

ARTICLE

Open Access

# Concept and proof for an all-silicon MEMS micro speaker utilizing air chambers

Bert Kaiser<sup>1</sup>, Sergiu Langa<sup>1,2</sup>, Lutz Ehrig<sup>1</sup>, Michael Stolz<sup>1</sup>, Hermann Schenk<sup>1</sup>, Holger Conrad<sup>1</sup>, Harald Schenk<sup>1,2</sup>, Klaus Schimmanz<sup>2</sup> and David Schuffenhauer<sup>1</sup>

## Abstract

MEMS-based micro speakers are attractive candidates as sound transducers for smart devices, particularly wearables and hearables. For such devices, high sound pressure levels, low harmonic distortion and low power consumption are required for industrial, consumer and medical applications. The ability to integrate with microelectronic circuitry, as well as scalable batch production to enable low unit costs, are the key factors benchmarking a technology. The Nanoscopic Electrostatic Drive based, novel micro speaker concept presented in this work essentially comprises in-plane, electrostatic bending actuators, and uses the chip volume rather than its surface for sound generation. We describe the principle, design, fabrication, and first characterization results. Various design options and governing equations are given and discussed. In a standard acoustical test setup (ear simulator), a MEMS micro speaker generated a sound pressure level of 69 dB at 500 Hz with a total harmonic distortion of 4.4%, thus proving the concept. Further potential on sound pressure as well as linearity improvement is outlined. We expect that the described methods can be used to enhance and design other MEMS devices and foster modeling and simulation approaches.

## Introduction

There is great demand for high performance micro electromechanical systems (MEMS) micro speakers by the consumer and by the hearing aid industry. Key factors, such as high sound pressure level (SPL), low total harmonic distortion (THD), low power consumption, small device footprint and batch fabrication possibilities are needed to outperform classic micro speaker technologies like moving coil and balanced armature receiver technology. MEMS-based micro speakers achieve all these key factors. Successful devices made with MEMS, like inertial sensors, timing devices and pressure sensors, are ubiquitous in many electrical applications. These applications range from vehicles to wearables. A prominent example is smartphones, which have all the types of MEMS mentioned above. They also include acoustic transducers, most likely appearing as multiple MEMS microphones, to

aid their original primary application, voice communication. In contrast to fabrication, pricing and integration logic, today's smartphones include fine mechanical engineered, non-MEMS micro speakers. Hearables, which are wearables plugged directly into the ear canal and relying mostly on acoustics for their human-machine interface, are expected to take over the acoustic functionality of smartphones and extend their application to that of personal assistant. In earbuds, the available space is limited, thus demanding exceptional miniaturization and integration of all components including electronics, sensors, batteries, and acoustic transducers.

Several concepts for MEMS-based micro speakers have been reported in the literature. Among them are the following examples. Cheng et al.<sup>1</sup> published an electromagnetically driven MEMS micro speaker comprising an electroplated, micro machined membrane with a diameter of 3.5 mm and a permanent magnet assembled underneath. The device reached 93 dB at 5 kHz in a 2 cm<sup>3</sup> closed cavity, consuming 320 mW of power. No results for the THD were published. Chen et al.<sup>2</sup> used a 3.5 mm

Correspondence: Bert Kaiser (bert.kaiser@ipms.fraunhofer.de)

<sup>1</sup>Fraunhofer-Institute for Photonic Microsystems, 01109 Dresden, Germany

<sup>2</sup>Chair of Micro and Nano Systems, Brandenburg University of Technology Cottbus-Senftenberg, 03013 Cottbus, Germany

© The Author(s) 2019



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

diameter polymer membrane on a silicon substrate, incorporating a single loop copper coil and two magnets for electromagnetic driving of the speaker. They were able to demonstrate 106 dB at 1 kHz in a 2 cm<sup>3</sup> closed cavity, with an input power of 1.76 mW. No values for the THD were reported. Albach et al.<sup>3</sup> published the concept of a MEMS-based micro speaker driven by a changing magnetic field. The concept describes a parallel arrangement of bending actuators, using a magnetostrictive active layer with a photoresist passive layer. The actuators form an acoustically closed membrane of about 7.5 mm<sup>2</sup>, despite the gaps between the individual cantilevers. In a hybrid setup, Albach et al. drove the MEMS devices with externally generated magnetic fields. A SPL 101 dB was measured in a 2 cm<sup>3</sup> closed cavity at 400 Hz. No values for either the THD or power consumption were presented. A sound transducer based on the piezoelectric bending of actuators was presented by Stoppel et al.<sup>4</sup> They describe a concept where the laterally expanded, triangular actuators deflect out-of-plane according to the bimorph effect. For a driving voltage of 2 V<sub>pp</sub> they were able to measure a peak SPL of 86 dB (2 cm<sup>3</sup> closed cavity) with a THD of less than 2% including a preprocessed input signal. A sensitivity of more than 105 dB/mW was achieved. Roberts et al.<sup>5</sup> presented the concept of an electrostatically actuated MEMS micro speaker. It features a silicon carbide membrane on a silicon back-plate with an 8 μm electrostatic gap. They were able to measure a SPL of 73 dB at 16.6 kHz at a distance of 10 mm in free field radiation with a driving voltage of 200 V<sub>pp</sub>. From the deflection and active area data given, this equals an approximate SPL of 65 dB in a closed 2 cm<sup>3</sup> cavity. Neither the THD nor the power consumption data were reported. Fei et al.<sup>6</sup> presented a thermoacoustic device based on graphene foam. The device exhibits an active area of 100 mm<sup>2</sup>. An SPL of 75 dB at 10 kHz in a free field measurement at 3 cm is reported for an input power of 1000 mW. This translates to approximately 84 dB in a 2 cm<sup>3</sup> cavity. For the sake of completeness it should be noted that several works have been conducted with parametric transducers, utilized in arrays<sup>7</sup>, for digital sound reconstruction<sup>8–11</sup>. These approaches, which are presented in literature, are assumed to be relatively large because of the large number of single membrane drivers arranged in parallel. This renders the devices not beneficial for miniaturization in the first place.

The literature overview reveals various mechanical approaches, with at least three different actuating mechanisms. Except for the thermoacoustic devices, they all actuate single or multiple membranes or similar working entities which deflect out-of-plane. These membranes or plates exhibit different sizes and shapes and are sometimes closed or slitted.

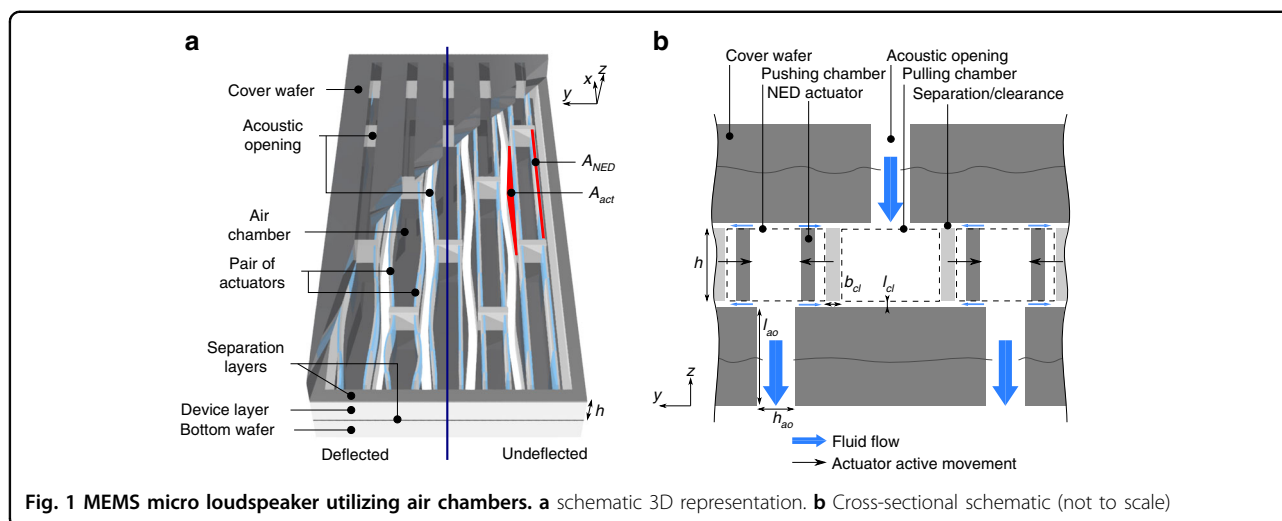
Highly miniaturized MEMS concepts which do not require complex or elaborate assembly technologies, but rather allow the use of broadly available manufacturing technologies (are compatible with complementary metal-oxide semiconductor (CMOS) technology) are highly favored for scalable mass production.

In this paper we present a novel design concept for MEMS micro speakers based on the fabrication technologies referred to as silicon bulk micro machining, which is fully CMOS compatible. The design comprises all-silicon electrostatic, in-plane bending micro actuators working in air chambers, thus utilizing the chip's bulk volume for air displacement and sound pressure generation. This micro speaker design approach allows the substitution of chip area for volume. In this work governing principles and parameter relations are presented and the concept's potential is estimated. By fabricating and characterizing fabricated specimen, the concept is proven. To validate a concept, performance in terms of attainable sound pressure has to exceed the hearing threshold and should at least reach sound pressure values of normal speech, e.g., 55 dB at 1 kHz, and surpass previously reported values for electrostatic transducers<sup>5</sup>. Distortion in terms of THD should be below 10% and more likely below 5% as indicated by listening test<sup>12</sup>.

## Results

The nanoscopic electrostatic drive (NED) based micro speaker concept is shown in Fig. 1a. The device comprises clamped-clamped electrostatic bending actuators, the NED<sup>13</sup>, placed pairwise in rows and columns within the device layer of a bonded silicon on insulator (SOI) wafer and covered by another wafer bonded to the SOI wafer with a tiny separation. Between each neighboring row of actuators, acoustically effective openings are integrated in the top and bottom of the wafer in an alternating way to allow the emission of sound from the device without acoustic shortcut.

The clamped-clamped configuration, when based on a suitable design, gives an efficient double-S bending shape to the NEDs. Every other row of paired actuators is shifted along the longitudinal axis by half the length of one actuator to make denser packing of the actuators possible whilst considering the bending shape. When a driving voltage is applied across the electrodes of the individual but identical actuators, they bend laterally in-plane according to their orientation in an odd mode relative to each other, thus balancing inertial forces within the device. With the top and bottom wafers as closing covers, each space between any of the actuators creates an air chamber. The function of each chamber at any one moment is to push air out into or to pull air in from the surroundings, depending on the direction of its associated actuators' active deflection.



**Fig. 1** MEMS micro loudspeaker utilizing air chambers. **a** schematic 3D representation. **b** Cross-sectional schematic (not to scale)

As a consequence of the deflection of the actuators, a fraction of the air volume  $\Delta V_{\text{act}}$  of a pushing chamber is transferred out of the device through the associated acoustic opening. The same quantity of air is transferred into the device, i.e., into the pulling chambers. When the driving voltage is turned off, the restoring mechanical forces of the actuator beam bring it back to its rest position and invert the previous volume flow. Therefore, each pushing chamber becomes a pulling chamber and vice versa. This process may be driven at any frequency  $f$  in the audible range, thus generating sound. Considering inertial forces in a dynamic actuation regime the NEDs deflection may also get inverted, i.e., deflection passes beyond the mechanical rest position. A fact to be considered in the device design. The actuators must not pass the acoustic opening aperture in any operational state, as this will cause loss of sound pressure due to acoustic shortcutting.

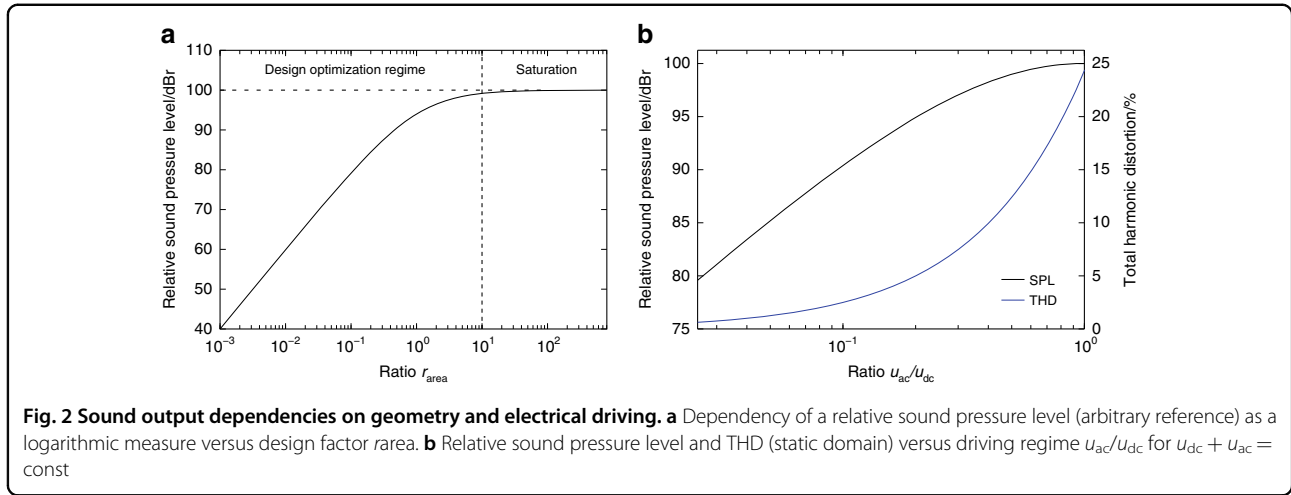
The tiny separation between the in-plane deflecting actuators and the top and bottom wafer ensures movability of the actuators (see Fig. 1b). However, the separation represents a fluidic path between oppositely pressured chambers, thus enabling an acoustic shortcut. Geometrically, the clearance should be designed in such a way that losses are negligible, most likely by minimizing the distance between the actuators and the covers. In contrast, electrostatic forces generated by fringing electrostatic fields around the actuator tend to deflect the actuators in the  $z$ -direction, closing the clearance, and are reduced by a larger clearance. A vertical movement in the  $z$ -direction must be considered, since perfect symmetry of the upper and lower clearances is not to be expected in practice. Since the vertical stiffness of an actuator is proportional to  $l^{-3}$ , the size of the clearance limits the viable actuator length  $l$ . If not designed properly, the vertical movement ultimately results in pull-in if the

voltage exceeds the critical value  $U_{\text{pi}}$ . This snap-in effect<sup>14</sup> would create undesired friction between the actuator and the bottom or cover wafers, resulting in distortion (“rub & buzz”<sup>15</sup>) and substantially deteriorated sound quality. Furthermore, vertical pull-in may also result in electrical shortcutting and may lead to device failure.

The height of the actuators equals the SOI layer thickness and specifies the displaced volume for a given lateral actuator displacement. In a closed acoustic emission cavity, the sound pressure is proportional to the quotient of the volume change and the initial volume of the cavity. Thus, a given layout may be extended to an increased height into the SOI layer, thereby increasing the generable sound pressure without affecting the footprint of the device. This is a unique characteristic of the micro speaker concept presented in this paper and is in contrast to any other approach relying on membranes. The design possibilities, some of which have already been described above, comprise actuator design, actuator arrangement, actuator count, fluidic channels and damping entities. These all further shape the uniqueness of the concept. Nevertheless, the various layout options result in a rather complex design process and must be closely aligned with the technological boundary conditions, most likely relating to minimal widths and lengths, to exploit the full potential of the concept. To estimate this potential, fundamental aspects are presented in the following section.

### Theoretical description

Attainable sound pressure is the prominent performance indicator for micro speakers. In the following section, the limits are estimated using a geometric approach. For a given SOI layer thickness  $h$ , the acoustic performance is then specified by the arrangement and number of actuators, each exhibiting a certain geometry and area consumption  $A_{\text{NED}}$ . While being displaced, the



actuators sweep across an area  $A_{act}$ , thus defining the ratio  $r_{area} = A_{act}/A_{NED}$ . The area  $A_{act}$  of the double-S bending shape may be simply calculated from the actuators' maximum deflection at its middle yielding  $A_{act} = w_{max} * l / 2$ . The area of the whole device  $A_{dev}$  equals the sum of the area of the actuators, considered a passive area, and the actively swept area, i.e.,  $A_{dev} = A_{NED} + A_{act} = A_{NED} * (1 + r_{area})$ . Comparing the actively scanned area to the device area defines the ratio  $r_{fill} = A_{act}/A_{dev} = 1/(1 + r_{area})$ . Essentially,  $r_{fill}$  is in the range  $0 < r_{fill} < 1$  and is practically limited to  $r_{fill} \ll 1$  because additional areas are required within  $A_{dev}$ , such as for mechanical anchors and fluidic or electrical paths, as well as areas related to technological constraints. These areas may be subsumed into  $A_{NED}$  for the sake of simplicity.

The SPL (symbol:  $L_p$ ) in a closed acoustic emission space of volume  $V_0$ , such as with earphones plugged into the ear canal, is not dependent on the position within this volume. Assuming wavelengths twice as big as the largest geometric dimension of  $V_0$  and that  $V_0$  is without leakage, the SPL can be simply calculated as follows. First, the pressure increase  $\Delta p$  is calculated as:

$$\Delta p = \kappa p_0 \Delta V / V_0 \quad (1)$$

where  $\kappa$ ,  $p_0$  and  $V_0$  are the adiabatic index, the initial pressure of air and the volume of the acoustic cavity, respectively. The total displaced volume  $\Delta V$  is the sum of all the volume changes caused by the single actuators,  $\Delta V_{act}$ . The SPL is then calculated as the logarithm of the pressure increase relative to an acknowledged referential pressure  $p_{ref}$ :

$$L_p = 20 \log \left( \frac{\Delta p}{p_{ref}} \frac{1}{4} \sqrt{2} \right) \quad (2)$$

The  $\sqrt{2}/4$  factor accounts for the RMS calculation of the sound pressure, assuming a harmonic sinusoidal

movement of the actuators. Substituting  $\Delta V = A_{act} * h = A_{dev} * r_{fill} * h$  gives

$$L_p = 20 \log \left( \frac{\kappa p_0 A_{act} h}{p_{ref} V_0} \frac{1}{4} \sqrt{2} \right) \quad (3)$$

$$L_p = 20 \log \left( \frac{\kappa p_0 A_{dev} r_{fill} h}{p_{ref} V_0} \frac{1}{4} \sqrt{2} \right)$$

Obviously, the maximum SPL is attained when  $r_{fill} = 1$  and when device area and SOI thicknesses are maximized. This implies that the ratio  $r_{area}$  has a practically useable limit of about  $r_{area} = 10$ , giving an  $r_{fill}$  of approximately 9/10. In other words, there is no interest in actuators which exhibit a deflection greater than ten times their associated lateral dimensions. A ratio  $r_{area} > 10$  would not increase attainable sound pressure significantly. Hence, the regime with values of  $r_{area} < 10$  is referred to as the design optimization regime. Values of  $r_{area} > 10$  are in the saturation regime. Figure 2a plots the attainable SPL against  $r_{area}$ , showing the different regimes. The SPL is normalized using arbitrarily chosen parameters  $A_{dev}$ ,  $h$  and  $V_0$  so that the maximum achievable sound pressure for  $r_{fill} = 1$  and  $r_{area} \gg 100$ , respectively yields 100 dB, since it is a relative measure. Additionally, SPL may be estimated for different measurement cavity volumes using Eq. (3).

The previous SPL calculation assumes that the fluid transfer in and out of the device is through large enough openings which do not lead to, e.g., significant damping in the audible frequency range. Another important aspect is the split of fluid flows within the device, i.e., the differences in flow resistance between different fluidic paths. In the following, a static analysis is given. The flow resistivity of the acoustic openings may be estimated from the law of Hagen-Poiseuille<sup>16</sup>:

$$\dot{V}_{ac} = K * \min(b_{ao}, h_{ao})^3 * \frac{\max(b_{ao}, h_{ao})}{12\eta l_{ao}} \Delta p_{ao} \quad (4)$$

where  $b_{ao}$ ,  $h_{ao}$  and  $l_{ao}$  are the width, height and length of the acoustic opening channels respectively. It should be noted that the fluidic flow is in the direction parallel to the length of the opening channel  $l_{ao}$ . The viscosity of air is encoded in  $\eta$ . The correction factor  $K$  may be estimated as follows<sup>16</sup>:

$$K = 1 - \sum_{n=1}^{\infty} \frac{1}{(2n-1)^5} \cdot \frac{192}{\pi^5} \cdot \frac{\min(b, h)}{\max(b, h)} \tanh\left((2n-1) \frac{\pi \max(b, h)}{2 \min(b, h)}\right) \quad (5)$$

The pressure amplitude  $\Delta p_{ao}$  may be derived from Eq. (1), taking  $V_0$  as the volume of one pushing or pulling chamber within the device (refer to Fig. 1b) and substituting  $\Delta V$  with two single actuator displaced volumes  $\Delta V_{act}$ , so that  $\Delta p_{ao} \propto 2 \Delta V_{act}$ . The fluidic losses through the clearance, the rectangular opening along the actuator, may also be estimated. Two clearances of similar geometry are active for each actuator, with the  $\Delta p$  being doubled because they have opposite signs on either side of the actuator, giving  $\Delta p_{cl} \propto 2 \Delta V_{act}$ . In contrast to the acoustic openings, a high fluidic flow resistance is desired for the clearance channels. The clearance channel has length  $l_{cl}$ , width  $b_{cl}$  and height  $h_{cl}$ . According to a typical layout, it is assumed that  $b_{ao} = b_{cl}$ ,  $h_{ao} \gg h_{cl}$  and  $l_{ao} \gg l_{cl}$  holds true. The ratio between the flow resistances of the clearance and acoustic openings may then be simplified as follows:

$$\frac{\dot{V}_{cl}}{\dot{V}_{ao}} \propto 2 \frac{h_{cl}^3}{h_{ao}^3} * \frac{l_{ao}}{l_{cl}} \ll 1 \quad (6)$$

Thus, the fluidic flow around the actuators through the clearance is significantly lower than through the acoustic openings. Consequently, no reduction in the acoustically effective volume flow must be assumed because of the air chamber concept.

The SPL described before defines the objective loudness of the acoustic transducer, i.e., acoustic quantity. In contrast, the THD defines the acoustic finesse, i.e., the acoustic quality of an acoustic transducer, making it an important measure for comparing sound reproduction. The THD shall therefore be discussed briefly and theoretically estimated in the following.

The working principle of the NED transducer is based on Coulomb forces with a quadratic dependence on the driving voltage as well as on the inverse of the electrode distance (Coulomb nonlinearity), which diminishes with increasing voltage<sup>13</sup>. The latter aspect is greatly influenced by the mechanics defined through the layout and how this affects the electrodes', and consequently the actuators', movements. For a clamped-clamped actuator configuration as given in our approach, a stress, resulting from deflection and inducing stiffening, i.e., nonlinear mechanical behavior known as the Duffing effect<sup>17</sup>, is to be expected. This effect develops contrarily to the

previously mentioned electrode separation nonlinearity with increasing deflection of the whole actuator. Therefore, the focus is on the quadratic proportionality on the driving voltage only (a simplification). Adding an offset voltage  $u_{dc}$  is commonly known to enhance linearity, i.e., to decrease the distortion. Thus, the driving voltage is:

$$U = u_{ac} \cos(\omega t) + u_{dc} \quad (7)$$

The THD (symbol  $k$ ; according to International Electrotechnical Commission IEC) is defined as the sum of all power  $p_i$  radiated in frequencies other than the desired excitation frequency, relative to the total emitted sound power.

$$k = 100\% \sqrt{\frac{\sum_{i=2}^n p_i}{\sum_{i=1}^n p_i}} \quad (8)$$

The actuator force  $F$  and deflection  $x$ , and thus its sound generation capability, are proportional to the square of the driving voltage  $U$ :

$$F(t) \sim (u_{ac} \cos(\omega t) + u_{dc})^2 = \left(\frac{u_{ac}^2}{2} + u_{dc}^2\right) + 2u_{ac}u_{dc} \cos(\omega t) + \frac{u_{ac}^2}{2} \cos(2\omega t) \quad (9)$$

The average radiated power  $p_i$  is then proportional to the square of the signal:

$$\bar{P} \propto \frac{1}{T} \int_0^T dt x^2(t) \quad \text{with } x \propto F \quad (10)$$

Comparing the amplitudes in Eq. (9),  $k$  reveals its dependency as follows:

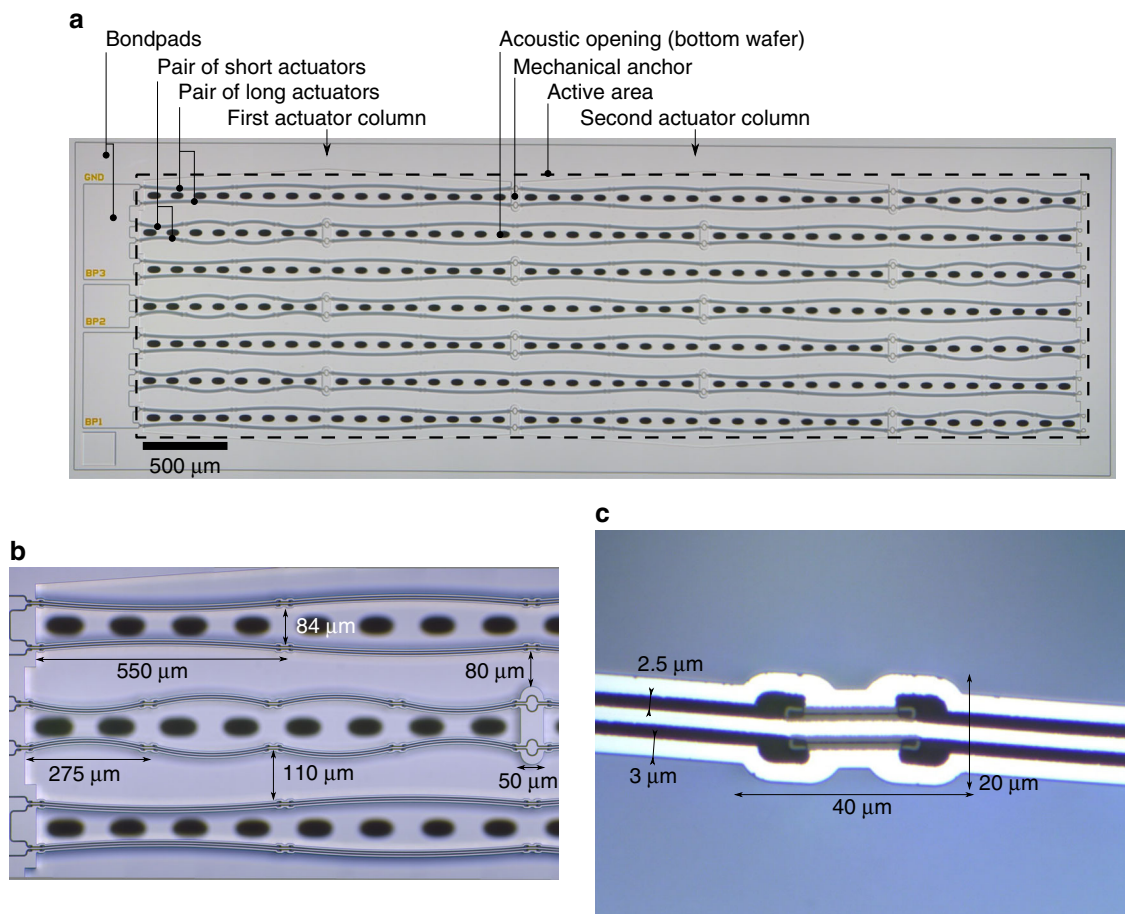
$$k = 100\% \sqrt{\frac{u_{ac}^2}{u_{ac}^2 + 16u_{dc}^2}} \quad \text{with } |u_{ac} + u_{dc}| < U_{pi} \quad (11)$$

Because of the voltage limits due to pull-in, the sum of offset voltage and signal voltage need to be smaller than the pull-in voltage  $U_{pi}$  at any time. Therefore, the ratio  $u_{ac}/u_{dc}$  needs to be minimized to minimize the THD. The SPL is dominated by the proportionality of  $u_{dc}$  and  $u_{ac}$ . The SPL is then maximal for  $u_{ac}/u_{dc} = 1$ . In Fig. 2b, SPL and THD are plotted against the ratio  $u_{ac}/u_{dc}$ . From this, a ratio  $u_{ac}/u_{dc} < 1$ , e.g.,  $u_{ac}/u_{dc} = 1/8$ , seems reasonable to get low distortion of around 3% and a relative SPL of  $-8$  dB (factor 0.4). For real devices, the SPL and THD follow more complex dependencies, including the mechanical nonlinearities mentioned above and fluidic interactions, e.g. from the acoustic load.

### Transducer design

A first design was created and fabricated. The layout of the SOI layer is shown in Fig. 3a. It shows an arrangement





**Fig. 3 Geometric setup of fabricated MEMS speaker** **a** Top view of the device layer showing actuators and their arrangement. Sample was taken out during fabrication without the cover wafer being bonded. **b** Enlarged partial view of three rows of actuators. **c** Detail of an actuator including dimensions

of 14 pairs of long actuators in a 2-serial-7-parallel (2s7p) configuration, meaning that there are two actuator pairs in a row, and seven rows aligned parallel to the device plane. The actuators are bundled into three individually addressable groups (3-1-3 rows). To increase the fill factor of the design, remaining space to the edge of the preferred rectangular shaped device is filled with actuators of half the length. This adds seven pairs of half-length actuators in total. The footprint of the rectangular active area is  $9.3 \text{ mm}^2$ . Each of the double clamped actuators has a length of 2.2 mm and 1.1 mm respectively. Vertical pull-in for the maximum driving voltage of 48 V is safely avoided, even after taking fabrication variations into account. The NED actuators' shape was designed with finite element analysis-based optimization with the figure of merit being the displacement achievable while respecting technological boundary conditions (electrode thicknesses, trench widths, material properties). Details of the actuator arrangement and design properties are shown in Fig. 3b, c. The approximate width of the actuators is less than

20  $\mu\text{m}$ . Each pair of actuators exhibits a distance of about 84  $\mu\text{m}$  while each row of actuators is placed at about 110  $\mu\text{m}$  from each other.

The device layout incorporates large ( $250 \times 250 \mu\text{m}^2$ ) bond pads for wire bonding afterwards. Acoustic openings in the bottom and cover wafers are designed as rectangles located between every other actuator pair. The size of the opening is determined by the technological boundaries for etching, which defines their minimal width (30  $\mu\text{m}$ ). The bottom and top wafers are 400  $\mu\text{m}$  thick. The distance (clearance) between the actuator and the bottom wafer equals the thickness of the 1  $\mu\text{m}$  buried oxide layer of the bonded SOI wafer. This is also the design value for the separation to the top wafer, which can be defined by a deposited and patterned bonding interface layer. The SOI layer thickness is 75  $\mu\text{m}$ .

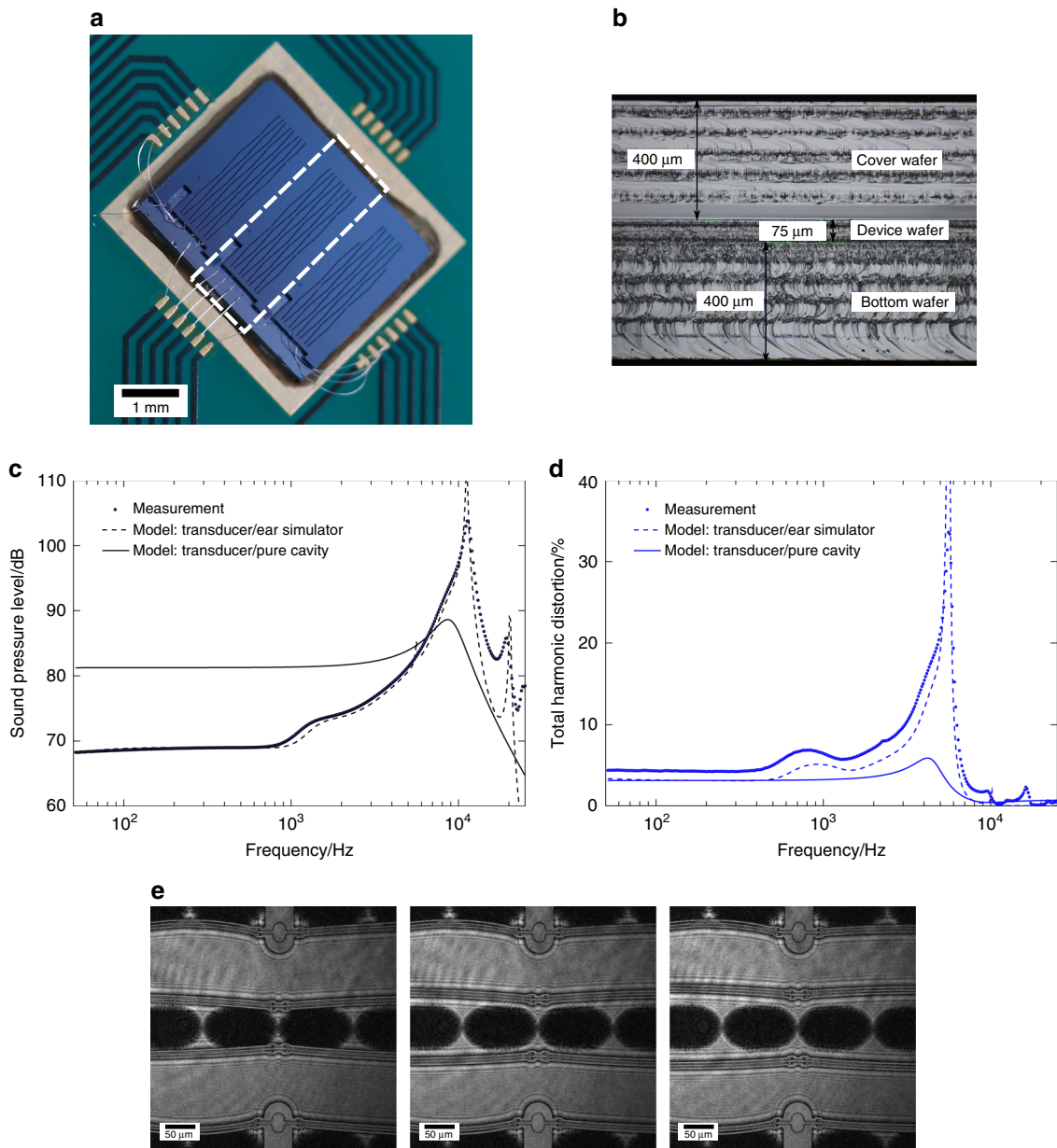
Three different transducer devices were fabricated on one single chip. Fabrication was performed on silicon wafers, solely using CMOS compatible processes. The main process is deep reactive ion etching (DRIE). Details

of the fabrication are supplied in the methods and materials section.

**Measurement setup and results**

To characterize the acoustic performance, SPL and harmonic distortion tests were performed. The bare

MEMS die layer stack, comprising the actual device, the bottom wafer and the cover wafer bonded on top, was glued to a carrier board which allows for electrical contacts using wire bonding (Fig. 4a). The layer stack is shown in Fig. 4b. A specially designed adaptor embracing the carrier board and the MEMS device allowed for the



**Fig. 4** MEMS device setup and acoustical as well as optical measurement data **a** Assembled device with cover wafer on top showing its acoustic openings; the MEMS chip is glued to a carrier board enabling wire bonds for electrical contacting. The chip actually comprises three different micro speakers, which were not diced for this test set-up. The middle one was used as shown in **(a)**. **b** Side view of the bonded wafer stack showing the three wafers and their thicknesses. Scallops resulting from cleavage through lacer dicing are clearly visible. **c** Measurement and modeling results for the SPL. Device was driven with  $u_{dc} = 40\text{ V}$  and  $u_{ac} = 10\text{ Vpp}$ . An SPL of 69 dB was achieved at 500 Hz. **d** THD measurement and modeling. THD value at 500 Hz is 4.4%. **e** Sequence of stroboscopic pictures from the holographic microscope showing three consecutive (time wise) deflections states in resonance of one actuator pair (specimen without cover wafer)

mounting of a standardized IEC 60318-4 ear simulator (formerly IEC 60711,<sup>18</sup>) incorporating a condenser microphone measuring in a closed cavity of 1.26 cm<sup>3</sup> (at the back end of the chip). Additionally, the adaptor ensured decoupling of the acoustic openings in the top and bottom wafers by framing the MEMS device. The front side of the chip was left open to the surroundings. Using an audio analyzer (NTI Flexus 100) and a voltage amplifier (Krohn-Hite Model 7602 M) to drive the device, the acoustic response to sinusoidal excitation ( $u_{dc} = 40$  V,  $u_{ac} = 10$  V<sub>pp</sub>) while sweeping the frequency was obtained, as shown in Fig. 4c, d. In addition, these figures also show the simulation results of the MEMS transducer and ear simulator combined, as well as of the MEMS transducer with an acoustic output impedance as specified in ref. <sup>19</sup>. Details of the simulation models are supplied in the “Materials and methods” section.

For frequencies up to 1 kHz, the SPL measurement ( $p_{ref} = 20$  μPa) shows a comparatively slight variation between 68 and 70 dB. In this frequency range, the quasistatic behavior of the MEMS transducer driving the acoustic load is visible. At a referential frequency of 500 Hz, the SPL was measured at 69 dB without being influenced by dynamical effects of either the transducer itself or the ear simulator. Beyond 1 kHz, these dynamical effects become visible, resulting in an increase in the SPL. Two characteristic frequencies may be distinguished, 1.7 kHz and 11 kHz. Both are attributed to the IEC ear simulator characteristics and may be estimated from equivalent circuit models (see methods and materials section). A maximum SPL of 104 dB was measured at 11.4 kHz. This is a superposition of the ear simulator resonance and the first mode transducer resonance. This first mode transducer resonance is also visible as a shoulder at a frequency of 9.1 kHz, agreeing well with electromechanical measurements of the long actuators. The contribution to sound generation from the short actuators may be neglected due to their low number, short length (cumulative length ratio 1/4) and very low deflection ( $x \propto l^{-2}$ ). From the curve of the transducer with a pure cavity, it can be concluded that the system is quite strongly damped (quality factor approximately  $Q = 2$ ). This curve shows a significantly higher SPL than the measurement for frequencies below 5 kHz. The losses in the measurements are attributed to dissipation within the ear simulator (i.e., its RCL-network).

The THD remains below 5% for frequencies up to 500 Hz, representing the quasistatic behavior of the transducer itself. An increase to <7% at half the 1.7 kHz resonance frequency of the ear simulator can be seen. The THD exhibits maximum values at a frequency of 5.7 kHz, the equivalent of the resonance of the ear simulator at 11.4 kHz as mentioned before. Again, the signature of the transducer itself is visible as a shoulder in the curve at

around 4.6 kHz, i.e., half its own resonance frequency. In contrast, the THD drops at the resonances of the transducer and the ear simulator. As before, the curve of the transducer with a pure cavity load exhibits no specific peaks except for the maximum at half the actuator resonance frequency.

For comparison, pure electromechanical measurements have also been carried out using a digital holographic microscope (Lyncee Tec R-2203). By taking a specimen without cover wafer (see Fig. 3) the lateral deflection could be obtained as a function of the driving signal frequency. The measurement results reveal a quasi-static deflection of about 1.5 μm at the middle of the 2.2 mm long actuators for  $u_{dc} = 40$  V and  $u_{ac} = 10$  V<sub>pp</sub> ( $f = 50$  Hz). These values fit well with the simulated deflection of the clamped-clamped beams. At resonance a deflection of about 22 μm was measured. Figure 4e shows a time wise sequence of consecutive deflections of one actuator pair.

The device’s total electrical capacitance was measured as 65 pF with an impedance spectrometer (Keysight E4990A) at a frequency of 500 Hz according to DIN 60268-7. Internal losses and dissipative effects, such as hysteresis losses, eddy current losses and heat generation, may be neglected for electrostatic transducers<sup>20</sup>. Assuming an efficiency of 40% for an appropriate driving circuit, handling the mostly reactive load of the electrostatic transducer, this gives a calculated sensitivity of 100 dB/mW (500 Hz, 1 mW, IEC 60318-4 ear simulator).

## Discussion

The fabricated and characterized device had a maximum SPL of 69 dB at 500 Hz in a IEC-ear simulator with a significant loss of SPL after resonance at around 10 kHz. A commonly known referential frequency response of headphones is the “Harman In-Ear Target Response Curve”<sup>21,22</sup>, which is the result of extensive listening tests to identify the preferred frequency responses of headphones. It shows a steep decline of the SPL from 9 kHz ongoing. A similar qualitative trend may be identified in the measurement curve. Thus, the applicability of the device does not seem to be limited by the specific resonance frequency of the actuators.

The measured SPL is enough to validate the new loudspeaker concept as an audible sound reproduction device. The THD for the transducer shows values of around 4.5% without any signal preprocessing at all. From the very simple theoretical estimations for the driving regime, one would expect a THD of about 3%. The more complex and still purely linear equivalent circuit model considers several more effects, such as the frequency dependencies of the clearance and the deflection. It is then able to predict the behavior in good agreement with the measurements for frequencies below any resonance.



To lower the THD values in general, i.e., below 4.4%, design optimization is proposed. This incorporates the use of mechanical and fluidic nonlinearities counteracting the inherent Coulomb nonlinearities. Additionally, setting up the transducer for its acoustic load, i.e., the ear simulator, e.g., by diversifying the resonances, will help to enhance the audio quality further.

Based on the measured values of the quasistatic deflection (1.5  $\mu\text{m}$ ) of the double-S bending shape, an actively swept area of approximately  $A_{\text{act}} = 1650 \mu\text{m}^2$  per actuator (28 in total) may be calculated. The estimated sound pressure then yields 77 dB in a 1.26  $\text{cm}^3$  volume. This value arises from the geometric approach presented in the theory section neglecting any loss in the RCL network of the ear simulator. Consequently, the value is greater than the measured SPL result (69 dB) in the ear simulator with the same volume.

Assuming a width of 20  $\mu\text{m}$  of the actuator ( $A_{\text{ned}} = 20 \mu\text{m} * 2.2 \text{mm}$ ), a factor  $r_{\text{area}} = 3/80$  may be estimated. This calculation neglects the additional required areas (e.g., bondpads, technology constraints, etc.) which would lower the factor because the effective  $A_{\text{ned}}$  would increase. As such, the very same factor calculated from the sound pressure yields  $r_{\text{area}} = 1/500$  ( $r_{\text{fill}} = 1/500$ ). It follows that the present proof-of-concept is in the design optimization regime. For the boundaries (chip area of about 9.3  $\text{mm}^2$  and SOI thickness of 75  $\mu\text{m}$ ), the maximum theoretical SPL with the presented concept is roughly 123 dB in a 1.26  $\text{cm}^3$  cavity. To paraphrase, a rather large improvement of 54 dB (123–69) can be expected theoretically. Practically, an improvement to the design, mostly by denser packing of more effective actuators, i.e., an increase in  $r_{\text{area}}$ , will allow an increase in  $r_{\text{fill}}$  to  $> 1/20$ , thus potentially adding  $> 20$  dB.

The work presented shows a MEMS micro speaker concept and its proof with a fabricated specimen. Several important advantages of the concept comprise CMOS compatibility including integration capabilities with driving circuitry, precision mass production capabilities, possible low area foot print for high sound quality and quantity and very low power consumption.

## Outlook

Further work focusing on SPL increase and THD decrease is being conducted. Denser packing of the actuators thanks to design optimization is expected to allow higher efficiencies, i.e., higher SPL per area footprint. The nature of the concept allows the volume of the MEMS die to be used for sound generation. This capability is extended significantly if the thicknesses of the device layer is increased. Theoretically, the layer thickness can be increased from 75  $\mu\text{m}$  to values as high as 725  $\mu\text{m}$ , a typical silicon wafer thickness. This would add 20 dB for the same area footprint and fill factor. This scaling is

connected with significant challenges in the technology, most prominently the DRIE process, which limits the achievable aspect ratio. The concept also allows for lateral acoustic openings, thus enabling the stacking of multiple MEMS dies to increase the SPL per area. Further optimization potential regarding distortion is expected to be released, with an optimized mechanical layout which makes use of the flexural behavior due to the actuators shape.

Future work will also address the speaker system, i.e., the development of driving circuitry comprising a step-up converter, signal amplifiers and, optionally, signal processing specifically designed for the device's needs. The challenge is to maintain the advantages of the concept regarding ultra-low power consumption. Recently, Hansler et al. have published work on capacitive loads and highly efficient driving circuitry<sup>23</sup>. They describe several amplifier classes as well as an approach to regain energy from reactive power. This is also of significant importance to the capacitive driven speaker concept presented here.

## Materials and methods

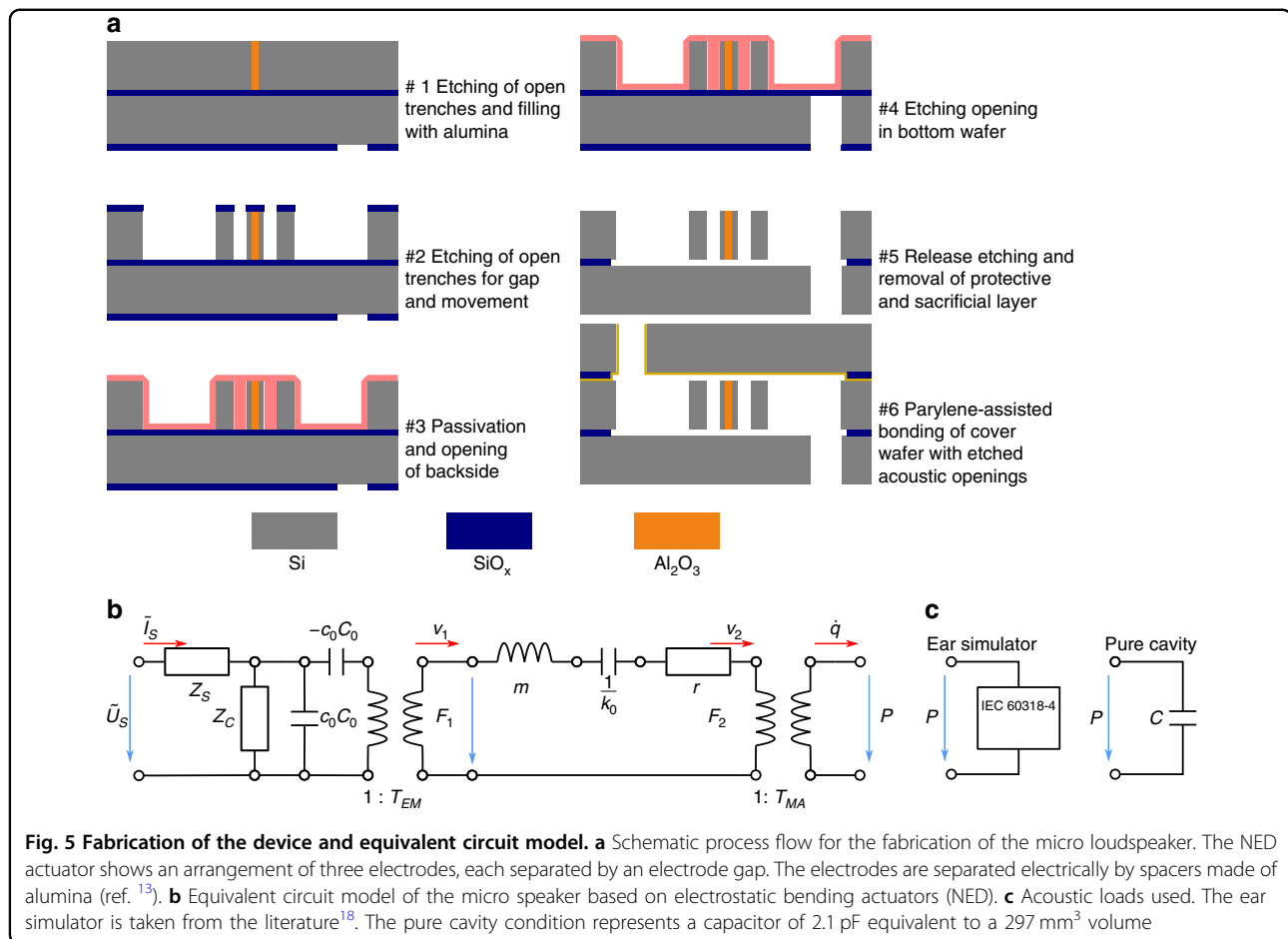
### Fabrication

The MEMS micro speaker was fabricated using parylene-assisted wafer bonding, DRIE etching and atomic layer deposition (ALD). The process flow is illustrated in Fig. 5a. A 200 mm diameter, highly boron doped ( $10^{18} \text{cm}^{-3}$ ) bonded SOI wafer with a 75  $\mu\text{m}$  device layer thickness was used as a substrate. First, open trenches were etched and subsequently filled with alumina using ALD, forming the spacers as mechanical connectors but electrical insulators in the NED actuators. Next, open trenches were etched, using DRIE again to build the actuators, anchors and the air chambers. Then, acoustic openings were etched from the backside through the handle layer (400  $\mu\text{m}$ ) of the bonded SOI wafer. The cover wafer was prepared by depositing and patterning the 1  $\mu\text{m}$  oxide layer facing the device layer and by etching the acoustic openings. Finally, the cover wafer (400  $\mu\text{m}$ ) was bonded on top of the device wafer using parylene (500 nm deposited on the cover wafer, 300  $^{\circ}\text{C}$ ) to assist wafer bonding<sup>24</sup>.

### Equivalent circuit modeling

The acoustic measurement results show the response of the MEMS transducer to the load of the IEC ear simulator. In this section, both the transducer and the IEC ear simulator are modeled using an equivalent acoustic circuit.

The equivalent circuit for the NED actuators is based on a simple plate-capacitor model presented in the literature<sup>25</sup>, which is transferred to an equivalent circuit according to the principles presented in<sup>26</sup>. The equivalent circuit model of the NED-based drive of the micro



speaker comprises the electrostatic, mechanical and acoustic domains. The model is shown in Fig. 5b. In this model, the NED effect is implemented by a leveraging factor translating the plate displacement to the beam displacement, and is encoded in  $T_{EM}$ . The actuator mass  $m$  is estimated from the layout. The electrical capacitance  $C_0$  and the resonance frequency were taken from measurements. The resonance frequency of 9.1 kHz was estimated from electromechanical measurements (Lyncee Tec R-2203) of the transducer without a cover wafer, taking into account the shift due to the offset voltage  $u_{dc} = 40$  V. Effective stiffness  $k_0$  is then calculated from resonance frequency and mass. The model is fitted using the pull-in voltage, leveraging factor ( $T_{EM}$ ) and damping coefficient  $r$  as fitting parameters. Pull-in voltage was fitted to 49 V as it agree with the simulation. The leveraging factor was also extracted from simulation in which a value of 51 was calculated. The difference to the fitted value of 42 is attributed to the rather complex mechanics. The damping has a substantial influence on the fitting and the fitted damping coefficient correlates to a quality factor of about 2.3. This value was qualified by acoustic measurements yielding a value of 1.5. For comparison, a

quality factor of approximately 30 was calculated from the mechanical measurements in ambient atmosphere. The observed large damping in the devices with covers is attributed to the fluid dynamics of the vents and the involved flow pattern inside the 28 actuator chambers.

The equivalent circuit for the IEC ear simulator is derived from the literature<sup>18</sup>, and uses the parametrization published by Brüel and Kjær<sup>19</sup>. The pure cavity condition reflects all but one component removed from the ear simulator network. The remaining component is the capacitor as shown in Fig. 5c right. The value of the capacity is 2.1 pF, which is equivalent to a 297 cm<sup>3</sup> volume of a cavity.

The models of both the transducer and the ear simulator are used to estimate the SPL and THD. The coupling of both models takes into account a certain amount of leakage between the ear simulator and the MEMS device. For the curves of the transducer with a pure cavity, presented in Fig. 4b, c, the most part of the IEC ear simulator circuit is removed. Consequently, these curves reveal behavior dominated by the MEMS transducer. The models will be presented in greater detail as well as with further development in a forthcoming publication.

**Acknowledgements**

This work was funded by the German Federal Ministry of Education and Research (BMBF) as part of the MEMSound project in the funding measure “Validation of the technological and social innovation potential of scientific research - VIP+” under the grant number 03VP01800.

**Authors' contributions**

B.K. created the layout of the MEMS device, accompanied fabrication and coordinated characterization. S.L. was responsible for fabrication in clean room facilities and consulted on layout. L.E. conducted acoustic setup and characterization and co-created equivalent circuit models. M.S. performed mechanical and electrical measurements. H.S. proposed and developed the equivalent circuit models. H.C. and H.S. proposed the concept and revised the manuscript. K.S. proposed and discussed the governing modeling approaches. D.S. supported the optimization of the NED actuators.

**Conflict of interest**

The authors declare that they have no conflict of interest.

Received: 27 February 2019 Revised: 7 August 2019 Accepted: 12 August 2019

Published online: 07 October 2019

**References**

- Cheng, M.-C., Huang, W.-S. & Huang, S. R.-S. A silicon microspeaker for hearing. *J. Micromech. Microeng.* **14**, 859–866 (2004). <https://doi.org/10.1088/0960-1317/14/7/004>.
- Chen, Y. C. & Cheng, Y. T. A low-power milliwatt electromagnetic microspeaker using a PDMS membrane for hearing aids application. 1213–1216 (2011).
- Albach, T. S., Horn, P., Sutor, A. & Lerch, R. Sound Generation Using a Magnetostrictive Micro Actuator. *J. Appl. Phys.* **109**(7), <https://doi.org/10.1063/1.3554200> (2011).
- Stoppel, F., Männchen, A. & F. Niekil, F. New integrated full-range mems speaker for in-ear applications. MEMS (Belfast, 2018).
- Roberts, R. C. et al. Electrostatically Driven Touch-Mode Poly-SiC Microspeaker. *IEEE Sensors* (2007).
- Fei, W., Zhou, J. & Guo, W. Low-voltage Driven Graphene Foam Thermoacoustic Speaker. *Communications*, 1068–1071 (Belfast, Northern Ireland, UK, 2018).
- Glacer, C., Dehé, A., Nawaz, M. & Laur, R. Reversible acoustical transducers in MEMS technology (DTIP 16–18 April 2013, Barcelona, Spain, 2013).
- Diamond, B., Neumann, J. & Gabriel, K. Digital sound reconstruction using arrays of CMOS-MEMS microspeakers, <https://doi.org/10.1109/MEMSYS.2002.984260>, (2002).
- Pedersen, B. R. Performance analysis of digital loudspeaker arrays. *J. Audio Eng. Soc.* **56**(No. 9), 2008 (2008).
- Dejaegera, R. et al. Development and Characterization of a Piezoelectrically Actuated MEMS Digital Loudspeaker, Proc. EuroSensors XXVI, September 9–12, 2012 (Kraków, Poland, 2012).
- Arevalo, A., Castro, D., Conchouso, D., Kosel, J. & Foulds, I. G. MEMS digital parametric loudspeaker, MEMS 2016, Shanghai, China, 24–28 January 2016 (2016).
- Temme, S., Olive, S., Tatarunis, S., Welti, T. & McMullin, E. The Correlation Between Distortion Audibility and Listener Preference in Headphones, Audio Engineering Society, 137th Convention, 1–15 (Los Angeles, USA, 2014).
- Conrad, H. et al. A small-gap electrostatic micro-actuator for large deflections. *Nat. Commun.* (2015). <https://www.nature.com/articles/ncomms10078>, <https://doi.org/10.1038/ncomms10078>.
- Zhang, W.-M., Yan, H., Peng, Z.-K. & Meng, G. Electrostatic pull-in instability in MEMS/NEMS: A review. *Sens. Actuators A* **214**, 187–218 (2014).
- Klippel, W., & Seidel, U. Measurement of Impulsive Distortion, Rub and Buzz and other Disturbances (AES, 2003).
- Bruus, H. *Theoretical Microfluidics* (Oxford University Press, 2008).
- Brennana, M. J., Kovacicb, I., Carrellaa, A. & Watersa, T. On the jump-up and jump-down frequencies of the Duffing oscillator. *J. Sound. Vibration* **318**(4–5), 1250–1261 (2008).
- International Electrotechnical Commission, Electroacoustics—simulators of human head and ear—Part 4: Occluded ear simulator for the measurement of earphones coupled to the ear by inserts (International Electrotechnical Commission, Geneva, Switzerland, 1981).
- Brüel & Kjær. *Application Note Ear Simulator for Telephonometry* (Brüel&Kjær Sound & Vibration Measurement A/S, Denmark, 2012).
- Hunt, F. V. *Electroacoustics - The Analysis of transduction, and Its Historical Background*. (Harvard University Press, Cambridge, Massachusetts, 1956).
- Olive, S., Welti, T. & Khonsaripour, O. The Influence of Program on Sound Quality Ratings of In-Ear Headphones, 142nd Audio Eng. Soc. Convention (Berlin, 2017).
- Olive, S., Welti, T. & Khonsaripour, O. The Preferred Low Frequency Response of In-Ear Headphones, AES International Conference on Headphone Technology, paper 6-1 (2016).
- Hänsler, M., Strle, D. & Auer, M. Design of a MEMS speaker driver. *Elektro. Inf.* **135/1**, 10–17 (2018).
- Kim, H., & Najafi, K. Characterization of low-temperature wafer bonding using thin-film parylene, *J. Microelectromech. Syst.* **14**, 1347–1355 (2005).
- Pelesko, J. A. & Bernstein, D. H. *Modeling MEMS and NEMS* (CHAPMAN & HALL/CRC, 2003).
- Beranek, L. L. *Acoustics* (Acoustical Society of America, Cambridge, Massachusetts, 1993).