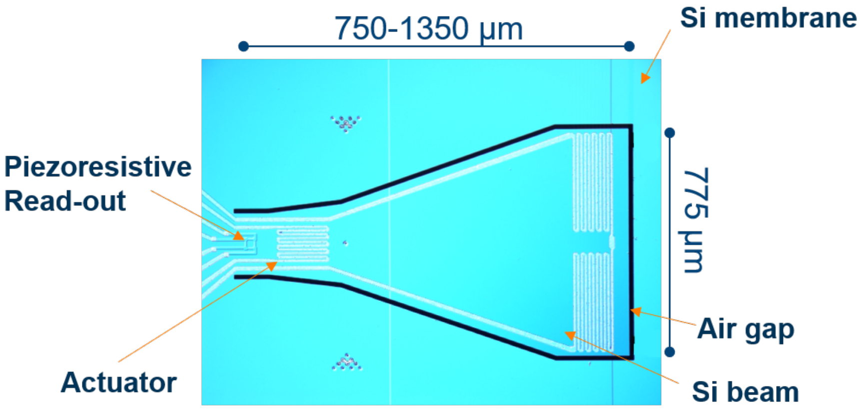

FIG. 2: Sound-optimized design with increased length and larger surface for lower resonance frequency and larger sensitivity.

TABLE 1: Parameters of the new, sound-optimized design for modelling the sensor system

variable	value
width of beam w_{Si}	$775\ \mu$m
length of beam l_{Si}	$1200\ \mu$m
thickness of beam d_{Si}	$2 - 10\ \mu$m
frequency of beam f	$2.2 - 6.9$ kHz
heater resistance R	$12\ \Omega$
calibration factor k (piezoresistive sensing and preamplification)	$0.6 \cdot 10^6$ V/m
time constant of high-pass filter τ	10^{-3} s
transfer factor α (temperature to deflection)	$19.2\,\mathrm{m}^2/\mathrm{s}^2\mathrm{K}$
time constant β (temperature change)	$1.0066 \cdot 10^3\,1/\mathrm{s}$
transfer factor γ (voltage to temperature)	$1.62 \cdot 10^7\ \mathrm{K}\Omega^2/\mathrm{sV}^2$

Acoustic properties of the new, sound-optimized design were tested in an anechoic chamber, see Fig. 3(left). The resulting properties are shown in the Fig. 3(right). As can be seen, the sensor gain in passive mode is 3 V/Pa and can be increased up to around 50 V/Pa in the active nonlinear mode. The dotted and dashed lines are fits to the data to guide the eye. The red, green, and gray shaded areas indicate the minimal detectable sound-pressure level (SPL), that is, the self-noise floor, of a measurement microphone (e.g. Microtech Gefell M221), a typical MEMS microphone [1], and the dynamic MEMS cochlea, respectively.

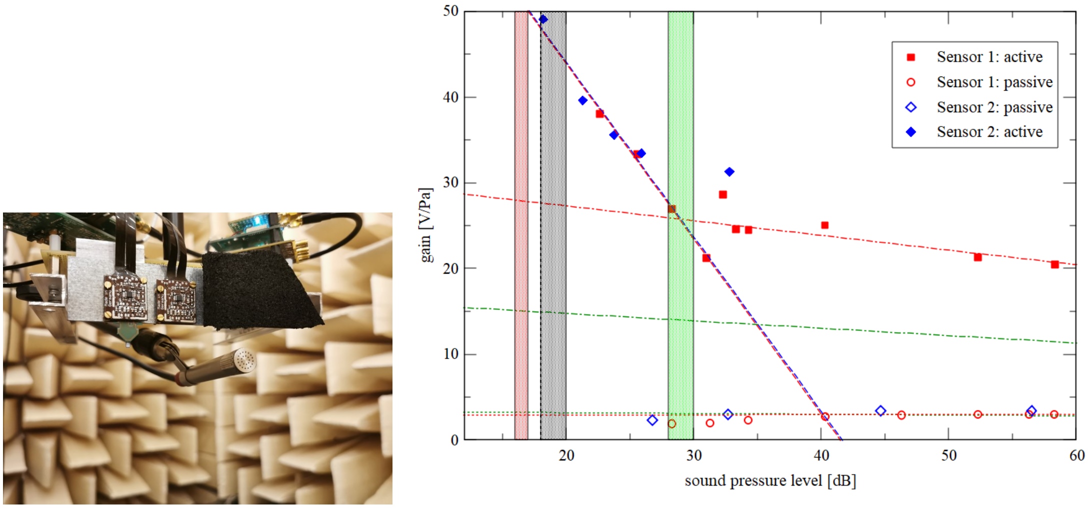

FIG. 3: (left) Measurement setup for determining the self-noise floor and the sensitivity of the sensor system in an anechoic chamber. (right) Sensor gain in dependence of the sound pressure level for two different sensors of the optimized design (cf. Fig. 2). In passive mode (open symbols), self-feedback is turned off. In active mode (full symbols), the feedback strength is larger than 0 and nonlinear characteristics are observed (SPL-dependent gain). Dotted and dashed lines are fits to the data. Shaded areas mark the self-noise floor, i.e., the minimal detectable SPL, of: a measurement microphone MT Gefell M 221 (red), a typical MEMS microphone (green) and the dynamic MEMS cochlea (gray).

3 Bifurcation analysis

In this section, we present a discussion of the bifurcation observed in the dynamic MEMS cochlea that gives rise to periodic oscillations. First, we will briefly derive a formula for the fixed point. Then, we elaborate on the transition to a limit cycle by means of a Hopf bifurcation.

Setting all time derivatives in Eqs. (3) of the main text to zero (cf. section *Theoretical description* in *Methods*) yields the following analytical expression for the fixed point x^*, that is, the deflection of the sensors free end:

$$x^* = \frac{\alpha\gamma}{\omega_0^2\beta}\left(\frac{u_{\rm DC}}{R}\right)^2 . \tag{1}$$

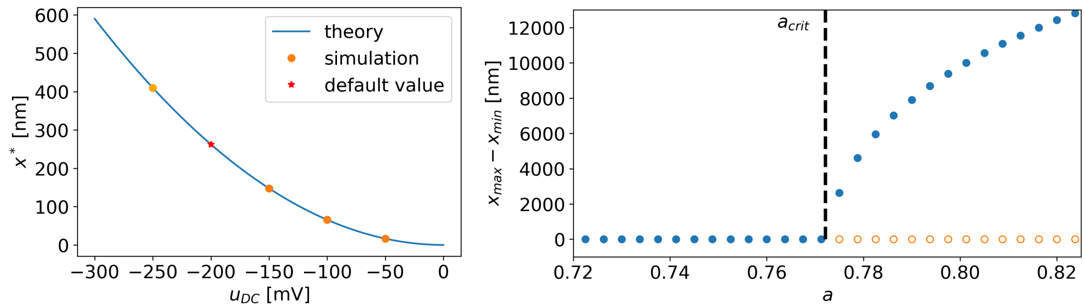

FIG. 4: (left): Fixed point x^* vs. bias voltage $u_{\rm DC}$. The blue curve corresponds to the analytical formula (1), and the symbols depict data from numerical simulations. The feedback strength is fixed at $a = 0.6$. (right): amplitude of oscillations ($x_{\rm max} - x_{\rm min}$) in dependence on the feedback strength a for a fixed bias voltage $u_{\rm DC} = -200\,\rm mV$ (default value in left panel). All other parameters as in Tab. I of the section *Theoretical description* in *Methods*. Filled/empty dots correspond to stable/unstable solutions. The dashed vertical line marks the Hopf bifurcation point $a_{\rm crit}$.

Figure 4(left) depicts the value of the fixed point in dependence on the bias voltage $u_{\rm DC}$ for a fixed feedback strength $a = 0.6$, i.e., below the bifurcation. Equation (1) is given as blue curve and the results of numerical simulations are shown as symbols. The red star for $u_{\rm DC} = -200\,\rm mV$ refers to the default value used in the present study.

Figure 4(right) displays numerical results that illustrate the emergence of periodic oscillations in a Hopf bifurcation. Starting from a stable fixed point, a limit cycle exists for increasing feedback strength a above the bifurcation point. The bias voltage is fixed at $u_{\rm DC} = -200\,\rm mV$. The bifurcation point $a_{\rm crit} \approx 0.772$ is confirmed by the bifurcation continuation software AUTO [2]. Above $a_{\rm crit}$, we find a limit cycle, whose amplitude is plotted as $x_{\rm max} - x_{\rm min}$, and the fixed point becomes unstable (empty dots).

4 Comparison to bio-inspired acoustic sensing systems

In this section, we compare basic properties of bio-inspired sensor systems with human cochlea as well as the developed, dynamic MEMS cochlea, see Tab. 2. Here, we took into account only bio-inspired sensors, which incorporate all three functionalities, namely frequency decomposition, nonlinear amplification and adaptation. These are the resonant gate transistor (RGT) microphone by Koickal et al. [3], the tunable membrane by Guerreiro et al. [4] and the basilar-membrane (BM) like sensor by Tsuji and Yamazaki et al. [5, 6]. In all three devices, the transducer itself implements the above described functionalities, similar to our developed sensor system.

Furthermore, we compare the dynamic MEMS cochlea system to analog cochleas that have been reported as (i) the Dynamic Audio Sensor (DAS) by Liu et al. [7] and its low-power version by Yang et al. [8], (ii) the Hopf cochlea by Hamilton et al. [9, 10] and (iii) the bionic ear processor by Sarpeshkar et al. [11], as well as to digital cochleas: (a) the CAR cochlea, the CAR-FAC cochlea and the CAR-Lite cochlea by Xu, Thakur, and Singh et al. [12–15] and (b) the binaural neuromorphic auditory sensor (NAS) by Jimenez-Fernandez et al. [16]. In all these devices, the acoustic signal is sensed and transduced by a 'standard' broadband microphone and the functionalities like frequency decomposition, nonlinear amplification and adaptation are achieved by subsequent circuits or FPGA calculations.

The developed, dynamic MEMS cochlea considered in the present study has several advantages: (i) large change of gain by feedback comparable to the human cochlea, (ii) simple feedback calculation that requires only a small amount of transistors (in comparison to the more complex leaky-integrate-and-fire (LIF) based or sound detection-based feedback calculations [3, 4, 6]), (iii) fast feedback calculation (below μs), (iv) fast adaptation for the feedback values (μs-ms), (v) relatively small actuation voltages (below 1 V), (vii) negligible latency for processing (filtering and amplification) after transduction, and (ix) bandwidth tuning. Furthermore, the standardized fabrication process chain based on CMOS technology allows for fabrication of a large number of sensors in one run (about 1000 sensors per wafer). Since the feedback parameters can be tuned in a large range to obtain the desired dynamics, the system has a high resilience against tolerances in device properties (in the feedback circuit as well as the transducer level). Both features together, make the system easily scalable by extension to multi-beam systems, enabling the coverage of large frequency ranges.

In contrast, alternative cochlea concepts have a number of drawbacks. For instance, only a small number of RGT microphones have been developed, and a full system with adaptation is yet to be realized. Guerreiro et al. reported on the realization of 1-2 single sensors only, which require relatively large actuation voltages (1 V-60 V) [4] and an exact timing of feedback spikes, making the feedback calculation more complex. The design proposed by Tsuji et al. comes with a slow (ms) and complex feedback calculation process [5]. All bio-inspired sensor systems achieve a relatively small gain increase only by the active process. Analog and digital cochleas include either complex circuits per channel, facing problems with device mismatch [7–11], or require complex calculations per channel [12, 13, 16, 17]. In both cases, limitations due to the broadband transduction by the microphone like masking of small signals by loud signals due to a fixed gain or saturation of microphone signals cannot be overcome.

TABLE 2: Comparison of human (mammalian) cochlea with dynamic MEMS cochlea (present study, gray), microphone, bio-inspired sensors, and artificial cochlea.

			bio-inspired sensors			artificial cochlea	
	human cochlea (mammalian) [18, 19]	**dynamic MEMS cochlea**	**resonant gate transistor (RGT) microphone [3, 20]**	**Guerreiro et al. [4]**	**Tsuji/Yamazaki et al. [5, 6]**	**analog cochlea [7, 9–11]**	**digital cochlea [12–16]**
frequency range	20 Hz-20 kHz (up to 100 kHz)	500 Hz-100 kHz	1-30 kHz	3-4 kHz	3-12 kHz	200 Hz-20 kHz [9, 10] 6-10 kHz [11]	20 Hz-20 kHz [12, 13], 9.6-14.6 kHz [16]
sensor principle	multiple sensors	multiple beams	multiple beams	multiple membranes	single membrane	1-2 microphones (typically MEMS)	1-2 microphones (typically MEMS)
number of sensors	3000 inner hair cells	> 50 realized and tested	10-20 sensors tested	2 sensors described	up to 12 electrodes in one system for read out	commercially available	commercially available
self-noise (minimal SPL)	0 dB SPL	18-20 dB SPL	n.a.	n.a.	50 dB SPL	24-30 dB SPL (MEMS microphone)	24-30 dB SPL (MEMS microphone)
maximal SPL or dynamic range (DR)	120 dB SPL (DR: 120dB)	> 110 dB SPL (DR: > 90 dB)	n.a.	n.a.	up to 105 dB SPL	DR: 77 dB [11],DR: 73 dB [8]	DR: 86dB [16], DR: 70dB [14]
size of sensor	BM area: 35 mm × 0.35 mm (single hair cell: up to 10 μm diameter and length 10-80 μm)	350 μm × 150 μm, 775 μm × 1500 μm (single beam)	∼20 μm × 300-7700 μm (single beam)	∼4 mm×4 mm (single membrane)	25 mm×8 mm	∼4 mm×3 mm (MEMS microphone)	∼4 mm×3 mm (MEMS microphone)
frequency decomposition	mainly resonance of membrane with varying properties and fluid interaction (BM), 3 dB-bandwidth (neuronal tuning curve): 500 Hz - 5 kHz, $Q_{10}: 3-10$	resonance of differently sized beams, 3 dB-bandwidth: 20 Hz-500 Hz (Q: 20-200)	resonance of differently sized beams, Q: $0.7-6$	resonance of differently sized membranes, Q: 30	resonance of membrane with varying properties (width/thickness), similar to BM, Q≈ 1	16-83 band-pass filters [9, 10, 21]	low-pass filter in spike domain [16], up to 108 bandpass filters [14, 15]
gain change, Q change	40-60 dB	44 dB, Q change: 40 to 200 (only amplification)	n.a.	5-7dB	amplitude change: factor 3-4; max. Q change: 1 to 1.63	20 dB [9, 10, 21], 17 dB [8], max. Q change: 1 to 40 [8]	n.a.
output dynamic range	40 dB (0.1-10 nm deflection)	in combination with feedback strength similar to input dynamic range	n.a.	n.a.	n.a.	46 dB [9, 10], 57 dB [21], 36 dB [7]	n.a.

References

[1] Zawawi, S. A., Hamzah, A. A., Majlis, B. Y. & Mohd-Yasin, F. A review of MEMS capacitive microphones. *Micromachines* **11**, 484 (2020).

[2] Doedel, E. J. & Oldeman, B. E. *Auto-07P: Continuation and bifurcation software for ordinary differential equations.* Concordia University, Montreal, Canada (2009).

[3] Koickal, T. J. *et al.* Design of a spike event coded rgt microphone for neuromorphic auditory systems. In *2011 IEEE International Symposium of Circuits and Systems (ISCAS)*, 2465–2468 (IEEE, 2011).

[4] Guerreiro, J. *et al.* Enhancing Acoustic Sensory Responsiveness by Exploiting Bio-inspired Feedback Computation. In *ICASSP* (2019).

[5] Tsuji, T., Nakayama, A., Yamazaki, H. & Kawano, S. Artificial cochlear sensory epithelium with functions of outer hair cells mimicked using feedback electrical stimuli. *Micromachines* **9**, 273 (2018).

[6] Yamazaki, H., Yamanaka, D. & Kawano, S. A preliminary prototype high-speed feedback control of an artificial cochlear sensory epithelium mimicking function of outer hair cells. *Micromachines* **11**, 644 (2020).

[7] Liu, S.-C., van Schaik, A., Minch, B. A. & Delbruck, T. Asynchronous binaural spatial audition sensor with $2\times64\times4$ channel output. *IEEE Trans. Biomed. Circuits Syst.* **8**, 453–464 (2013).

[8] Yang, M., Chien, C.-H., Delbruck, T. & Liu, S.-C. A 0.5V 55μW 64×2-channel binaural silicon cochlea for event-driven stereo-audio sensing. In *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 388–389 (2016).

[9] Hamilton, T. J., Tapson, J., Jin, C. & Van Schaik, A. Analogue vlsi implementations of two dimensional, nonlinear, active cochlea models. In *2008 IEEE Biomedical Circuits and Systems Conference*, 153–156 (IEEE, 2008).

[10] Hamilton, T. J., Tapson, J., Jin, C. & Van Schaik, A. Analogue VLSI implementations of two dimensional, nonlinear, active cochlea models. In *2008 IEEE Biomedical Circuits and Systems Conference*, 153–156 (IEEE, 2008).

[11] Sarpeshkar, R. *et al.* An ultra-low-power programmable analog bionic ear processor. *Biomedical Engineering, IEEE Transactions on* **52**, 711 – 727 (2005).

[12] Thakur, C. S., Hamilton, T. J., Tapson, J., van Schaik, A. & Lyon, R. F. FPGA implementation of the CAR model of the cochlea. In *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1853–1856 (2014).

[13] Thakur, C. S. *et al.* Sound stream segregation: a neuromorphic approach to solve the "cocktail party problem" in real-time. *Front. Neurosci.* **9**, 309 (2015).

[14] Xu, Y. *et al.* A FPGA implementation of the CAR-FAC cochlear model. *Front. Neurosci.* **12** (2018).

[15] Singh, R. K. *et al.* CAR-Lite: A multi-rate cochlear model on FPGA for spike-based sound encoding. *IEEE Trans. Circuits Syst. I: Regul. Pap.* **66**, 1805–1817 (2019).

[16] J-Fernandez, A. *et al.* A binaural neuromorphic auditory sensor for FPGA: A spike signal processing approach. *IEEE Trans. Neural Netw. Learn. Syst.* **28**, 804 – 818 (2017).

[17] Nouri, M. *et al.* A Hopf resonator for 2-D artificial cochlea: Piecewise linear model and digital implementation. *IEEE Trans. Circuits Syst. I: Regul. Pap.* **62**, 1117–1125 (2015).

[18] Robles, L. & Ruggero, M. A. Mechanics of the mammalian cochlea. *Physiol. Rev.* **81**, 1305–1352 (2001).

[19] Kandel, E. R., Schwartz, J. H. & Jessell, T. M. (eds.) *Principles of Neural Science* (Elsevier, New York, 2000), fourth edn.

[20] Latif, R., Majlis, B. & Cheung, R. Mems design and modelling based on resonant gate transistor for cochlear biomimetical application. *Microsystem Technologies* **23**, 2329–2342 (2017).

[21] Sarpeshkar, R. Universal principles for ultra low power and energy efficient design. *IEEE Trans. Circuits Syst. II: Express Briefs* **59**, 193–198 (2012).