Neuromorphic acoustic sensing using an adaptive microelectromechanical cochlea with integrated feedback

Claudia Lenk [1] ✉, Philipp Hövel [2,3], Kalpan Ved[1], Steve Durstewitz[1], Thomas Meurer[3,4], Tobias Fritsch[5], Andreas Männchen[5], Jan Küller[5], Daniel Beer[5], Tzvetan Ivanov[1] & Martin Ziegler [1,6]

Many speech processing systems struggle in conditions with low signal-to-noise ratios and in changing acoustic environments. Adaptation at the transduction level with integrated signal processing could help to address this; in human hearing, transduction and signal processing are integrated and can be adaptively tuned for noisy conditions. Here we report a microelectromechanical cochlea as a bio-inspired acoustic sensor with integrated signal processing functionality. Real-time feedback is used to tune the sensing and processing properties, and dynamic switching between linear and nonlinear characteristics improves the detection of signals in noisy conditions, increases the sensor dynamic range and enables adaptation to changing acoustic environments. The transition to nonlinear behaviour is attributed to a Hopf bifurcation and we experimentally validate its dependence on sensor and feedback parameters. We also show that output-signal coupling between two coupled sensors can increase the frequency coverage.

Our hearing exhibits remarkable sensing properties: a dynamic range of sound pressure level (SPL) of 120 dB, a frequency resolution of 0.1%, an intensity discrimination of 1 dB and adaptive capabilities (particularly at small SPLs and in noisy environments; these environments are known as the 'cocktail party effect')[1]. This functionality is due to three properties of the system: pre-processing, local frequency-selective feedback and dynamic adaptation. Pre-processing includes frequency filtering and frequency-selective nonlinear amplification of the signal before it reaches the auditory nerve, and encoding of the signal into spike trains at the auditory nerve[2,3]. Local frequency-selective feedback of the sensor changes the sensor gain by 40–60 dB (ref. 4), based mainly on changes to the outer hair cell motility[2,5,6]; feedback enables the detection of sounds below the thermal noise level and in the presence of noise or masking sounds[7–9]. Dynamical adaptation

occurs at multiple stages of the auditory pathway, including signal processing before transduction (that is, middle ear transfer function by acoustic reflex), during transduction (that is, inner ear processes) and at subsequent processing stages. This provides improved sensing conditions for varying hearing environments[7,10–12].

It remains challenging to recreate the features of biological hearing with technology. Learning-based sound processing systems such as convolutional, recurrent and spiking neural networks have been developed for tasks such as keyword spotting, speaker identification and speech analysis[13–16]. It has been shown that incorporating bio-inspired pre-processing increases the performance considerably[17]. Processing of a microphone signal, such as nonlinear filtering, frequency decomposition and feature extraction, occurs at the audio front end before feeding it into a neural network at the back end (Fig. 1, left).

[1]Department of Microelectronic and Nanoelectronic Systems, Technische Universität Ilmenau, Ilmenau, Germany. [2]School of Mathematical Sciences, University College Cork, Cork, Ireland. [3]Faculty of Engineering, Kiel University, Kiel, Germany. [4]Department of Chemical and Process Engineering, Karlsruhe Institute of Technology, Karlsruhe, Germany. [5]Fraunhofer Institute for Digital Media Technology Ilmenau, Ilmenau, Germany. [6]Institute of Micro- and Nanotechnology, Technische Universität Ilmenau, Ilmenau, Germany. ✉e-mail: claudia.lenk@tu-ilmenau.de
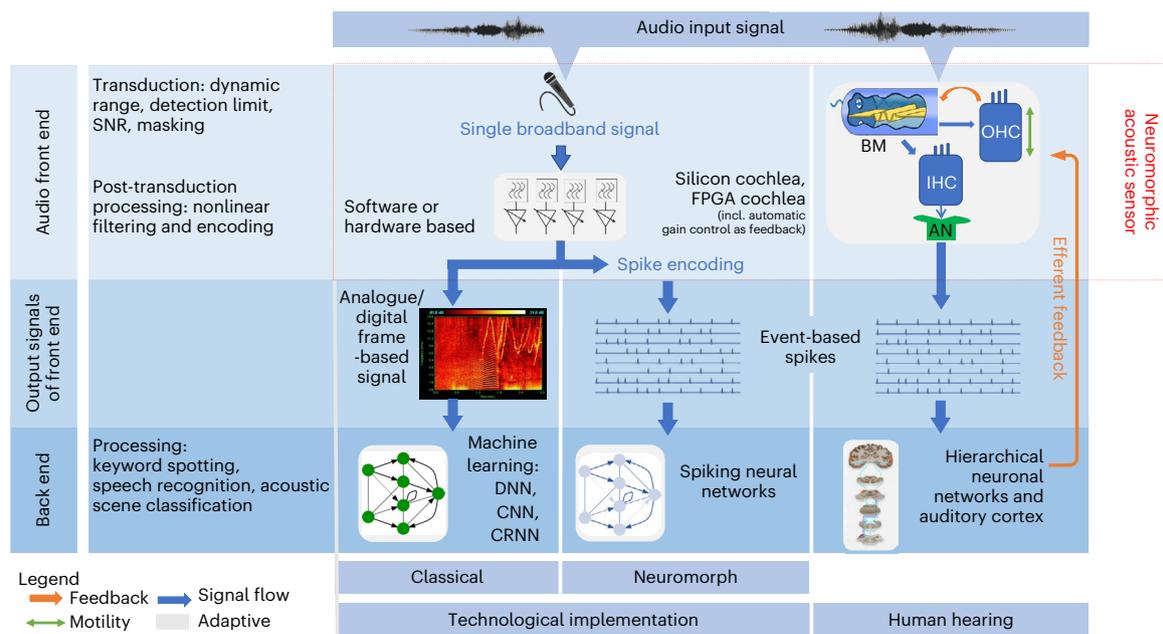
Fig. 1 | Concepts of speech processing: technological versus biological audio front ends and speech processing units. The bright grey boxes indicate the adaptive parts. The orange arrows represent feedback at the same level or from higher levels to change sensing and/or processing properties. The red bracket indicates the target levels and properties for the design of neuromorphic acoustic sensors. BM, basilar membrane; OHC, outer hair cell; IHC, inner hair cell; AN, auditory nerve; DNN, deep neural network; CNN, convolutional neural network; CRNN, convolutional, recurrent neural network. Credit: BM image, Wikimedia Commons under a Creative Commons license CC BY 2.5; auditory cortex, Wikimedia Commons under a Creative Commons license CC BY 4.0.

However, room reverberation, interfering noise and other perturbations to the signal that can affect the underlying feature representation can limit performance, particularly at low signal-to-noise ratios (SNRs)[18–20]. In addition, it is difficult to separate individual sound sources from a mixed acoustic signal and to generalize to unknown acoustic conditions. Automatic adaptation that can overcome some of these issues is currently being implemented at the signal processing level (that is, nonlinear filtering) or at the network stage[21–23]. Although basic tasks like voice activity detection, keyword spotting and speech detection with a limited vocabulary have been implemented on low-power devices[24,25], this has yet to be achieved for more complex applications such as acoustic scene analysis.

Neuromorphic sound sensors—such as silicon and field-programmable gate array (FPGA) cochleae[23,26–35]—already incorporate some adaptive signal pre-processing in the form of frequency decomposition and nonlinear frequency-selective amplification. However, these sensors rely on standard microphones that have limited pre-processing and adaptation capabilities for transduction. Bio-inspired acoustic sensors have been developed that can implement frequency decomposition (of up to 12 channels), nonlinear amplification (change in gain of up to 7 dB) and directionality[36–38], but these only cover a small frequency range, have low frequency resolution (quality (Q) factor of around 1) and do not include adaptability. Artificial cochleae that can adapt to their acoustic environment can greatly improve the performance and efficiency of processing. This has been achieved by damping the neighbouring frequency bands[39,40] and incorporating a leaky-integrate-and-fire model for feedback[41]. Such an approach is expected to outperform state-of-the-art signal processing capabilities in terms of detecting/processing large SPLs, latency, energy efficiency and reduction in masking of quiet sounds in noisy environments[42].

In this Article, we report an adaptive neuromorphic microelectromechanical system (MEMS)-based cochlea with integrated signal processing. The acoustic sensor system consists of one or more acoustic transducers—a silicon beam integrated with a thermo-mechanical actuator and four piezo-resistive elements for deflection sensing—and a feedback system that connects the sensor elements to the actuator. This feedback is used to tune the sensing and processing properties of the system in real time based on the acoustic signal properties (Fig. 1, right). The feedback and coupling properties can be controlled to suit different acoustic environments. For example, the signal amplitude can be used to implement an amplitude-dependent shift of the dynamic range[43] to increase the sensitivity of measurements at low noise levels.

In the case of the MEMS cochlea with a single transducer, the self-feedback strength is used to switch the transfer characteristics—the voltage as a function of sound amplitude—among linear, nonlinear, mixed or amplitude-independent regimes (for example, in the nonlinear regime, the sensitivity is high at lower sound pressures and decreases at higher pressures, which improves signal detection in noisy conditions). Our dynamic MEMS cochlea exhibits a gain change of up to 44 dB, comparable with mammalian hearing. Stability analysis of the nonlinear response indicates that a Hopf-type bifurcation occurs in the system. In addition to the operation based on the self-feedback of a single transducer, we show that two transducers in the MEMS cochlea can be coupled using feedback based on the output signal of the other. This can be used to adjust the sensitivity and frequency coverage.

Sensing properties

The sensor system proposed here consists of an acoustic transducer and a feedback loop (Fig. 2a,b). The transducer is realized as a silicon beam with integrated piezo-resistive deflection sensing and integrated thermo-mechanical actuation (Methods). Such transducers have already been successfully applied in atomic force microscopy and scanning probe lithography[44], as well as other sensing tasks like gas flow sensing and gas mixture analysis[45,46]. The integration of actuation and sensing into the beam has the advantage of being able to implement real-time feedback loops. Here the feedback loop is either self-feedback for a single transducer (Fig. 2) or an output-signal coupling of two transducers (Fig. 5), with feedback calculation time in the range of less than 10 μs.
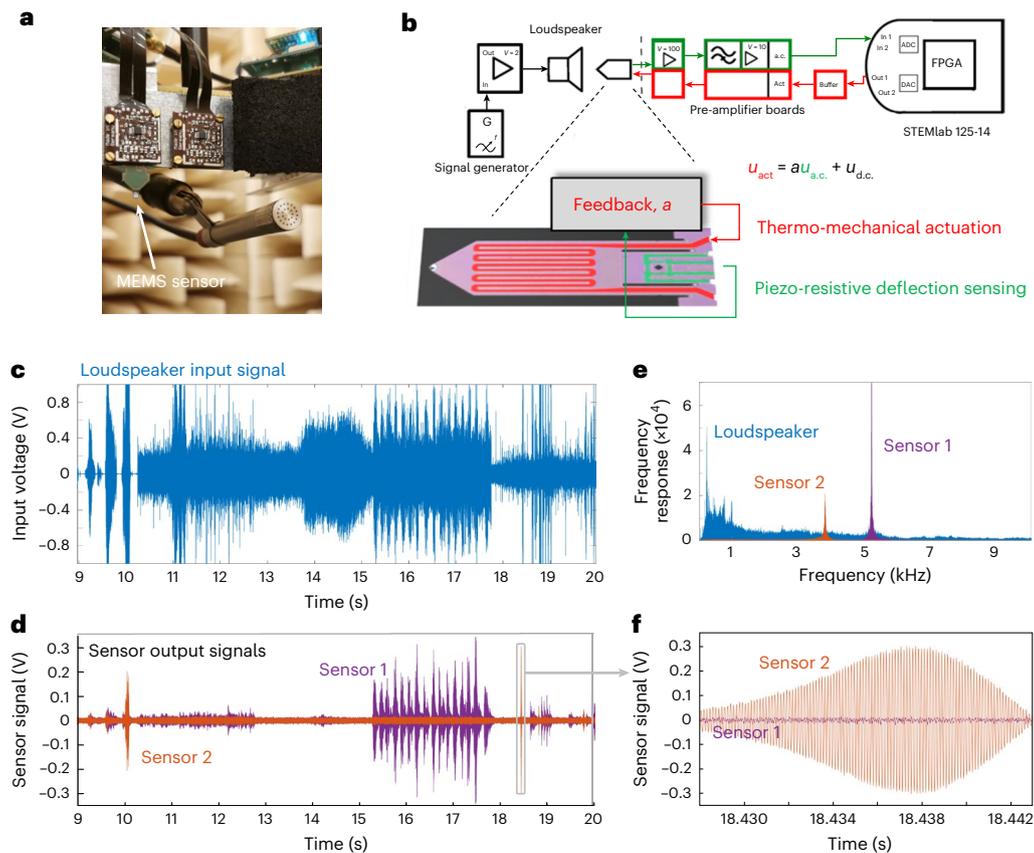
Fig. 2 | System overview and frequency filtering. a,b, Photograph (**a**) and schematic (**b**) of the sensor system with self-feedback. The inset shows a coloured microscopy image of the transducer. ADC, analogue-to-digital converter; DAC, digital-to-analogue converter. **c,** Time series of the sound input given by natural sounds (from the natural sound dataset 1 in ref. 47), used to study the sensor response to complex inputs. **d,** Sensor response (in the active, nonlinear mode) to sound input (shown in **c**) obtained from the measurements in an anechoic chamber with two sensors of different resonance frequencies: 5.19 kHz for sensor 1 (purple line), 3.73 kHz for sensor 2 (red line). **e,** Frequency responses of sensor 1, sensor 2 and the exciter signal for natural sound input (as shown in **c**). The filtering effect is clearly visible. **f,** An enlarged section of the peak in **d**, to demonstrate the sine-wave response of the sensor despite the complex sound input. Note that the measurements in **c–f** are performed with a sensor having the design shown in Supplementary Fig. 2.

The responses of two sensors with different resonance frequencies (f_1 = 5.19 kHz; f_2 = 3.73 kHz) to a complex sound signal (consisting of natural sounds like ferret calls, speech and running water, from a natural sound database[47]) are shown in Fig. 2c–f. The frequency filtering effect is visible from different onsets of response, as well as the frequency spectra. Due to these strong filtering properties, the dynamics are not affected by extrinsic noise down to a very small SNR (Supplementary Section 1). Furthermore, as shown by the frequency response (Fig. 2e) and the enlargement of the time series (Fig. 2f), harmonic oscillations are observed only at the first mode, despite the complex input.

Self-feedback strength a provides a route for tuning the transfer characteristics, that is, sensing voltage (proportional to deflection of the beam) as a function of the amplitude of sound pressure, for single-tone excitation (Fig. 3). Besides the passive mode ($a = 0$), four different types of response characteristic can be observed in the active, amplification mode ($a > 0$; Fig. 3a): a linear response for $a < 0.50$, a nonlinear response for $0.70 < a < 0.74$, a mixture of linear and nonlinear response for $0.50 < a < 0.70$ and a sound-amplitude-independent response for $a > 0.74$. In the following, these regimes are discussed in more detail.

In the linear regime, the sensitivity increases and lower sound pressures can be detected, if the feedback strength is increased (Fig. 3a). At the same time, the equivalent noise level (self-noise) is reduced by 3 dB SPL due to the active operation. The relative gain for the active operation mode compared with the passive mode without feedback can be increased by a factor of 4–5, where the highest gain is observed at the highest sound pressures (Fig. 3c).

In the nonlinear regime (Fig. 3b), in contrast, the highest change in gain (around 9) is observed for the lowest sound pressure (0.05 Pa) and the lowest gain (around 3) for the highest sound pressure (0.43 Pa). Thus, the sensor becomes more sensitive to lower sound pressures than larger sound pressures. This effect resembles compressive amplification, which is observed in the human hearing system in the perception of loudness, that is, at the processing stage[48], and at the transduction stage, that is, the hair cells in the inner ear[5,49,50]. Furthermore, this effect is applied in many acoustic sensing systems as post-transduction processing by using nonlinear amplification (Fig. 1). Compressive amplification yields an amplitude-dependent resolution/sensitivity[48], and is observed for most of the biological senses such as vision and touch.

The change in gain could be further increased by optimizing the design of the transducer for acoustic sensing (Supplementary Section 2). This design shows a change in gain by a factor of 10 for the active linear mode (compared with the passive mode) and a factor of 16 for the active nonlinear mode. Furthermore, the self-noise, that is, the lowest detectable SPL (at resonance), was reduced to 26–28 dB SPL in the passive mode, comparable with standard MEMS microphones[51], and can be further reduced to 18–20 dB SPL in the active nonlinear mode, which is almost at the level of higher-quality measurement microphones (≈15–16 dB SPL).

In both linear and nonlinear regimes, the sensing of single tones is possible even in the presence of band-limited white noise down to SNRs below 0 dB (Supplementary Fig. 1). Here the SNR of the sensing
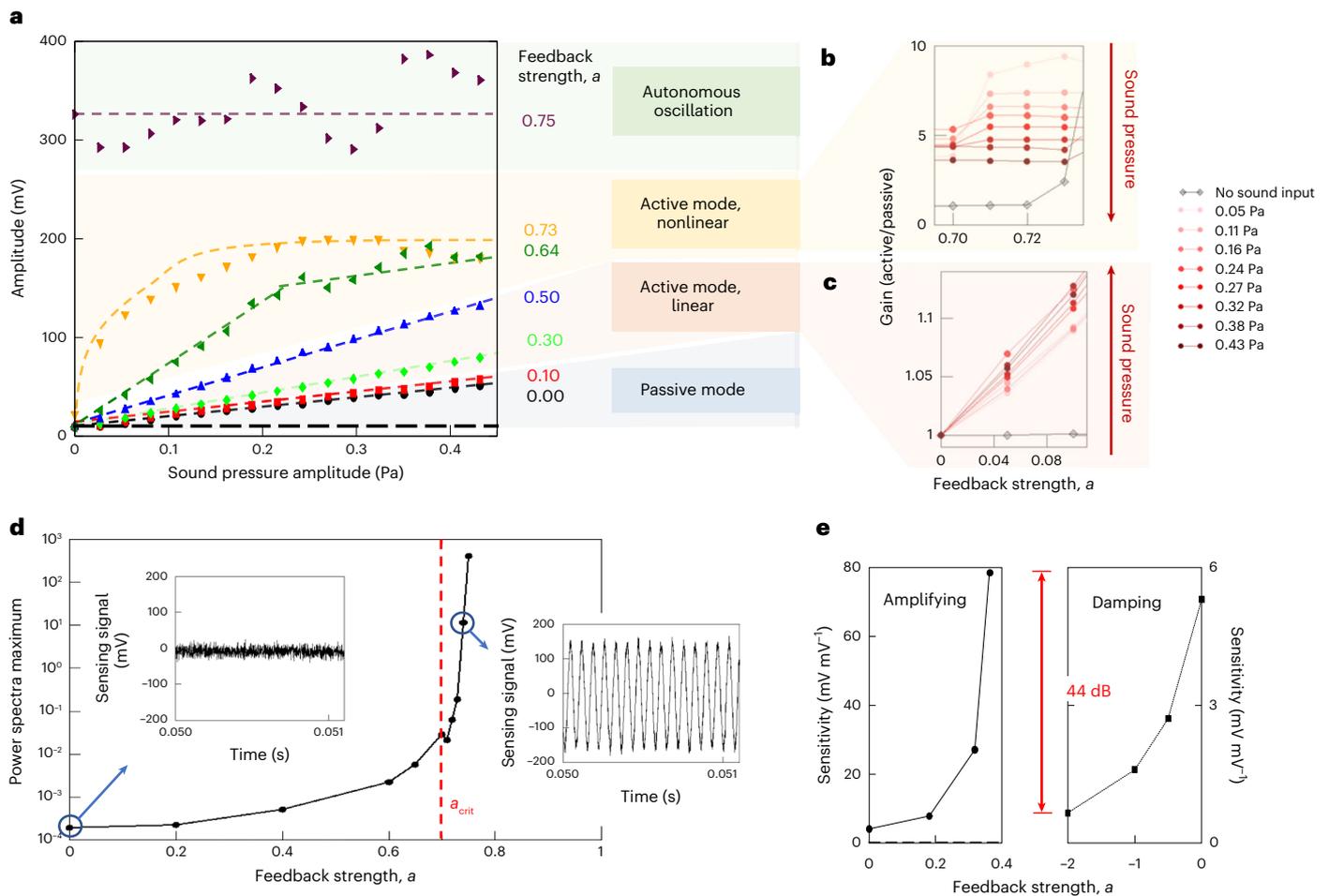
Fig. 3 | Sensing properties as a function of feedback. a, Amplitude of the sensor signal versus sound pressure amplitude for different values of feedback strength a to study the transfer characteristics of the sensor system ($u_{d.c.} = -200$ mV). Measurements were performed using a transducer with a resonance frequency of 14.2 kHz (Table 1 and Methods list the other properties) and chirped sound signals (12–16 kHz). Depending on a, the sensing behaviour in the active mode ($a > 0$) can be divided into an active, linear mode for $a < 0.50$; an active, nonlinear mode for $0.70 < a < 0.74$; a mixture between the linear and nonlinear mode for $0.50 < a < 0.70$; and sound-amplitude-independent, autonomous oscillations for $a > 0.74$. The intrinsic noise level due to electronics and so on is given by the dashed black line. **b,c**, Gain as a ratio of the active-mode amplitude to passive-mode amplitude for various sound pressure amplitudes in the two modes:

active, linear mode ($a < 0.50$) (**b**); active, nonlinear mode ($0.70 < a < 0.75$) (**c**). Compressive amplification, yielding a higher gain for lower sound pressure amplitudes, is observed for the active, nonlinear mode. **d**, Power spectra maximum depending on positive feedback strength a without applied sound. Autonomous oscillations without a sound input occur for $a > 0.74$. **e**, Sensitivity of bio-inspired sensors, given by the slope of the ratio of sensing voltage to driving voltage of the loudspeaker, as a function of feedback strength a. A positive feedback strength strongly increases the sensitivity near the bifurcation point, namely, $a \approx 0.5$ (left), whereas a negative feedback strength reduces the sensitivity (right). Using positive and negative feedback strengths together, the sensitivity (or gain) of the sensor can be varied by 44 dB, which is close to the 40–60 dB change in gain in the human cochlea due to the outer hair cell operation.

signal is constant for a large range of SNRs of sound signals (≈25 dB) and the SNR of the sensing signal can be improved by the active mode.

For much larger feedback strengths ($a > 0.74$), the sensing amplitude is almost independent of the SPL, and the sensor oscillates even without applying any sound (Fig. 3a,d). This behaviour is typical of nonlinear systems at a Hopf bifurcation[52].

Introducing a negative feedback strength results in damping of the acoustic response (Fig. 3e). Combining the amplification and damping regime, the sensor offers a change in gain of up to 44 dB, which is comparable with the added gain of 40–60 dB by outer hair cell activity in the mammalian cochlea[4].

Modelling of sensing properties

To understand the nonlinear response of the acoustic sensor and find out whether the observed autonomous oscillation is caused by a Hopf bifurcation, we analysed the dynamics of the sensing system. The mathematical description is based on another model[53]. The derived model

describes the change in deflection x of the free end of the beam due to thermo-mechanical actuation $\alpha\theta$ and external forcing ($\tilde{F}_{ext}$), (such as sound), by a damped oscillator equation derived using Euler–Bernoulli beam theory:

$$\ddot{x}(t) + \frac{\omega_0}{Q_0}\dot{x}(t) + \omega_0^2 x(t) = \alpha\theta(t) + \tilde{F}_{ext}(t), \quad (1)$$

where ω_0 is the resonance frequency, Q_0 is the quality factor and θ is the change in beam temperature, which is caused by the applied actuation voltage u_{act}:

$$\dot{\theta}(t) + \beta\theta(t) = \gamma\left(\frac{\tanh u_{act}(t)}{R}\right)^2 \quad (2)$$

obtained from the feedback loop

$$u_{act}(t) = a u_{a.c.}(t) + u_{d.c.}, \quad (3)$$

Table 1 | Parameters of the sensor system used for modelling

Variable	Value
Width of beam, w_{Si}	150 μm
Length of beam, l_{Si}	350 μm
Thickness of beam, d_{Si}	1.25–1.45 μm
Frequency of beam, f	[14.2, 13.92, 12.75] kHz
Density of silicon, ρ_{Si}	2,329 kg m^{-3}
Elastic modulus of silicon, E_{Si}	170 GPa
Heater resistance, R	[25.0, 20.5, 20.5] Ω
Calibration factor, k (piezo-resistive sensing and pre-amplification)	[1.00, 0.80, 0.25]×10^6 V m^{-1}
Time constant of high-pass filter, τ	10^{-3} s
Transfer factor, α (temperature to deflection)	749.3702 m^2 s^{-2} K^{-1}
Time constant, β (temperature change)	1.0066×10^3 s^{-1}
Transfer factor, γ (voltage to temperature)	4.2588×10^7 K Ω^2 s^{-1} V^{-2}
Mode factor δ_1 for frequency calculation	1.8751
Density of air, ρ_{gas}	1.189 kg m^{-3}
Dynamic viscosity of air, η_{gas}	18.232×10^{-6} Pa s

Device-specific parameters for sensors 1, 2 and 3 are provided in the brackets.

where R is the heater resistance. Here $u_{a.c.}$ is obtained from the transformation of deflection x into a sensing voltage u_s ($u_s = kx$) by the piezo-resistive elements, including high-pass filtering and amplification:

$$\dot{u}_{a.c.}(t) = -\frac{u_{a.c.}(t)}{\tau} + \dot{u}_s(t). \qquad (4)$$

The feedback introduces a nonlinearity into the system. This model includes various sensor properties such as the resonance frequency ω_0, heater resistance R and quality factor Q_0, and it can be easily adjusted to other beam dimensions and frequency ranges. Parameter values of the analysed system are given in Table 1 and Methods.

To determine the origin of the nonlinear response and autonomous oscillations, we studied the stability of the fixed points. This revealed a Hopf bifurcation depending on the feedback parameters, that is, feedback strength a and bias voltage $u_{d.c.}$ (Supplementary Section 3). For feedback strengths below the critical value a_{crit} at which bifurcation occurs, the system is quiescent in the absence of sound, whereas for $a > a_{crit}$, self-excited, autonomous oscillations occur together with a strong increase in amplitude (Fig. 3d, insets).

From this stability analysis and comparison with the normal form of Hopf-type oscillators, we derived an analytical equation for the critical feedback strength a_{crit}, enabling us to obtain the nonlinear regime in the dependence of sensor properties and feedback parameter $u_{d.c.}$ (Methods). The comparison of the derived equation with experimental data (black dots) shows excellent, quantitative agreement between theory and experiment (Fig. 4a). It is noteworthy that the critical feedback strength stays finite even for higher frequencies. Thus, the nonlinear regime should occur not only in the audible frequency range but also for ultrasound. Indeed, in the experiments with sensors having different resonance frequencies between 2 and 96 kHz, all the sensors exhibited autonomous oscillations as an indication of Hopf bifurcation.

To study the origin of the increase in sensitivity with increasing feedback strength in the linear regime, we compared the sound response in the experiment with the response to external forces in the model (Fig. 4b). We found that the sensitivity increase in the linear regime originates from an effective change in the quality factor with increasing feedback strength, similar to Q control. The slope of the effective quality factor strongly depends on the initial quality factor Q_0, which is determined by the geometric dimensions of the beam (Methods). With an increase in the initial Q_0, the slope of the effective quality factor strongly increases due to the influence of feedback. This enables us to control the sensitivity of the sensor by the choice of sensor design (setting Q_0) and feedback strength a.

Two coupled sensors

In addition to the discussed nonlinear operation, the human auditory system is argued to be further enhanced by the coupling of sensory elements (hair cells)[12,54,55]. This can improve the sensing performance by reducing damping due to cochlear fluid, increasing sensitivity and response amplitude, enhancing the reliability of sound encoding and stabilizing the operation mode by increasing the range of non-linear operation[12,54–56]. If we introduce output-signal coupling of two transducers as feedback (Fig. 5b) in our sensing system (instead of the self-feedback described above), we obtain increasing sensitivity with increasing coupling strength, a switching from linear to nonlinear sensing characteristics, and self-excited, autonomous oscillations indicating a Hopf bifurcation. The latter was observed even if the resonance frequencies of the two sensors were more than 10 kHz apart.

Another effect of coupling feedback is shown by the power spectra of both sensors (Fig. 5a). If the sensors are uncoupled (coupling strength $b_{12} = b_{21} = b = 0$), each sensor responds to sound at its resonance frequency (black curves). If the beams are mutually coupled, for example, increasing the coupling strength to $b = 1.05$, an increased response of the respective sensor at its own resonance frequency is observed, and each sensor exhibits a slight response at the resonance frequency of the other sensor (red curves). For even higher coupling strengths ($b = 1.875$), a substantial response of the sensors occurs even in the frequency range between both resonance frequencies. This effect strongly increases the bandwidth of the sensor system: initially from 500 Hz up to approximately 5 kHz (Fig. 5a). A further increase in the coupling strength results in self-excited oscillations of the sensor system.

Hence, the output-signal coupling can modify the sensitivity of each sensor and its transfer characteristics, similar to the self-feedback and coupling effects in the hearing system, and it can also modify the bandwidth of the coupled system consisting of both sensors. This effect helps to reduce the number of sensors needed to cover a certain frequency range, since the sensors do not only respond at their resonance frequency (with a typical bandwidth of 20–500 Hz, depending on the design) but also in the frequency range between the resonance frequencies of the coupled sensors.

Dynamical adaptation

Biological senses, like vision, hearing and touch, are focused on detecting the relative values and changes rather than absolute values[3]. Therefore, adaptation is not only used to tune the sensing properties like sensitivity, resolution and operation point of the system in a slowly changing environment but also to highlight fast changes in stimuli such as, for example, the onset of a stimulus[3,11,12,57]. These fast adaptation mechanisms support processing tasks like sound source localization, where performance is strongly dependent on exact onset detection[58–62]. Furthermore, the adaptation can increase the efficiency of the system and reduce the redundancy of information for processing, for example, by reducing the spike rate for constant stimuli (known as sensory adaptation). Onset/offset detection can help to reduce the power consumption and data streaming needs by reducing the feedback signal after the detection of the onset of constant sounds and triggering the start/end of data streaming to processing units. In this way, data will be transferred for further analysis only when sound occurs in a specific frequency band set by the sensor.

In our sensing system, dynamic adaptation is implemented by the self-guided adjustment of the feedback parameters: feedback strength a, bias voltage $u_{d.c.}$ and coupling strength b. The feedback
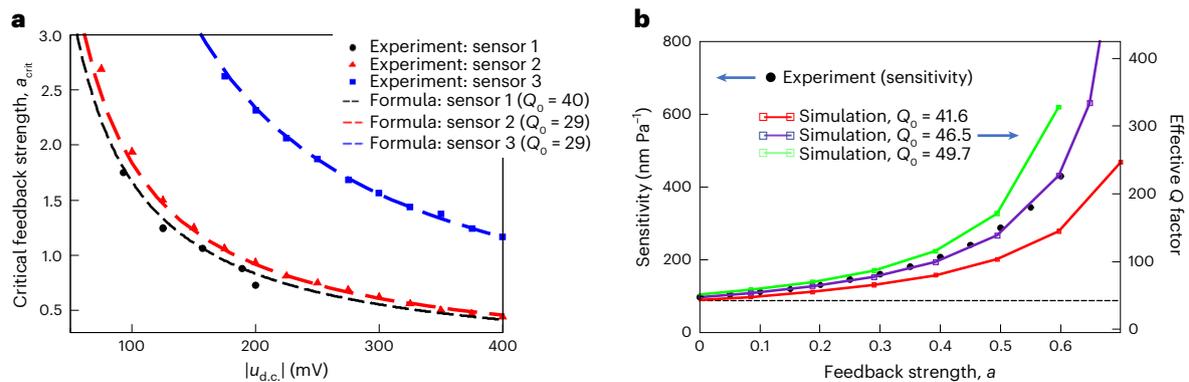
Fig. 4 | Comparison of experiments and model. a, Critical feedback strength a_{crit}, for the occurrence of a Hopf bifurcation without sound input (autonomous oscillations), depending on bias voltage $u_{d.c.}$ obtained from experiments (black, red and blue dots) with three different sensors and from a fixed-point analysis (equation (10)) with experimental parameters f, R, Q and k (Methods). The formula describes the dependence of the critical feedback strength on the sensor properties and feedback parameter $u_{d.c.}$. **b**, Sensitivity of the experimental sensor

system ($u_{d.c.} = -200$ mV) depending on feedback strength a for single-tone sound signals with a frequency of 14 kHz (pressure range, 0.10–0.32 Pa) as black dots. The effective quality factor depending on feedback strength a obtained from the simulations of equations (1)–(5) for different initial quality factors: $Q_0 = 41.6$ (red line), $Q_0 = 46.5$ (purple line) and $Q_0 = 49.7$ (green line). The comparison shows that an increase in sensitivity in the active, linear mode with increasing feedback strength a can be attributed to the change in quality factor due to feedback.
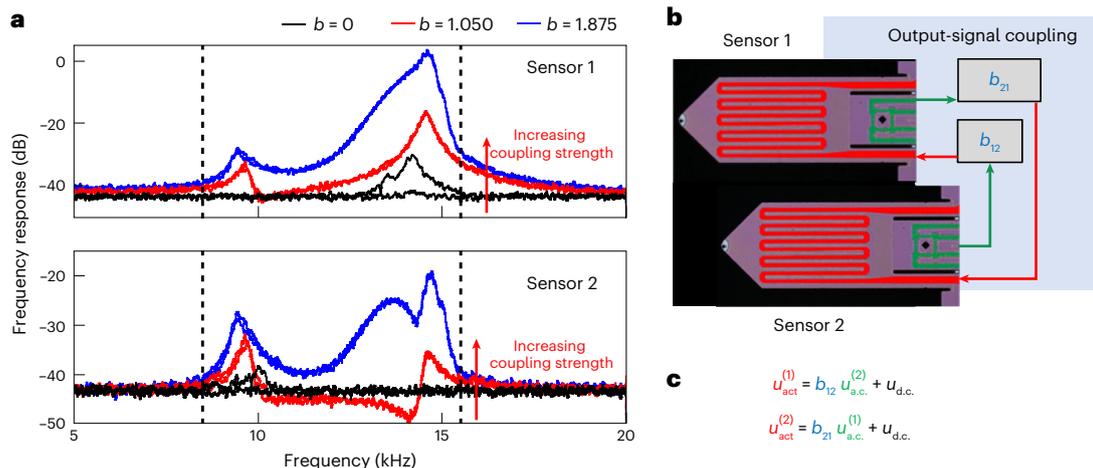
Fig. 5 | Two output-coupled sensors. a, Frequency response ($f_1 = 14.2$ kHz, $f_2 = 10.1$ kHz) shown for three different coupling strengths $b_{12} = b_{21} = b = 0$ (black), $b = 1.050$ (red) and $b = 1.875$ (blue). The sound signal consists of a linear frequency chirp in the range of 8.5–15.5 kHz with a sound pressure amplitude of 0.107 Pa. Increasing the coupling strength yields a stronger response of the sensors at their

respective resonance frequency as well as an increase in sensor bandwidth, that is, the range of sound frequencies, to which the sensors respond (the blue curve between 10 and 14 kHz). **b**, Schematic of the sensor system with output-signal coupling. **c**, Equations of actuation signals for output-signal-coupled sensors (Methods).

strength controls the linearity of transfer characteristics (amplification behaviour), sensitivity and filtering properties by changing the quality factor of the system. The bias voltage shifts the critical feedback strength for the nonlinear regime. The coupling strength changes the sensitivity and bandwidth of the system. Since all the three parameters can be individually controlled, short-term and long-term adaptations targeting amplitude and frequency ranges can be easily implemented. This enables the combination of, for instance, a fast adaptation of the sensor to the onset of sound signals (similar to sensory adaptation[3]) or automatic gain control to avoid damage due to high SPLs with slow adaptation, similar to homoeostatic control keeping the sensing amplitude in a pre-defined range[11,22]. Such adaptations can be used to increase the dynamic range, implement event-based sensing and spike-rate-based encoding of sound properties, as well as cover large frequency ranges with only a few transducers and still retaining high-frequency resolution.

We implemented a dynamical adaptation in our sensor system using a fast adaptation of feedback strength a depending on the sensing amplitude (Fig. 6a; switching time below 10 µs). Here a is switched from a_1 to a lower value a_0 if the amplitude crosses a pre-defined threshold V_{th}. It is reset to the initial value a_1 either if the amplitude decreases below a second threshold (to model sensory adaptation) or after a pre-defined time interval τ_2 (to model a refractory period).

Experimental implementation of the refractory period adaptation shows a spike-like output of the sensing system (Fig. 6b), which can be used to generate event-based spikes based on the acoustic input. The spiking frequency depends on the refractory period τ_2, as well as on the sound pressure amplitude. Increasing the sound pressure amplitude results in a reduction in rise time τ_1 of the sensor signal until reaching the threshold for switching the feedback strength, as evident from a comparison of the response to two different sound levels (Fig. 6b, red and blue curves). Thus, the sound amplitude is encoded as a spike rate
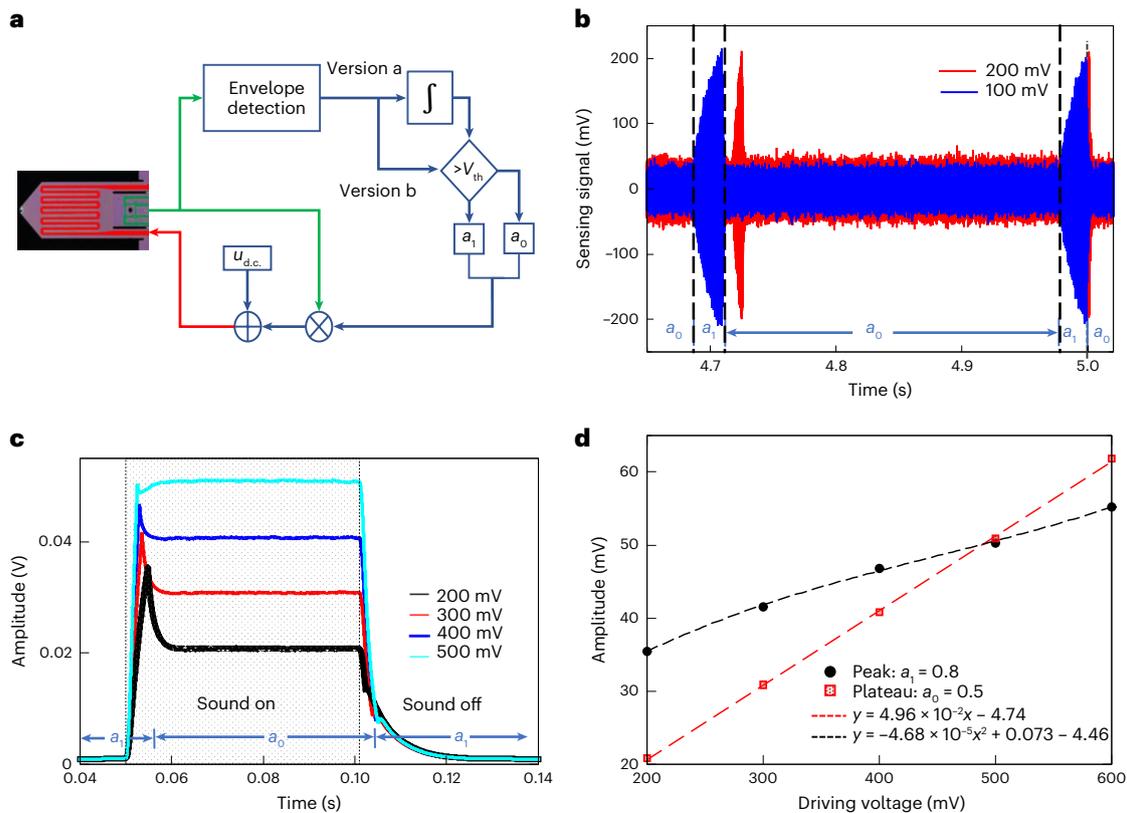
Fig. 6 | Dynamical adaptation of feedback strength. a, Schematic describing the dynamic adaptation algorithm used in **b** and **c**. **b**, Time series of sensor signals for two different sound amplitudes (blue and red) obtained from experiments with the FPGA-based implementation of feedback strength switching (schematic shown in **a**) between $a_1 = 0.7$ and $a_0 = 0$. Here the feedback strength is kept at its lower value for a constant time interval τ_2 before resetting it to the high-sensitivity regime. This yields a spike-like response of the sensor system to the constant sound input. The comparison for both sound amplitudes reveal an amplitude-dependent spike rate, which is determined by the sound-amplitude-dependent part τ_1 and the fixed time interval τ_2 for reset. **c**, Numerical implementation of sensory adaptation obtained from LTspice simulations of system with adaptation circuit (schematic shown in **a**). The envelope of the sensor signal with dynamic adaptation of feedback strength is shown in the

case of a constant sound input in the interval of 0.05 to 0.10 s. Switching events changing the feedback strength are marked with the blue dashed lines. The dynamic adaptation increases the resolution and dynamic range by enabling the sensing of low sound pressures before switching (nonlinear regime $a_1 = 0.8$), which are otherwise below the noise level, and the discrimination of large sound amplitudes after switching (linear regime $a_0 = 0.5$). The resolution decreases with increasing sound amplitude and large sound amplitudes can drive the sensor into saturation in the nonlinear regime (Fig. 3a). **d**, Peak (black dots) and plateau (red squares) amplitude of the time series obtained from dynamic adaptation simulations shown in **c**: a decreasing resolution with increasing sound amplitude is observed for the peak amplitudes due to the nonlinear-operation range ($a_1 = 0.8$). In contrast, for the linear-operation regime ($a_0 = 0.5$), the resolution remains constant. The dashed curves are fits as a guide to the eye.

of the sensing signal. Furthermore, the experimental implementation of the latter loop using an FPGA demonstrates the stability of the sensor even under stepwise changes in feedback strength a.

Simulations (Fig. 6c) and measurements of the sensory adaptation case, implemented using analogue circuits with discrete devices, show that the onset of sound is highlighted in the sensing signal (which is important, for instance, for localization tasks) and that the dynamic range of the sensor is increased. The latter is achieved by generally operating the sensor in the most sensitive regime (a_1 close to a_{crit}) to enable highly resolved detection of small SPLs. However, as shown in Fig. 6d (black), this yields decreasing resolutions for increasing SPLs up to saturation with sound pressure amplitude. Switching to lower sensitivities after the initial response yields a better discrimination for larger SPLs (Fig. 6d (red)). Furthermore, the switching signal can be used to trigger either a data streaming unit, sending the sensing signals to a processing system, or a processing unit. Thus, data streaming or sound processing is initiated only if the sound signals are detected, which reduces the power consumption and streaming requirements for tasks like machine supervision or systems like hearing aids.

Conclusions

We have reported a neuromorphic acoustic sensing system that consists of MEMS cochlea and integrated real-time feedback, either to itself or as output-signal coupling to a pair of sensors. The system shows high tunability and adaptive sensing properties, such as variable sensitivity or switching between linear and nonlinear transfer characteristics, as well as the integration of signal processing steps such as frequency filtering and nonlinear compressive amplification. We also showed that dynamical switching between linear and nonlinear characteristics improves the detection of signals in noisy conditions, increases the dynamic range of the sensor and enables adaptation to changing acoustic environments. Furthermore, output-signal coupling strongly increases the frequency coverage.

Our dynamic MEMS cochlea has several advantages over previously reported neuromorphic acoustic sensing systems (including bio-inspired acoustic sensors with integrated signal processing/adaptation[38–41]) and silicon and FPGA cochleae[23,26–35]. Its sensing properties—particularly, its gain change of up to 44 dB—are comparable with the mammalian cochlea, and the simplicity of the feedback algorithm enables fast and efficient feedback and adaptation mechanisms

with a small overhead per channel. Our sensor can also be fabricated based on standard complementary metal–oxide–semiconductor processes and shows high resilience against device tolerances and device mismatches, due to the large operation ranges for the feedback parameters (a comparison of the dynamic MEMS cochlea with mammalian cochlea and other neuromorphic acoustic sensing systems is given in Supplementary Section 4 and Supplementary Table 2).

The adaptive sensing of our system is of particular interest in noisy or multi-source situations. Due to the adaptive properties of the sensor, the sensitivity can be increased at low SNRs using the nonlinear-operation mode to improve detection or reduced at high SPLs using the linear mode to avoid saturation of the sensing signal. Since each sensor can be individually and dynamically tuned by the integrated amplification mechanism, it is possible to avoid the masking of certain frequency bands by larger SPLs in other bands, as can occur for microphone-based systems. Furthermore, because the input dynamic range is directly compressed at the sensor level, there are no constrictions of the dynamic range by subsequent electronics. Both these features are hard to achieve using standard microphone technology.

The bio-inspired merging of sensing and processing in the dynamic MEMS cochlea provides compact (in terms of circuit elements per channel) and robust (in terms of device mismatch and tolerances) systems with minimal signal processing latency due to the integration of signal processing into the sensing process. These properties make our system a potential alternative to conventional 'microphones plus subsequent signal processing' as the input stage for speech processing systems.

Methods

Experimental implementation

The acoustic sensor system (Fig. 2a,b) consists of two parts: the acoustic transducer and a feedback loop[63]. The transducer comprises a three-layer structure with a silicon layer as the base of the beam structure with 150 μm width, 350 μm length and thickness varying between 1 and 5 μm (fabrication details are given elsewhere[44]). The other two layers on top of the silicon are a silicon dioxide layer (thickness, ≈100 nm) for electrical isolation and an aluminium layer (thickness, ≈5 μm; Fig. 2b, red), which is used as an actuator for the beam. The size of both additional layers is negligible compared with the silicon base, which, thus, determines the resonance frequency and sensor properties (such as quality factor, Q_0). The aluminium layer on top of the beam is used as a thermo-mechanical actuator. Applying a voltage at the aluminium loop leads to a current through the actuator that introduces heating of the beam due to its resistance. Since the thermal expansion coefficients of silicon and aluminium differ, the temperature change yields a deflection of the beam, which is proportional to the introduced power. In addition to the integrated actuator, deflection sensing of the transducer is realized by four piezo-resistive elements (Fig. 2b, green) near the base of the beam. They are arranged in a Wheatstone bridge configuration to reduce the influence of noise. The deflection can be inferred as a voltage change, since a deflection of the beam results in a resistivity change in the piezo-resistive elements.

The second part of the sensor system is the feedback loop (Fig. 2b) that is used to tune the sensing properties by changing the dynamics of the transducer. The sensing voltage is amplified, high-pass filtered to neglect its d.c. part and converted into a digital signal by the analogue-to-digital converter of the STEMlab 125-14 board (sample rate 125 MHz and 14-bit resolution). The feedback signal is calculated in an FPGA structure on that board, too. Finally, the feedback signal is converted into an analogue signal by the digital-to-analogue converter of the STEMlab 125-14 board (sample rate 125 MHz; limitation ±1 V) and used to drive the actuator of the transducer.

Two types of feedback mechanism are applied: self-feedback, which uses the sensing voltage of a single transducer for feedback, and an output-signal coupling, which takes the sensing signal of one transducer to drive the actuator of a second transducer. The self-feedback signal u_{act} is given by

$$u_{act}(t) = a u_{a.c.}(t) + u_{d.c.} \tag{5}$$

with high-pass-filtered sensing voltage $u_{a.c.}$, the self-feedback strength $a \geq 0$ and bias voltage $u_{d.c.}$. In the case of output-signal coupling, the feedback signals $u_{act}^{(i)}, i = 1, 2$ for the two coupled transducers are given by

$$u_{act}^{(1)}(t) = b_{12} u_{a.c.}^{(2)}(t) + u_{d.c.}^{(1)}, \tag{6a}$$

$$u_{act}^{(2)}(t) = b_{21} u_{a.c.}^{(1)}(t) + u_{d.c.}^{(2)}, \tag{6b}$$

where $u_{a.c.}^{(1)}(t)$ and $u_{a.c.}^{(2)}(t)$ denote the high-pass-filtered sensing signals of sensors 1 and 2, respectively; coupling strength $b_{ij}, i, j = 1, 2$; and bias voltages $u_{d.c.}^{(i)}, i = 1, 2$. The coupling strengths and bias voltages can be different for the two sensors, but in the following, we take the same values for both.

The implementation of the feedback loop with the FPGA architecture of the STEMlab 125-14 board allows a near real-time feedback (≈0.1–1.0 μs delay, corresponding to maximum 1.4% of the oscillation period of the resonator). The sensor signal is saved into a file with a sample rate of 1.98 MHz for a subsequent analysis using MATLAB (versions 2019b and 2022b).

The acoustic sensing properties are tested using sound excitation with a piezo-loudspeaker (Kemo Electronic L010) driven by a signal generator (Agilent 33521A). Three types of acoustic signals are used: (1) single-tone studies using a sine-wave signal (for self-feedback and dynamical adaptation experiments); (2) chirp tones with a sine wave, whose frequency is linearly swept (for output-coupling experiments); and (3) a sum of a sine-wave signal with band-limited white noise (for self-feedback experiments). The driving voltage for the loudspeaker determines the SPL, where the sound pressure amplitude is linearly dependent on the driving voltage.

Theoretical description

For the theoretical description of the sensor system, we use a modified form of the modal description for the first mode derived earlier[53].

$$\ddot{x}(t) + \frac{\omega_0}{Q_0}\dot{x}(t) + \omega_0^2 x(t) = \alpha\theta(t) + \frac{F_{ext}(t)}{m_{eff}} \tag{7a}$$

$$\dot{\theta}(t) + \beta\theta(t) = \gamma\left(\frac{u_{act}(t)}{R}\right)^2 \tag{7b}$$

$$\dot{u}_{a.c.}(t) = -\frac{u_{a.c.}(t)}{\tau} + \dot{u}_s(t). \tag{7c}$$

Here $x(t)$ represents the deflection of the beam, $\theta(t)$ is the temperature difference between the beam structure and its surrounding, $u_{a.c.}(t)$ is the high-pass-filtered sensing signal and $u_{act}(t)$ is the actuation voltage. The latter is calculated according to equation (5) or equation (6), depending on which case is studied. To prevent damage to the transducer, the actuation voltage is limited to the range of ±0.5 V. In the analysed deflection range, the sensing voltage u_s is linearly related to the deflection: $u_s(t) = kx(t)$ with calibration factor k, which also includes the pre-amplification of the signal. The eigenfrequency of the transducer is given by $\omega_0 = 2\pi f$. Since the width and length are kept constant, the thickness of the transducer determines the eigenfrequency of the sensor according to

$$f = \frac{\omega_0}{2\pi} = \delta_n^2 \frac{d_{Si}}{2\pi l_{Si}^2}\sqrt{\frac{E_{Si}}{12\rho_{Si}}}. \tag{8}$$

Here l_{Si} is the length of the sensor, d_{Si} is the thickness of the sensor, E_{Si} is the elasticity module, ρ_{Si} is the density of Si and δ_n is a pre-factor for the nth mode.

Quality factor Q_0 of an oscillating beam in air was derived elsewhere[45] and is mainly determined by damping due to the surrounding fluid. It can be calculated according to

$$Q_0 = \frac{\frac{4\rho_{Si}d_{Si}}{\pi w_{Si}\rho_{gas}} + 1.05333 + \frac{3.7997}{\sqrt{2Re}}}{\frac{3.8019}{\sqrt{2Re}} + \frac{2.7364}{2Re}}, \tag{9}$$

where Reynolds number Re for this system is given by

$$Re = \frac{2\pi f \rho_{gas} w_{Si}^2}{4\eta_{gas}}.$$

Here w_{Si} and $2\pi f$ describe the width and oscillation frequency of the silicon beam, respectively. Also, ρ_{gas} and η_{gas} denote the density and dynamic viscosity of the surrounding media (air), respectively.

Parameters α, β and γ are sensor-specific parameters that describe the transformation of temperature into deflection, the time constant for temperature changes and the transfer efficiency from actuation voltage into temperature changes, respectively. The resistance of the actuator is given by R. External forcing can be introduced by the force term $F_{ext}(t)/m$ related to mass m of the transducer. Note that mass m used to relate the force to the deflection is not the total mass m_{Si} of the transducer but additionally includes a so-called added mass term $m_{movedgas}$, which arises from thermo-viscous damping[45]: $m = m_{Si} + m_{movedgas}$. The added mass $m_{movedgas}$ can be calculated using

$$m_{movedgas} = \frac{1}{4}\rho_{gas}\pi w_{Si}^2 l_{Si}\left(1.0553 + \frac{3.7997}{\sqrt{2Re}}\right).$$

Critical feedback strength

From equations (5) and (7), a linear stability analysis can be performed to study the origin of the nonlinear response of the sensor. This yields the critical feedback strength a_{crit} at the bifurcation point in the absence of an external force. Specifically, a linearization around the fixed point leads to a characteristic equation. The solutions of this characteristic equation are the eigenvalues of the fixed point. They are given depending on the feedback parameters a and $u_{d.c.}$ and sensor properties ω_0, Q_0, R, α, β and γ. We find one real-valued eigenvalue and a pair of complex-conjugate eigenvalues. The bifurcation occurs when the pair of complex-conjugate eigenvalues crosses the imaginary axis, that is, when their real parts become zero. This is the signature of a Hopf bifurcation. Indeed, we observe this dynamical behaviour as we vary the feedback strength a. Fixing all the other system parameters determines the critical value a_{crit} at this bifurcation:

$$\begin{aligned}
a_{crit} = \frac{-R^2}{4\gamma\alpha\tau^2 k u_{d.c.}} &\left[\left(\beta + \beta^2\tau + \frac{\omega_0}{Q_0}\left(1 + \beta\tau + \beta^2\tau^2\right)\right.\right.\\
&+ \frac{\omega_0^2}{Q_0}\left(\frac{1}{Q_0} - Q_0\right)(\tau + \beta\tau^2) + \frac{\omega_0^3}{Q_0}\tau^2\right) + \left(1 + \beta\tau + \frac{\tau\omega_0}{Q_0}\right)\\
&\left.\sqrt{\left(\frac{\omega_0}{Q_0} + \tau\omega_0^2\right)^2 + \left(\beta + \beta\tau\frac{\omega_0}{Q_0}\right)^2 + 2\beta\omega_0\left(-\tau\omega_0 + \frac{1}{Q_0} + \frac{\tau\omega_0}{Q_0^2} + \frac{\tau^2\omega_0^2}{Q_0}\right)}\right].
\end{aligned} \tag{10}$$

The critical feedback strength (Fig. 3a) depends on the second control parameter, that is, bias voltage $u_{d.c.}$, for different quality factors Q_0.

Data availability

The data that support the plots within this paper and other findings of this study are available via Zenodo at https://doi.org/10.5281/zenodo.7640418.

Code availability

The custom-developed codes for the MATLAB simulation and data analysis are available from the corresponding author upon reasonable request. The code for the LTspice simulations is available via Zenodo at https://doi.org/10.5281/zenodo.7640418.

References

1. Cherry, E. C. Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* **25**, 975–979 (1953).
2. Hudspeth, A. J. Integrating the active process of hair cells with cochlear function. *Nat. Rev. Neurosci.* **15**, 600–6614 (2014).
3. Kandel, E. R., Schwartz, J. H. & Jessell, T. M. (eds) *Principles of Neural Science* 4th edn (Elsevier, 2000).
4. Robles, L. & Ruggero, M. A. Mechanics of the mammalian cochlea. *Physiol. Rev.* **81**, 1305–1352 (2001).
5. Ashmore, J. et al. The remarkable cochlear amplifier. *Hear. Res.* **266**, 1–17 (2010).
6. Peng, A. W. & Ricci, A. J. Somatic motility and hair bundle mechanics, are both necessary for cochlear amplification? *Hear. Res.* **273**, 109–122 (2011).
7. Guinan Jr, J. J. Olivocochlear efferents: their action, effects, measurement and uses, and the impact of the new conception of cochlear mechanical responses. *Hear. Res.* **362**, 38–47 (2018).
8. Sivian, L. J. & White, S. D. On minimum audible sound fields. *J. Acoust. Soc. Am.* **4**, 288–321 (1933).
9. Sasmal, A. & Grosh, K. The competition between the noise and shear motion sensitivity of cochlear inner hair cell stereocilia. *Biophys. J.* **114**, 474–483 (2018).
10. Schafer, P. B. & Jin, D. Z. A novel concept for dynamic adjustment of auditory space. *Sci. Rep.* **8**, 8335 (2018).
11. Milewski, A. R., Ó Maoiléidigh, D., Salvi, J. D. & Hudspeth, A. J. Homeostatic enhancement of sensory transduction. *Proc. Natl Acad. Sci. USA* **114**, E6794–E6803 (2017).
12. Fettiplace, R. Hair cell transduction, tuning, and synaptic transmission in the mammalian cochlea. *Compr. Physiol.* **7**, 1197–1227 (2017).
13. Adavanne, S., Politis, A., Nikunen, J. & Virtanen, T. Sound event localization and detection of overlapping sources using convolutional recurrent neural networks. *IEEE J. Sel. Topics Signal Process.* **13**, 34–48 (2018).
14. Alvarez, R. & Park, H.-J. End-to-end streaming keyword spotting. In *ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 6336–6340 (IEEE, 2019).
15. Abeßer, J. A review of deep learning based methods for acoustic scene classification. *Appl. Sci.* **10**, 2020 (2020).
16. Wu, J., Yılmaz, E., Zhang, M., Li, H. & Tan, K. C. Deep spiking neural networks for large vocabulary automatic speech recognition. *Front. Neurosci.* **14**, 199 (2020).
17. Araujo, F. A. et al. Role of non-linear data processing on speech recognition task in the framework of reservoir computing. *Sci. Rep.* **10**, 328 (2020).
18. Schafer, P. B. & Jin, D. Z. Noise-robust speech recognition through auditory feature detection and spike sequence decoding. *Neural Comput.* **26**, 523–556 (2014).
19. Barker, J., Vincent, E., Ma, N., Christensen, H. & Green, P. The PASCAL CHiME speech separation and recognition challenge. *Comput. Speech Lang.* **27**, 621–633 (2013).
20. Zai, A. T., Bhargava, S., Mesgarani, N. & Liu, S.-C. Reconstruction of audio waveforms from spike trains of artificial cochlea models. *Front. Neurosci.* **9**, 347 (2015).
21. Wen, B. & Boahen, K. A silicon cochlea with active coupling. *IEEE Trans. Biomed. Circuits Syst.* **3**, 444–455 (2009).
22. Kiselev, I. & Liu, S.-C. Event-driven local gain control on a spiking cochlea sensor. In *2021 IEEE International Symposium on Circuits and Systems (ISCAS)* 1–5 (IEEE, 2021).

23. Wang, S., Koickal, T. J., Hamilton, A., Cheung, R. & Smith, L. S. A bio-realistic analog CMOS cochlea filter with high tunability and ultra-steep roll-off. *IEEE Trans. Biomed. Circuits Syst.* **9**, 297–311 (2015).

24. Raychowdhury, A. et al. A 2.3 nJ/frame voice activity detector-based audio front-end for context-aware system-on-chip applications in 32-nm CMOS. *IEEE J. Solid-State Circuits* **48**, 1963–1969 (2013).

25. Price, M., Glass, J. & Chandrakasan, A. P. A 6 mW, 5,000-word real-time speech recognizer using WFST models. *IEEE J. Solid-State Circuits* **50**, 102–112 (2015).

26. van Schaik, A. & Liu, S.-C. AER EAR: a matched silicon cochlea pair with address event representation interface. *IEEE Trans. Circuits Syst. I, Reg. Papers* **5**, 4213–4216 (2005).

27. 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).

28. Liu, S.-C., van Schaik, A., Minch, B. A. & Delbruck, T. Asynchronous binaural spatial audition sensor with 2×64×4 channel output. *IEEE Trans. Biomed. Circuits Syst.* **8**, 453–464 (2013).

29. Yang, M., Chien, C.-H., Delbruck, T. & Liu, S.-C. A 0.5 V 55 μW 64×2-channel binaural silicon cochlea for event-driven stereo-audio sensing. In *2016 IEEE International Solid-State Circuits Conference (ISSCC)* **59**, 388–389 (IEEE, 2016).

30. 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).

31. 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 (IEEE, 2014).

32. 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).

33. Xu, Y. et al. A FPGA implementation of the CAR-FAC cochlear model. *Front. Neurosci.* **12**, 198 (2018).

34. Singh, R. K. et al. CAR-Lite: a multi-rate cochlear model on FPGA for spike-based sound encoding. *IEEE Trans. Circuits Syst. I, Reg. Papers* **66**, 1805–1817 (2019).

35. Nouri, M. et al. A Hopf resonator for 2-D artificial cochlea: piecewise linear model and digital implementation. *IEEE Trans. Circuits Syst. I, Reg. Papers* **62**, 1117–1125 (2015).

36. Kim, Y. et al. A novel frequency selectivity approach based on travelling wave propagation in mechanoluminescence basilar membrane for artificial cochlea. *Sci. Rep.* **8**, 12023 (2018).

37. Windmill, J. Biologically inspired acoustic sensors: from insect ears to miniature microphones. *J. Acoust. Soc. Am.* **143**, 1777 (2018).

38. 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).

39. 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).

40. 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).

41. Guerreiro, J. et al. Enhancing acoustic sensory responsiveness by exploiting bio-inspired feedback computation. In *ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 1478–1482 (IEEE, 2019).

42. Smith, L. S. Toward a neuromorphic microphone. *Front. Neurosci.* **9**, 398 (2015).

43. Wen, B., Wang, G. I., Dean, I. & Delgutte, B. Time course of dynamic range adaptation in the auditory nerve. *J. Neurophysiol.* **108**, 69–82 (2012).

44. Rangelow, I. W. et al. Review article: Active scanning probes: a versatile toolkit for fast imaging and emerging nanofabrication. *J. Vac. Sci. Technol. B* **35**, 06G101 (2017).

45. Zöllner, J.-P. et al. Gas-flow sensor based on self-oscillating and self-sensing cantilever. *Proceedings* **2**, 846 (2018).

46. Stauffenberg, J. et al. Determination of the mixing ratio of a flowing gas mixture with self-actuated microcantilevers. *J. Sens. Sens. Syst.* **9**, 71–78 (2020).

47. Schnupp, J. *NetworkReceptiveFields* (2016); https://osf.io/ayw2p/

48. Jesteadt, W., Wier, C. C. & Green, D. M. Intensity discrimination as a function of frequency and sensation level. *J. Acoust. Soc. Am.* **61**, 169–177 (1976).

49. Hudspeth, A. J. Making an effort to listen: mechanical amplification in the ear. *Neuron* **59**, 530–545 (2008).

50. Hudspeth, A. J., Choe, Y., Mehta, A. D. & Martin, P. Putting ion channels to work: mechanoelectrical transduction, adaptation, and amplification by hair cells. *Proc. Natl Acad. Sci. USA* **97**, 11765–11772 (2000).

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

52. Strogatz, S. H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering* (Westview Press, 2000).

53. Roeser, D. et al. Tip motion-sensor signal relation for a composite SPM/SPL cantilever. *J. Microelectromech. Syst.* **25**, 78–90 (2016).

54. Gomez, F., Lorimer, T. & Stoop, R. Signal-coupled subthreshold Hopf-type systems show a sharpened collective response. *Phys. Rev. Lett.* **116**, 108101 (2016).

55. Jean, P. et al. Macromolecular and electrical coupling between inner hair cells in the rodent cochlea. *Nat. Commun.* **11**, 3208 (2020).

56. Dierkes, K., Lindner, B. & Jülicher, F. Enhancement of sensitivity gain and frequency tuning by coupling of active hair bundles. *Proc. Natl Acad. Sci. USA* **105**, 18669–18674 (2008).

57. Zilany, M. S., Bruce, I. C., Nelson, P. C. & Carney, L. H. A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. *J. Acoust. Soc. Am.* **126**, 2390–2412 (2009).

58. Knudsen, E. I. & Konishi, M. Mechanisms of sound localization in the barn owl (*Tyto alba*). *J. Comp. Physiol.* **133**, 13–21 (1979).

59. Boder, D. P. & Goldman, I. L. The significance of audible onset as a cue for sound localization. *J. Exp. Psychology* **30**, 262–272 (1942).

60. Perrott, D. R. Role of signal onset in sound localization. *J. Acoust. Soc. Am.* **45**, 436–445 (1969).

61. Goodridge, S. & Kay, M. Multimedia sensor fusion for intelligent camera control. In *1996 IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems (Cat. no. 96TH8242)* 655–662 (IEEE, 1996).

62. Huang, J., Ohnishi, N. & Sugie, N. A biomimetic system for localization and separation of multiple sound sources. *IEEE Trans. Instrum. Meas.* **44**, 733–738 (1995).

63. Lenk, C., Seeber, L., Ziegler, M., Hövel, P. & Gutschmidt, S. Enabling adaptive and enhanced acoustic sensing using nonlinear dynamics. In *2020 IEEE International Symposium on Circuits and Systems (ISCAS)* 1–4 (IEEE, 2020).

Acknowledgements

N. Lam and S. Hayashi for fruitful discussions regarding the modelling of the system.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41928-023-00957-5.

Correspondence and requests for materials should be addressed to Claudia Lenk.

Peer review information *Nature Electronics* thanks Bastian Epp for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.