DOI: 10.61782/fa.2023.0860

forum acusticum 2023

HIGHLY MINIATURIZED MEMS SPEAKERS FOR IN-EAR APPLICATIONS

Johannes Fankhänel* **Fabian Stoppel** **Thorsten Giese**
Isa Pieper **Lenny Castellanos** **Christian Eisermann**
Fraunhofer-Institute for Silicon Technology, Germany

ABSTRACT

Product development is often determined by conflicting demands for more performance and functionality on the one hand and size, cost, and power consumption constraints on the other hand. The resulting trend toward miniaturization has led to the development of MEMS-based loudspeaker technologies specifically designed for in-ear applications. Their full miniaturization potential, however, has not yet been exploited. Based on a previous generation, in this paper, a new MEMS loudspeaker technology is discussed. It operates without a closed membrane and relies on an acoustic shield and a rectangular actuator to significantly increase the generated sound pressure level per unit area. Compared with the previous generation, the active area can be reduced by a factor of three while maintaining the acoustic performance and optimizing damping. The concept is demonstrated using a new in-ear demonstrator featuring two MEMS loudspeakers with an emitting area as small as 2.4 × 2.4 mm² and a maximum sound pressure level of 105 dB_{SPL} and 89.7 dB_{SPL}/mm^2, respectively. Moreover, a simulation study indicates further potential for design optimization.

Keywords: *MEMS, loudspeaker, in-ear, headphones, miniaturization, piezoelectric.*

*Corresponding author: johannes.fankhaenel@isit.fraunhofer.de.

1. INTRODUCTION

Driven by new wireless audio applications and the constant trend toward miniaturization, micro loudspeakers based on microelectromechanical systems (MEMS) are recently gaining interest. As with MEMS microphones, which have become state of the art for mobile devices, manufacturing based on semiconductor technology enables high miniaturization and cost-efficient production in high volumes. Various MEMS speaker technologies for in-ear applications (e.g. headphones and hearables) have been developed in recent years [1]. Among them are several electrodynamic approaches, which are either actuated by a permanent magnet fixed to a polymer membrane and a substrate-based coil [2] or vice versa [3-6]. Other concepts generate sound waves using unimorph piezoelectric bending actuators [7, 8]. Unimorph structures are also used in combination with magnetostrictive actuation and a magnetic field created by a substrate-based coil [9]. Additionally, electrostatic MEMS loudspeakers are being researched [10, 11].

Two of the most promising approaches in terms of sound pressure level (SPL) per active area are currently being commercialized by the companies USound and xMEMS. Their models Achelous UT-P2020 [12] and Cowell [13] reach maximum estimated SPL of 86.9 dB_{SPL}/mm^2 at 250 Hz and 30 V_{pp} and 86.2 dB_{SPL}/mm^2 at 1 kHz and 30 V_{pp}, respectively. Despite great efforts and promising results, the miniaturization potential has not yet been fully exploited [1]. Accordingly, the main challenge remains to increase the SPL per device area.

To address this, in this paper, a novel MEMS speaker concept is discussed. It builds on a previous generation [8], which mimics a classical, circumferentially clamped loudspeaker membrane. By introducing narrow gaps and dividing the membrane into four triangular, mechanically decoupled, piezoelectric actuators, the deflections and the

generated sound pressure per unit area were significantly increased compared to conventional loudspeakers. With a total active area of $4 \times 4\ mm^2$, an SPL of more than 105 dB_{SPL} was realized over the entire reproduction range in an IEC 60318-4 ear simulator.

The concept discussed in this paper is enhanced with an acoustic shield and a rectangular actuator, increasing both the SPL and miniaturization at the same time.

2. DESIGN CONCEPT

The novel MEMS speaker design concept is schematically illustrated in Figure 1 a. In analogy to the previous generation [8], sound waves are created by exiting unimorph piezoelectric bending actuators with an alternating electric voltage. The actuators consist of four essential layers, a passive 15 µm poly-silicon (Si) layer, a platinum (Pt) bottom electrode, a 2 µm piezoelectric lead zirconate titanate (PZT) thin film, and a gold (Au) top electrode (Figure 1 b). The additional silicon dioxide (SiO_2) layers are needed for manufacturing purposes and to protect the device against external and production-related influences. Except for one clamped edge, a narrow 3 µm gap surrounding the actuator enables free vibrations of the structure while making use of thermo-viscous effects to suppress any noticeable airflow through the gap. A comprehensive description of the manufacturing of the presented structure has previously been published in [14].

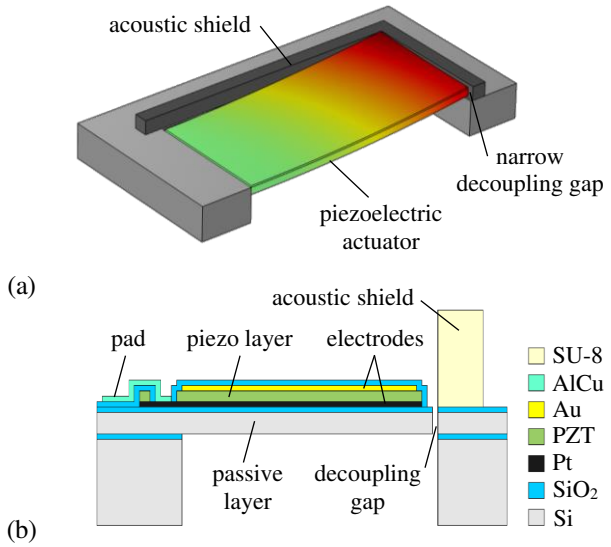

Figure 1. Illustration of the shield-based speaker concept (a); simplified cross-section of the final MEMS speaker (b).

The distinctive feature of this design concept is the addition of a 150 µm high acoustic shield made of SU-8 photoresist along the gap, which inhibits any significant gap widening when deflecting the actuator. Hence, the speaker acoustically behaves like a closed membrane while being mechanically decoupled at the same time. As opposed to the triangular actuators of the previous generation, this concept is not restricted to a symmetric, membrane-like design anymore and allows for arbitrary actuator shapes. Using rectangular actuators instead leads to a significant increase in average deflection and a drastically higher SPL per unit area.

3. VALIDATION OF THE DESIGN CONCEPT

To validate the design concept regarding its acoustic performance, finite element simulations were conducted with Comsol Multiphysics. The models comprise all relevant geometries, such as the MEMS speaker, the speaker housing, and an IEC 60318-4 ear simulator, as schematically illustrated in Figure 2. The influences of the microphone and the silicon earpiece connecting the housing and ear simulator are represented by two RCL impedances, which are calibrated using experimental measurements. To obtain the frequency response, the MEMS speaker is subjected to an offset electric AC voltage with varying frequency and $V_{AC} = V_{DC}$. The resulting speaker deflection is coupled to the pressure acoustics domain in terms of a two-way coupling. To correctly capture the boundary layer effects within the gap between the cantilever and the chip frame, a thermo-viscous-acoustics region is implemented. Ultimately, the frequency response is obtained by integrating the sound pressure over the rear face of the ear simulator opposing the headphones.

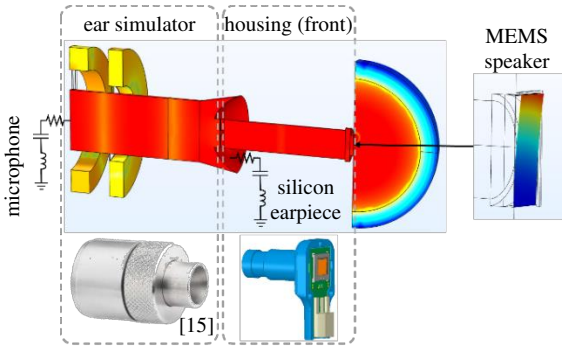

Figure 2. Schematic illustration of the simulation approach.

Figure 3 compares the simulated frequency responses of the previous generation with an active area of 4 × 4 mm^2 and triangular actuators and the new concept with a 2.4 × 2.4 mm^2 actuator. Due to the different sizes of the speakers, the curves are normalized to the active area. The new design outperforms the previous generation by approximately 10 dB_{SPL}/mm^2 at frequencies below 500 Hz. In other words, the active area of the new design concept can be reduced by roughly a factor of three while maintaining the SPL generated by the previous generation.

Generally, resonances are undesired in audio reproduction applications and need to be compensated for. Accordingly, loudspeakers with high damping and low quality factors are beneficial. Advantageously, increased damping is observed for the new concept, caused by thermo-viscous effects in the narrow gap between the bending actuator and the chip frame. Comparing the respective first speaker resonance (8.5 kHz for the previous generation (the resonance around 5 kHz is of acoustic nature) and 2.8 kHz for the new design), the resonance height relative to low frequencies is reduced by approximately 14 dB_{SPL}.

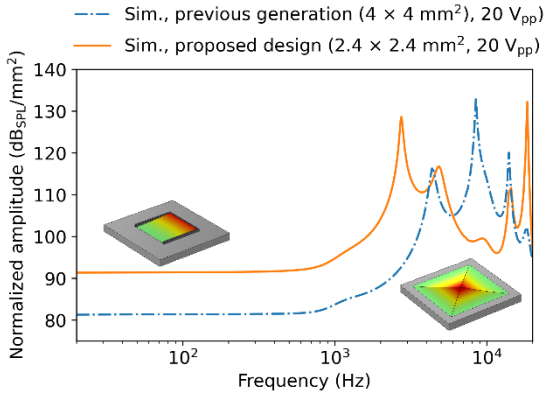

Figure 3. Comparison of the simulated, area-normalized frequency responses of the first- and second-generation MEMS speakers.

To experimentally validate the design concept, MEMS speakers with an active area of 2.4 × 2.4 mm^2 were fabricated. After the wafer processing, the speakers are singulated, assembled on PCB, mounted in a prototype earphone housing (Figure 4), and measured in an IEC 60318-4 ear simulator. A comprehensive description of the experimental setup can be found in [14].

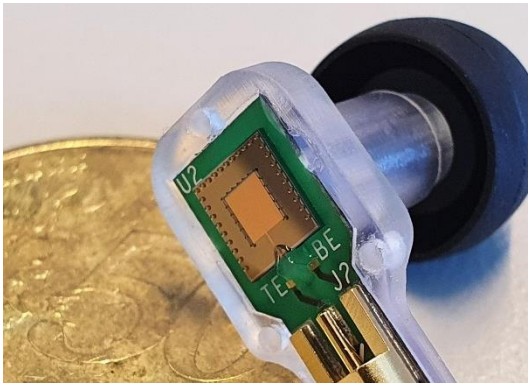

Figure 4. Partially assembled prototype earphones with a 2.4 × 2.4 mm^2 MEMS speaker.

Figure 5 shows the measured frequency responses for three different input signals. The non-equalized frequency response at 4 V_{pp} (black dotted curve) reaches an SPL of more than 87.5 dB_{SPL} for the entire reproduction range, except for a minimum of 83.9 dB_{SPL} around 12 kHz. It also reveals a mechanical resonance at 2.5 kHz, followed by several acoustic resonances caused by the prototype housing and the ear simulator. Figure 5 furthermore shows the frequency responses filtered to match the Harman target curve [16] at 4 V_{pp} (blue dash-dotted curve) and 20 V_{pp} (orange curve). For this, an FIR filter with a maximum order of 512 is implemented using the inverse transfer function overlaid with the target curve. For reference, the Harman curve at 4 V_{pp} is plotted (gray dashed curve). It is apparent that the measured frequency response reproduces the target curve well. Despite the small active area of 5.8 mm^2, a maximum SPL of 105 dB_{SPL} was reached at 30 Hz for 20 V_{pp}. This value surpasses typical SPL required for music or speech reproduction in in-ear applications, leaving room for use cases with higher SPL demands. For the latter case, the SPL are highly scalable, e.g., by varying the actuator design or integrating multiple actuators on one chip. Normalizing the measured data to the active area results in an SPL of 89.7 dB_{SPL}/mm^2 at 30 Hz and 20 V_{pp}. In comparison with the previous generation (80.3 dB_{SPL}/mm^2 at 30 Hz and 20 V_{pp}), an increase of 9.4 dB_{SPL}/mm^2 is observed in good agreement with the simulation study. This increase can be attributed to the higher area-averaged deflections of the rectangular actuators. Since the current and previous generations rely on the same piezoelectric material and film thickness, the higher area-normalized SPL furthermore indicates increased sensitivity due to the

almost three times lower capacitance of the smaller speaker.

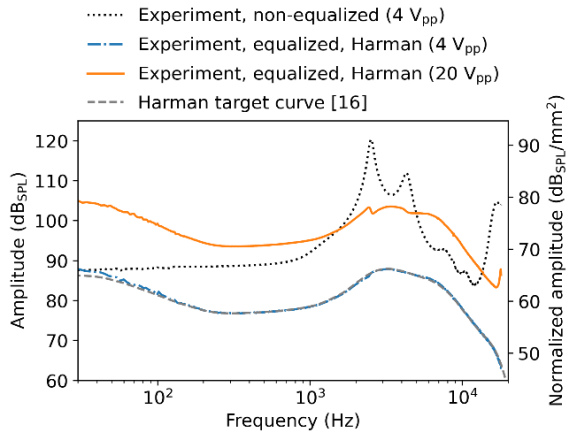

Figure 5. Measured frequency responses: non-equalized at 4 V_{pp} (black dotted curve) and equalized to match the Harman target curve [16] at 4 V_{pp} (blue dash-dotted curve) and 20 V_{pp} (orange curve).

4. NUMERICAL DESIGN STUDY

Besides the improved SPL, damping, and sensitivity, the new design furthermore provides higher design freedom. Due to the acoustic shield, no symmetric configuration of multiple actuators sealing each other is needed, as with the previous generation. Accordingly, the actuators and the acoustic behavior can be tailored much more freely. In the following, three example designs (Figure 6) are discussed and their acoustic performance is compared. The first variant is the previously introduced speaker with an active area of 2.4×2.4 mm^2. Additionally, a speaker with a 2.5×0.8 mm^2 actuator and an aspect ratio of three is considered as well as an array of three 2.5×0.8 mm^2 actuators. The latter two additionally feature a lower thickness of the passive poly-silicon layer of 10 µm.

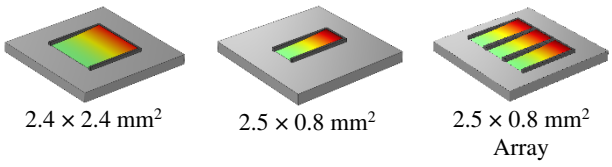

Figure 6. Overview of the simulated MEMS loudspeaker variants.

Figure 7 shows the comparison of the simulated frequency responses of the three variants. To compensate for the different speaker sizes, the frequency responses are normalized to the active area. Additionally, the driving voltage is chosen so that all speakers exhibit the same maximum deflection of 150 µm or, in other words, have an acoustic shield of similar height.

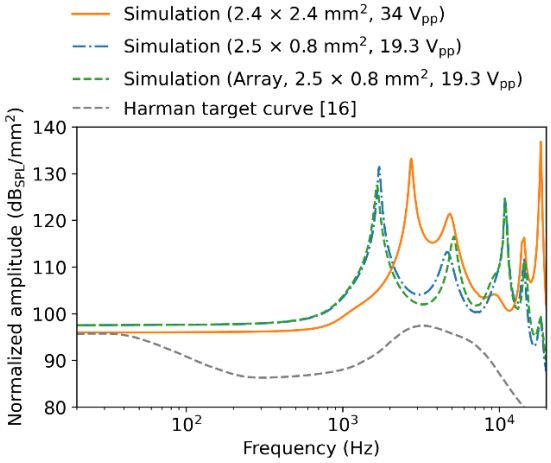

Figure 7. Comparison of the simulated, area-normalized frequency responses of the three MEMS speaker designs.

The 2.4×2.4 mm^2 and the 2.5×0.8 mm^2 MEMS speakers reach maximum normalized SPL of 95.9 dB_{SPL}/mm^2 and 97.5 dB_{SPL}/mm^2, respectively, at 20 Hz. Resulting from the lower thickness of the passive layer and the increased actuator length, the 2.5×0.8 mm^2 variant exhibits a slightly increased normalized SPL at approximately half the driving voltage, resulting in drastically increased sensitivity.
Furthermore, the speaker design also affects the position and height of the resonance peaks. Generally, a higher aspect ratio leads to an increased amount of gap volume per active area, which increases damping. As explained before, this is beneficial for sound reproduction applications. This effect is visible when comparing the heights of the respective first resonance peaks (Figure 7, heights relative to low frequencies: 2.4×2.4 mm^2 actuator: 37.3 dB_{SPL}; 2.5×0.8 mm^2 actuator: 33.9 dB_{SPL}). The array of three 2.5×0.8 mm^2 actuators imposes additional damping caused by the crosstalk of the actuators and resulting in a peak height of 29.9 dB_{SPL}. In addition, the resonance frequencies can be adjusted by varying the actuator length and thickness of the passive poly-silicon layer, e.g., to match the frequency of a possible maximum in the desired target curve. In this regard, the 2.4×2.4 mm^2 variant performs better than the other two when using the Harman target curve.

Additionally, total harmonic distortions (THD) are calculated using time domain simulations at 1 kHz and 94 dB_{SPL}. The simulation time is chosen so that after the transient excitation the deflection amplitudes reach the stationary regime. The THD is then calculated from a subsequent fast Fourier transform. The THD simulations presented here include geometric non-linearities and non-linearities caused by polarization effects of the PZT. Other material non-linearities and acoustic non-linearities are assumed to be negligible. To simplify the calculations, no deformation of the gap between the chip and the actuator (e.g. in terms of a moving mesh and remeshing) is modeled. Hence, the presented THD can be regarded as qualitative rather than quantitative values. The 2.4×2.4 mm^2 variant exhibits a THD of 0.39 % at 1 kHz and 94 dB_{SPL}. In comparison, the 2.5×0.8 mm^2 variant results in a THD of 0.41 %. At first sight, the latter variant performs slightly worse. However, to compensate for the lower active area, it requires significantly higher deflections to reach the same SPL of 94 dB_{SPL}. The array of three 2.5×0.8 mm^2 actuators compensates for this and results in a THD of 0.086 %, which is a drastic reduction compared to the square single actuator with a 2.4×2.4 mm^2 active area. Hence, considering both geometric non-linearities and polarization effects of the PZT, it can be concluded that the THD is smaller for higher aspect ratio speakers. In summary, the great design freedom of the new design concept enables a significant potential for optimizing the distortion of the MEMS speakers. More comprehensive simulations, e.g. including the deformation of the gap between the chip frame and actuator, and the comparison with THD measurements are subject of ongoing research.

5. SUMMARY AND OUTLOOK

In this paper, a novel concept for high SPL MEMS speakers has been discussed. The distinctive feature, an acoustic shield, allows for a mechanical decoupling between the actuator and chip while preventing any noticeable airflow through the gap. In combination with rectangular actuator geometries, high area-averaged deflections and SPL are achieved. Experimental studies on a prototype in-ear demonstrator with an emitting area of 5.8 mm^2 result in an SPL of 105 dB_{SPL} at 30 Hz for a driving voltage of 20 V_{pp}, measured in an IEC 60318-4 artificial ear simulator. Compared to the previous generation, the new concept allows for an increase of approximately 10 dB_{SPL}/mm^2 or a reduction of the active area by roughly a factor of three compared to the old design. Thereby, the full audible range from 20 Hz to 20 kHz is covered. Additionally, higher sensitivity and higher damping are observed for the new design. The latter is especially beneficial in terms of signal processing and distortion.

The new design furthermore enables significantly higher design freedom. The influence of several parameters, such as the cantilever aspect ratio and the thickness of the passive layer, on the acoustic performance was investigated in a numerical study. In summary, previously unused potential and large freedom for tailoring the designs and their acoustic output in terms of SPL, resonance frequency, and damping towards specific applications were discovered. Additionally, by varying the design or forming arrays of multiple actuators, the overall SPL is highly scalable.

In summary, with a broadband SPL of 89.7 dB_{SPL}/mm^2, the new concept promises a high miniaturization potential and surpasses the state of the art of MEMS loudspeakers [1] despite using lower driving voltages. Besides more comprehensive experimental validation and characterization, ongoing research focuses on further miniaturization and improvement of the MEMS speaker design, e.g., regarding damping. Additionally, the new piezoelectric material aluminum scandium nitride is being integrated, allowing for even higher sensitivity, SPL, and linearity, as well as a fully CMOS-compatible process.

6. REFERENCES

[1] H. Wang, Y. Ma, Q. Zheng, K. Cao, Y. Lu, and H. Xie, “Review of Recent Development of MEMS Speakers”, *Micromachines*, vol. 12, p. 1257, 2021.

[2] M. A. Harradine, T. S. Birch, J. C. Stevens, and C. Shearwood, “A micro-machined loudspeaker for the hearing impaired”, in *Tech. Dig. Papers of the Transducers’97 Conference*, (Chicago, USA), p. 429, 1997.

[3] M.-C. Cheng, W.-S. Huang, and S. R.-S. Huang, “A silicon microspeaker for hearing instruments”, *J. Micromech. Microeng.*, vol. 14, p. 859, 2004.

[4] S.-S. Je, F. Rivas, R. E. Diaz, J. Kwon, J. Kim, B. Bakkaloglu, S. Kiaei, and J. Chae, “A compact and low-cost MEMS loudspeaker for digital hearing aids”, *Trans. Biomed. Circuits Syst.*, vol. 3, p. 348, 2009.

[5] Y. C. Chen, and Y. T. Cheng, “A low-power milliwatt electromagnetic microspeaker using a PDMS membrane for hearing aids applications”, in

Proc. of the 24th Int. Conf. on MEMS, (Cancun, Mexico), p. 1213, 2011.

[6] B. Y. Majlis, G. Sugandi, and M. M. Noor, "Compact electrodynamic MEMS-speaker", in *Proc. of the Semiconductor Technology Int. Conf.*, (Shanghai, China), 2017.

[7] S. S. Lee, R. P. Ried, and R. M. White, "Piezoelectric cantilever microphone and microspeaker", *J. Microelectromech. Syst.*, vol. 5, p. 238, 1996.

[8] F. Stoppel, A. Männchen, F. Niekiel, D. Beer, T. Giese, and B. Wagner, "New integrated full-range MEMS speaker for in-ear applications", in *Proc. of the 31st Int. Conf. on MEMS*, (Belfast, UK), p. 1068 2018.

[9] T. S. Albach, P. Horn, A. Sutor, and R. Lerch, "Sound generation using a magnetostrictive microactuator", *J. Appl. Phys.*, vol. 109, 2011.

[10] H. Conrad, H. Schenk, B. Kaiser, S. Langa, M. Gaudet, K. Schimmanz, M. Stolz, and M. Lenz, "A small-gap electrostatic micro-actuator for large deflections", *Nat. Commun.*, vol. 5, p. 10078, (2015 Dec.).

[11] B. Kaiser, S. Langa, L. Ehrig, M. Stolz, H. Schenk, H. Conrad, H. Schenk, K. Schimmanz, and D. Schuffenhauer, "Concept and proof for an all-silicon MEMS micro speaker utilizing air chambers", *Microsyst. Nanoeng.*, vol. 5, 2019.

[12] USound, "Achlelous UT-P2020 | Datasheet", Feb. 2021 [Revised June 2021].

[13] xMEMS, "Products", 5 Dec. 2022, https://xmems.com/products/.

[14] F. Stoppel, J. Fankhänel, T. Giese, C. Eisermann, I. Pieper, D. Kaden, L. Castellanos, and S. Grünzig, "Highly miniaturized in-ear MEMS loudspeaker featuring high SPL", in *Proc. of the 22nd Int. Conf. on Solid-State Sensors, Actuators and Microsystems 2023 (Transducers 2023)*, (Kyoto, Japan), 2023.

[15] GRAS, www.grasacoustics.com, accessed 14 Sept. 2022.

[16] S. Olive, and T. Welti, "Factors That Influence Listeners' Preferred Bass and Treble Levels in Headphones", in *Proc. of the 139th AES Convention*, (New York City, USA), p. 9382, 2015.