

ARTICLE

Open Access

The push-pull principle: an electrostatic actuator concept for low distortion acoustic transducers

Bert Kaiser¹✉, Hermann A. G. Schenk², Lutz Ehrig², Franziska Wall¹, Jorge M. Monsalve¹, Sergiu Langa¹, Michael Stolz^{1,3}, Anton Melnikov¹, Holger Conrad², David Schuffenhauer¹ and Harald Schenk^{1,3}

Abstract

Electrostatic actuators are of particular interest for microsystems (MEMS), and in particular for MEMS audio transducers for use in advanced true wireless applications. They are attractive because of their typically low electrical capacitance and because they can be fabricated from materials that are compatible with standard complementary metal-oxide semiconductor (CMOS) technology. For high audio performance and in particular low harmonic distortion (THD) the implementation of the push-pull principle provides strong benefits. With an arrangement of three electrodes in a conjunct moving configuration on a beam, we demonstrate here for the first time a balanced bending actuator incarnating the push-pull principle operating at low voltages. Our first design already exhibits a harmonic distortion as low as 1.2% at 79 dB using a signal voltage of only $6V_p$ and a constant voltage of only $\pm 10V_{dc}$ in a standard acoustic measurement setup. Thus, exceeding our previously reported approach in all three key performance indications at the same time. We expect that our novel electrode configurations will stimulate innovative electrostatic actuator developments for a broad range of applications. In this paper we report the basic theory, the fabrication and the performance of our novel actuator design acting as an audio transducer.

Introduction

Highly linear MEMS μ Speaker technology is undoubtedly a current focal point of MEMS research. This interest is predominantly fueled by the desire for enhanced internet connectivity provided by true wireless headsets. And with the prospect of MEMS μ Speakers which draw only minute reactive currents at high sound pressure levels, innovative and cost-effective solutions – even for hearing aids – come within reach.

Quite impressive MEMS-based speakers have been reported in the literature. They are based on different transduction mechanisms, various materials and fabrications processes respectively.

One way of comparing the devices is their footprint, i.e. their geometric dimensions, in accordance with the attained sound pressure level in specific acoustic setups. For μ Speakers with dimensions in the sub-centimeter domain, a closed cavity approximating the ear canal (1.26 cm^3 , according to IEC 60318-4 standard) is proposed as the acoustic setup. Specifically, the harmonic distortion should be as low as possible at a specific sound pressure. Furthermore, the capacitive load of a device is a profound way to estimate the relative power consumption of the reported approaches. A capacitance of more than a few nano-Farads gives rise to reactive peak currents, which are unbearable for batteries typically used in modern true wireless headsets, at the required sound pressure levels (SPL) of 125 dB or above in the frequency domain of human speech.

MEMS μ Speaker approaches rely on the piezoelectric transduction mechanism^{1–8}, the electrostatic transduction mechanism^{9–15}, the magnetic transduction

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

¹Fraunhofer Institute for Photonic Microsystems IPMS, Dresden, Germany

²Arioso Systems GmbH; since Sep. 1st 2022: BOSCH Sensortec GmbH, Dresden, Germany

Full list of author information is available at the end of the article

© The Author(s) 2022



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/>.

mechanism^{16–20} and the thermoacoustic transduction mechanism^{21,22}. A thorough review paper about micro speakers was recently published by Wang et al.²³. A few publications shall be highlighted in the following.

Recently Wang et al. described a MEMS μ Speaker based on thin lead zirconium titanate ceramic⁸. The device measures $6.7 \times 6.7 \text{ mm}^2$ and has a circular membrane of about 4.2 mm diameter with a deflection direction, which is out-of-plane. The reported SPL reached 119 dB at a distance of 10 mm and for a frequency of about 9 kHz. This translates to approximately 124 dB in a 1.26 cm^3 closed cavity. An estimation from the graph reveals a value of 64 dB at a referential frequency of 500 Hz with the same assumed cavity. No harmonic distortion values are reported. Taking the given geometric dimensions, a capacitive load of at least $1.6 \mu\text{F}$ is estimated for the active area (4 mm diameter, lead zirconium titanate permittivity $\epsilon_r = 300$, layer thickness $5 \mu\text{m}$).

A push-pull configuration with static electrodes was reported by Zhou et al. for a speaker based on a graphene diaphragm¹¹. The speaker had two static electrodes and the diaphragm acted as a shuttle which translated the Coulomb forces into sound pressure waves. The drive used was a $100 \text{ V}_{\text{dc}}$ bias and a 10 V_{p} signal while at the same time very low power consumption, below $1 \mu\text{W}$, and very high power efficiency were achieved. No absolute values were reported for sound pressure or harmonic distortion.

Kim et al. reported on implementing the push-pull principle in a μ Speaker¹⁰. Sound pressure levels of 113.4 dB and 98.8 dB were attained at 1 cm distance using two and one electrodes respectively. This translates to about 121 dB and 96 dB in a closed cavity of 1.26 cm^3 volume. No values for harmonic distortion or electrical capacitance were reported.

Sugimoto et al.²⁴ reported an acoustic transducer based on two dielectric elastomer films which specifically took advantage of the push-pull principle to suppress harmonic distortion. A decline in harmonic distortion of 10 dB (factor ≈ 3) was described for push-pull driving over one-side driving.

Similarly Hsin-Yuan Chiang and Yu-Hsi Huang reported the experimental modeling and application of push-pull electrostatic speakers²⁵. Their devices are not in the micro-sized speaker domain because they incorporate a circular diaphragm of at least 60 mm in diameter. Specifically, the authors claim an outstanding sound quality for their push-pull speakers.

In our own work²⁶ we reported on a MEMS μ Speaker based on electrostatic bending actuators deflecting laterally in-plane. A sound pressure of 69 dB in a 1.26 cm^3 cavity was reported with a harmonic distortion below 4.4% at a driving voltage of 40 V_{dc} plus 5 V_{p} (pull-in voltage $U_{\text{pi}} = 49 \text{ V}$). The device's active area measured 9.3 mm^2 and it exhibited an electric capacitance of 65 pF.

This literature overview reveals several works addressed to MEMS μ Speaker. Some of them implement configurations based on the push-pull concept. These approaches are assumed to be either large and/or require quite high driving voltages because of the use of static electrodes. The electrostatic push-pull actuator concept offers strong benefits²⁷. Especially highest linearity achievable regarding voltage-deflection relation is a dominant quality of such actuators. Nevertheless, the implementation of an electrostatic push-pull actuator configuration remained challenging due to the arrangement of the electrodes and their respective relative movement with respect to each other. Likewise, this affects their voltage-deflection relation as well as the device design in which they are implemented in. Up to now, the general drawback of electrostatic actuators has been the necessity to operate them at comparably high voltages. This applies in particular to electrostatic actuators in a push-pull configuration, which in case of μ Speakers is required to achieve low harmonic distortion. Ultimately, the high voltage requirement is linked to the classical stator-shuttle electrode configuration and the limits posed by the pull-in phenomenon - a singularity.

Singularities attract many engineers' interest. In the case of electrostatically driven MEMS μ Speakers, such critical behavior is encountered in extremely miniaturized, high fidelity audio reproduction systems that are supposed to operate at high sound pressure levels. Typical MEMS μ Speaker designs use an elastic force to balance an electrostatic force. These two forces are quite different in nature. This asymmetry creates a singularity, called pull-in, when the input signal generates an electrostatic force strong enough to annihilate the relevant effective stiffness. For state-of-the-art μ Speaker designs, the demand for high audio reproduction quality in portable applications implies the need to operate them close to the pull-in. Only close to the pull-in is the slope of the deflection curve, i.e. the sensitivity of the transducer, large enough to produce high sound pressure levels at signal voltages sufficiently low for portable applications. Additionally, only when the μ Speaker is driven at low signal amplitudes the total harmonic distortion (THD) is acceptably low. The THD is defined as

$$THD = 100 \sqrt{1 - \frac{K_1^2}{\sum K_n^2}}, \quad (1)$$

with K_n being the harmonic components of the sounds generated.

For practical use, operating an electrostatic transducer close to pull-in is not desirable. Unless special protective measures have been taken, such a modus operandi is prone to catastrophic failure, caused, e.g., by minute mechanical shocks.

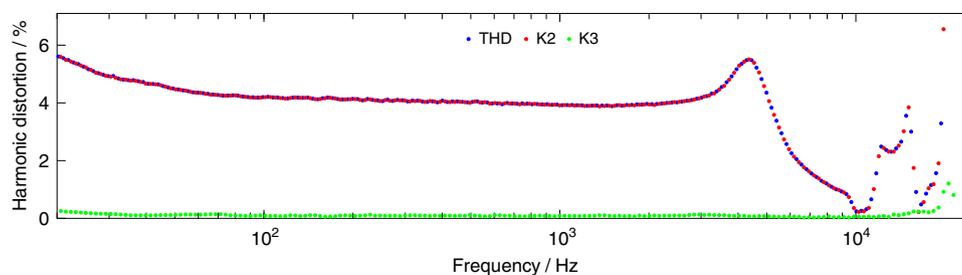


Fig. 1 μ Speaker with asymmetric actuator. Acoustically measured THD for a μ Speaker employing solely-pull bending actuators. The THD is dominated by the 2nd harmonic K2 due to the quadratic relation between signal voltage and displacement²⁶

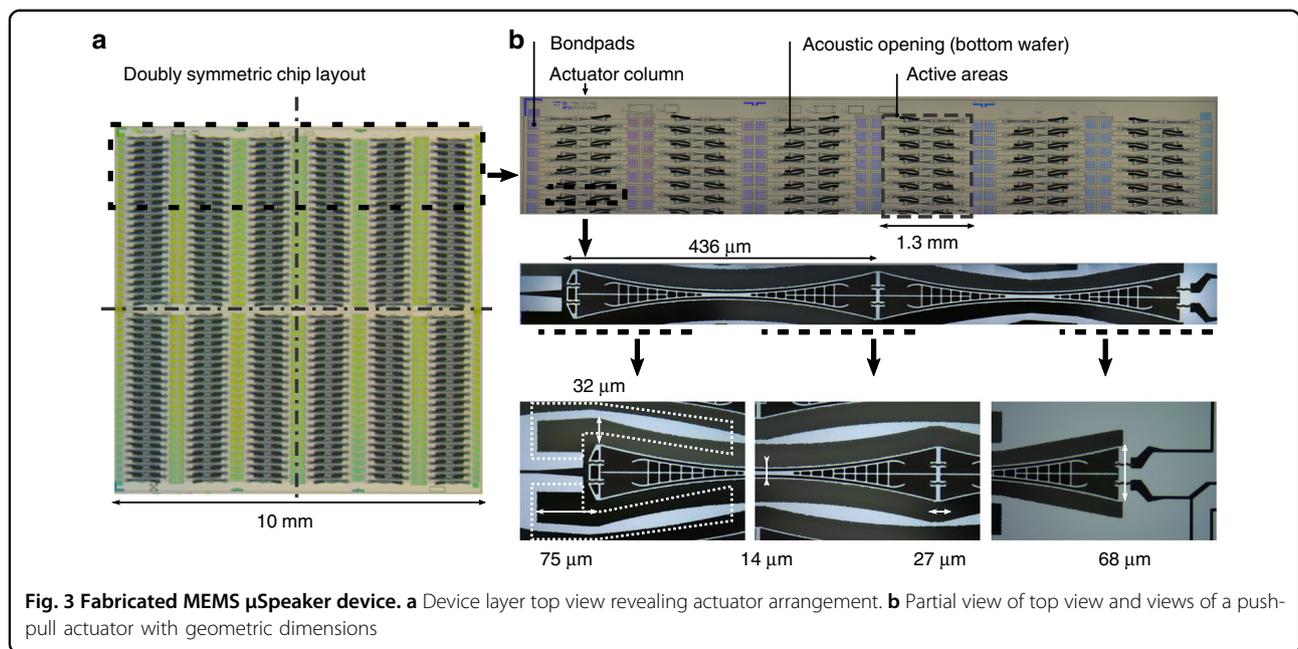
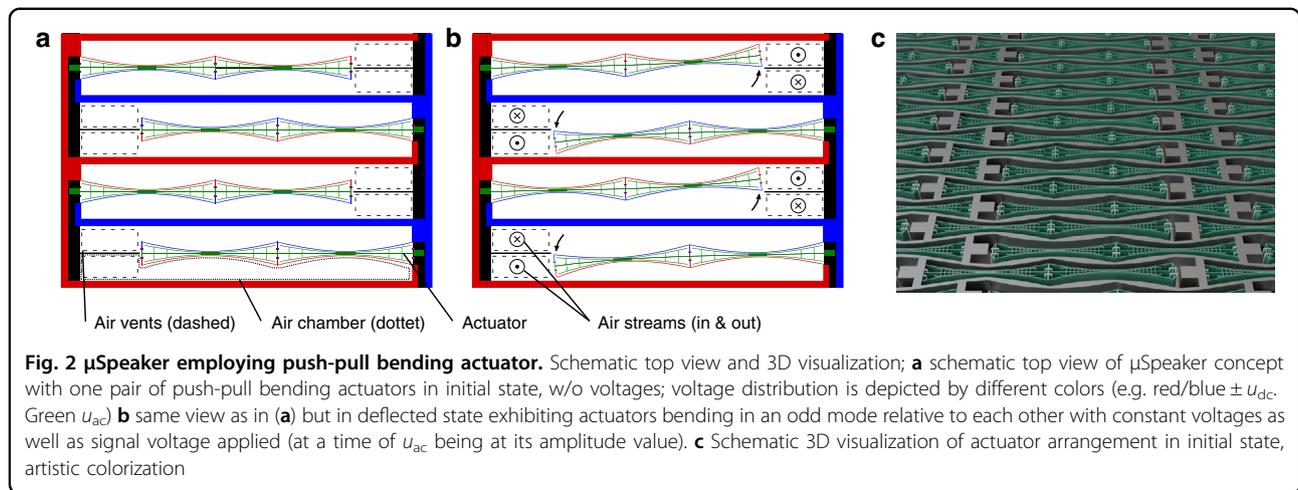
The other relevant implication of the electroelastic asymmetry is the generation of a pronounced second harmonic distortion (K2). That K2 tends to keep the THD well above an acceptable threshold of about 1% (Fig. 1) for μ Speaker employing solely-pull bending actuators. Compensating for such a deficiency by means of advanced signal processing is not desirable in portable devices that work off a tight battery budget. Instead the approach is to limit the share of K2 on the THD and go down to the third harmonic distortion (K3) fraction, which is significantly lower.

The question arises of how to design an electrostatic MEMS transducer, suitable for broad frequency range operation which have low harmonic distortion and do not need to be operated close to the pull-in singularity. The classical answer to this question is to use a push-pull configuration, placing a shuttle electrode between two stator electrodes²⁸. However, due to the required large stator distance, such push-pull designs, even when implemented as MEMS, operate at high voltages that are impractical for portable applications, and in particular for in-ear use. In previous papers^{26,29}, we reported an asymmetric design that solves the large voltage issue.

Here we show for the first time that the stator-shuttle configuration can in fact be replaced by a design that integrates all three electrodes into the moving parts of an electrostatic bending actuator. This means the electrode spacings or “gaps” can be kept tight, allowing voltages suitable for small portable devices and at the same time achieving a much-improved linearity of the resulting audio reproduction without sacrificing the maximum attainable sound pressure levels. In this work we describe such a concept for an electrostatic actuator based on MEMS technology. First, we introduce the transducer design and secondly, we introduce the bending actuators principle, theoretical description and fabrication. As a proof, we demonstrate a μ Speaker showing low THD levels as well as low voltage requirements while at the same time maintaining sufficient SPL levels.

Transducer design

The core component of the low distortion μ Speakers presented in this work is the balanced bending actuator based on the push-pull principle. The MEMS μ Speaker concept is based on a stack of wafers pile up to create a device layer sandwiched between a bottom and a cover wafer. A plurality of actuators and air chambers are preferably arranged densely and actuated simultaneously for sound reproduction. However, the push-pull concept has several ramifications for the design of the acoustic transducer. The bending actuator is clamped on one side only. One-side clamping is favored due to linearity concerns over the Duffing effect created by stress stiffening in two-side clamped beam designs. The push-pull actuators operate in single and individual air chambers. These air chambers are defined by the actuators themselves and additional sidewalls between each of the actuators. Fig. 2a shows the layout concept. The actuators are arranged in rows and columns. In each column the actuators alternate with respect to their clamped side. If actuated, the push-pull actuators bend towards each other or away from each other. Thus, two neighboring actuators form a pair, canceling out the resulting inertial forces. Close to its free end, where the maximum deflection and air displacement are achieved, each actuator has its own pair of air vents. The actuator tip is separated from the opposing silicon surface by a small gap. Like the clearance, this gap is small enough not to cause acoustic leakage and ensures the movability of the actuators. It should be noted that the neighboring actuators are shifted by about half a cell pitch (Fig. 2a), to allow for denser packing and higher area utilization. The sidewalls between the actuators transfer the electrical potential between the alternating clamping sides of the beams within the chip. Technological clean-room efforts, as well as wiring expenses, are greatly reduced because crossing potentials or external bridges or wires are avoided. While this is true for both constant voltages, which can be distributed within the device layer in the way described, it is necessary to integrate the bottom wafer as an electrical distribution layer for the audio signal. This is done by introducing through-silicon vias



electrically connecting the device layer and the bottom wafer specifically at the middle electrode. With these means of potential distribution, no external wiring is necessary except one general connection for each potential. Additionally, it allows for the odd mode movement of the neighboring actuators and also means the air chambers are free of an electrostatic field. This is highly desired in order not to generate any electrostatic air filtering effects. When a signal voltage is applied, the actuators bend towards or away from each other depending on the sign of the signal (Fig. 2b).

The description given in the previous paragraphs mainly focuses on a lateral view of the actuator and chip design, i.e. the layout composition parallel to the MEMS surface. The actual device is then an extrusion of that layout into

the chip’s volume, unleashing the uniqueness of the underlying μ Speaker concept.

Figure 3a shows the top view of the 75 μ m thick device layer of a fabricated MEMS speaker. The whole chip comprises 276 actuator pairs arranged in two groups. Both groups are of the same size and design, and actually mirror each other with respect to the center of the chip (horizontal dashed-dotted line). These groups are defined by the wiring scheme within the device plane and are individually addressable. The reason for having those two groups is purely down to testing capabilities.

Each actuator has a length of 872 μ m and maximum and minimum widths of about 68 μ m and 14 μ m respectively. A vertical pull-in between actuator and cover/bottom

wafer for the maximum driving voltage of 36 V is safely avoided. Additional holes have been placed within the center electrode. These holes do not affect the electro-mechanical behavior but do lower the mass of the actuator. The resonance frequency was calculated as about 5 kHz for the first mode, which is the desired in-plane swiping movement.

The actuators are placed pairwise in an active area (pitch) of 200 μm (vertically, rows) by 1646 μm (horizontally, columns). The vertical pitch is designed in such a way that the maximum deflection in resonance still gives enough room for the air to be exerted from the chambers through the vents. Specifically, a maximum quasistatic tip deflection of the actuators of about 8 μm peak to peak with a excitation in resonance and a remaining distance each of around factor two was assumed. This leads to a distance of 32 μm between actuator edge at the tip and closest wall (Fig. 3b). The air vents were designed to about 40 \times 260 μm in width and length respectively. The relatively large air vents overlap the sidewalls partially, making it possible that air is exerted from two neighboring actuators (see dotted indication in Fig. 3b). The chip design incorporates two individual bond pads (170 \times 70 μm^2) for the constant voltages in each group of actuators, i.e. four in total. The third potential is to be contacted via the bottom wafer as a backside contact or via individual bond pads at each actuator. The upper frame of the MEMS chip shows additional structures relevant for technological process monitoring purposes only. The whole chip measures 10 \times 10 mm^2 including the chip

frame and has a total thickness of about 1.1 mm. The device layer has a thickness of 75 μm and the bottom and cover wafers are 625 μm and 400 μm thick respectively. The separation between the device layer and the bottom layer is maintained by the 1 μm buried oxide layer. The cover wafer and device wafer showed a separation (clearance) of the exact same size.

Theory

Push-pull-concept

The push-pull concept for electrostatic bending actuators is shown in Fig. 4. It shows three independently addressable electrodes placed next to each other in close proximity. In fact, the distance is defined by the two electrostatic gaps. These gaps can be kept small in order to allow for small driving voltages. The minimum width is most likely defined by technological constraints. All three electrodes are mechanically linked together and form a group of simultaneously moving mechanical entities. The overall appearance of the group is like a beam because of the elongated geometric dimensions of the electrodes. The beam and group of electrodes are symmetrical with respect to their centered longitudinal axis. This symmetrical layout of the beam allows the non-linear electrostatic forces in the gaps at both sides of the neutral fiber to be counteracted. The actual shape of the actuator is less similar to a parallel plate and more likely to be recognized as suspension bridge arrangement. The reason for this is its mechanical susceptibility to bending in combination with the general desire to have the lowest area consumption possible.

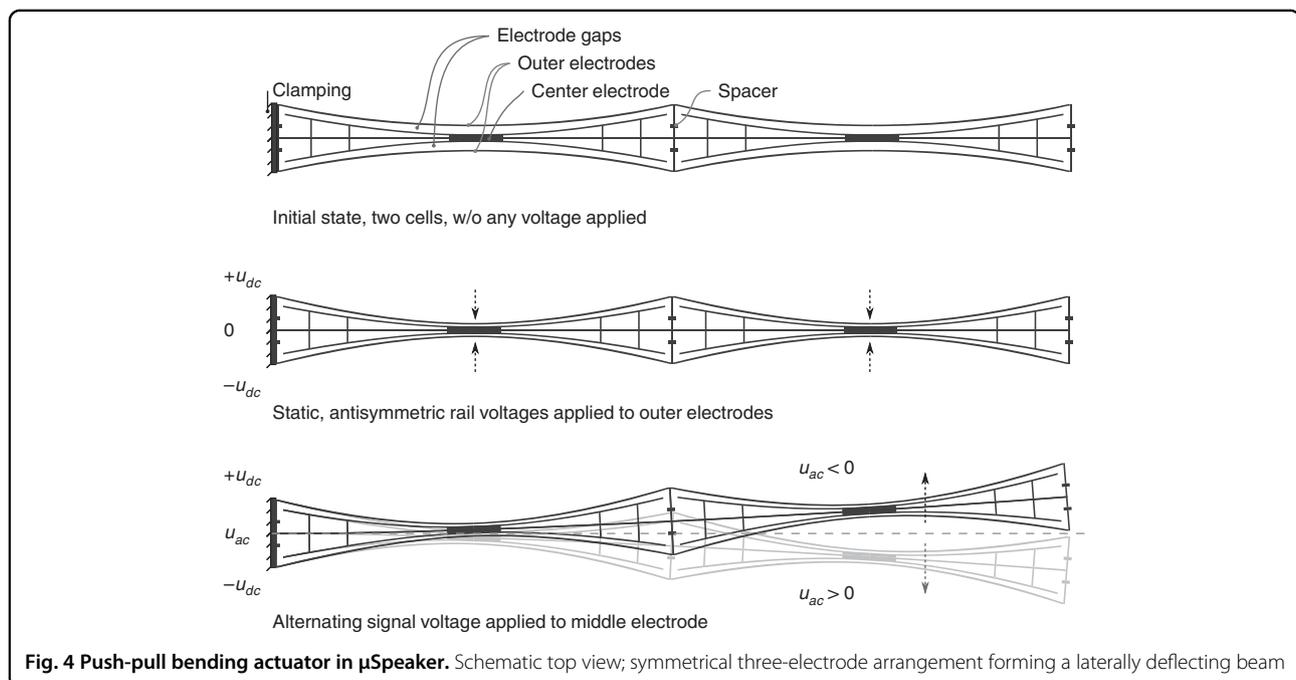
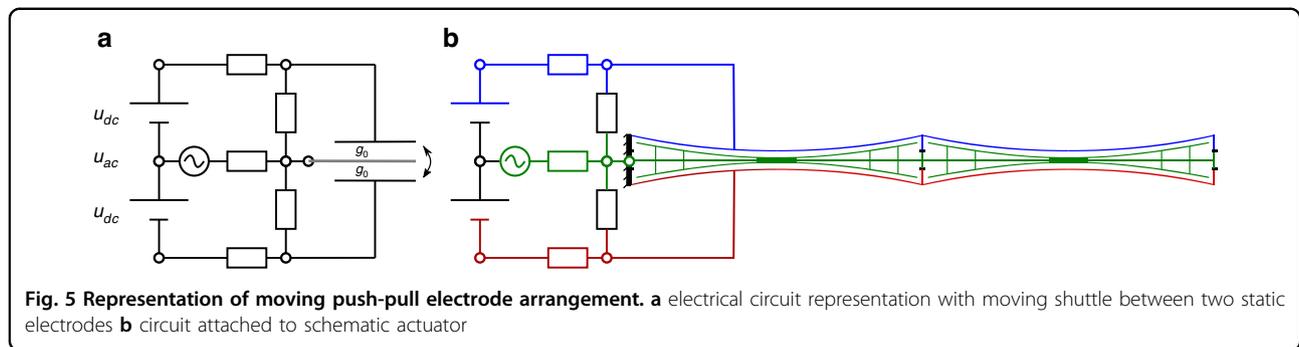


Fig. 4 Push-pull bending actuator in $\mu\text{Speaker}$. Schematic top view; symmetrical three-electrode arrangement forming a laterally deflecting beam



Bending of the beam-like actuator is established by incorporating the nanoscopic electrostatic drive (NED) principle²⁹. Coulomb forces in the gap are transferred into uniaxial stress by the shape of the electrodes, thus bending the actuator. Constant voltages u_{dc} of the same amplitude (with respect to the middle electrode) but with opposing signs are applied to the actuator's outer electrodes. In this state, no actual deflection in a lateral sense takes place. The audio signal u_{ac} is then applied to the middle electrode, thus unbalancing the forces on both sides of the actuator. In this state the actuator is actively bending in one of the two directions depending on the sign and amplitude of the audio signal. The direction is determined by which gap exhibits the greater voltage difference of $u_{dc} + u_{ac}$. In other words, one gap has electrostatic forces pulling more strongly in its direction - "pull" - and the opposing gap has electrostatic forces pulling less strongly in the opposing direction - "push". Interestingly, the actuator behaves like a push-pull system without pushing or repelling forces being present. Consequently, a linearized deflection is generated.

The design process can be understood as follows. The variable of merit is the achievable deflection for a given set of driving voltages. Respecting technological capabilities, which define the minimum electrode thicknesses as well as the minimum gap distances, the deflection will be maximized by using the thinnest electrodes and the smallest gaps possible. Assuming a fixed and already minimized stiffness, the mass of the actuators, i.e. their cross-sectional area, should be minimized for a resonance frequency suitable for in-ear applications. Consequently, the aforementioned suspension-like structure is superior to a reciprocal dome-like shape, due to less confined area by the actuators contour. The actual design details were determined using finite element analysis in a parameter sweep, in order to account for all the rather complex design rules. More details on the finite element analysis are given in section "Material and methods".

Some basic equations explaining the essence of the acoustic performance of our push-pull actuators are discussed in the following section.

Theoretical description

The classical electrostatic push-pull principle is depicted in Fig. 5a. It comprises a shuttle electrode that moves between two stator electrodes which are held at constant voltages in a two-loop circuit topology. When used in audio transducers, a classical electrostatic push-pull operation requires the use of very high drive voltages, typically of the order of 1000 V³⁰. Combining the push-pull principle with the NED concept as outlined in the preceding section allows, in contrast, for a low voltage operation, as is required for true wireless in-ear applications (Fig. 5b).

The most basic, though very effective, method modeling a NED-based push-pull actuator, is introducing a mechanical lever factor θ into the classical equation for an electrostatic push-pull actuator. In its dimensionless form this equation reads

$$\frac{d^2 a(\tau)}{d\tau^2} + \frac{1}{Q} \frac{da(\tau)}{d\tau} + k a(\tau) = \theta \left[\left(\frac{u_{dc} + u_s(\tau)}{1 - a(\tau)} \right)^2 - \left(\frac{u_{dc} - u_s(\tau)}{1 + a(\tau)} \right)^2 \right]. \quad (2)$$

Here $a(\tau)$ is the dimensionless stroke of the actuator²⁸ normalized to the initial electrode spacing g_0 . The dimensionless quantities τ , Q , u_{dc} and $u_s(\tau)$ are the dimensionless time, the quality factor, the dimensionless constant voltage and the dimensionless signal voltage respectively. The dimensionless stiffness k is typically normalized to unity. We keep it here to facilitate the interpretation of the formulae below. While the definition of the stroke is entirely obvious in the case of a rigid-plate-like shuttle, its meaning is less clear in the case of an elastic electrode that changes its shape under the influence of the Coulomb force. However, this dynamic variable can be identified, in a very well-defined sense, with the maximum deflection of a Coulomb-actuated Euler-Bernoulli beam, as used here^{31–33}.

The deflection characteristics in the case of a constant signal voltage are depicted in Fig. 8b. The domain of stable device operation extends between the two bifurcation

points where the slope of the deflection curve reaches infinity (pull-in). The slope of the deflection curve at the origin can be tuned by the value of the constant voltage.

As can be expected from the shape of this curve, the push-pull principle allows for device operation with the signal voltage oscillating about zero and with low harmonic distortion. This is in sharp contrast to the asymmetric electrostatic drive^{26,31}, where a working point close to the pull-in must be selected in order to minimize harmonic distortions. To analyze this claim in greater depth, we observe that according to Eq. (2) the direction of the stroke is reversed upon changing the sign of the signal voltage,

$$u_s(\tau) \rightarrow -u_s(\tau) \Rightarrow a(\tau) \rightarrow -a(\tau). \quad (3)$$

The consequences of this inversion symmetry can conveniently be understood by a deformation (homotopy) argument. To this end, we select a signal voltage oscillating around zero,

$$u_s(\tau) = u_{ac} \cos(\omega\tau), \quad (4)$$

and expand the stroke in terms of the signal amplitude,

$$a(\tau) = \sum_{n=0}^{\infty} a_n(\tau) u_{ac}^n. \quad (5)$$

From the inversion symmetry we can immediately infer that all coefficient functions with even indices vanish,

$$a_{2n}(\tau) = 0 \quad n \in \mathbf{N}. \quad (6)$$

Inserting Eq. (5) into Eq. (2), expanding the non-linear terms in u_{ac} and collecting all terms of a given power in u_{ac} , we find that the coefficient functions $a_n(\tau)$ satisfy the following deformation equations:

$$\frac{d^2}{d\tau^2} a_n(\tau) + \frac{1}{Q} \frac{d}{d\tau} a_n(\tau) + (k - 4u_{dc}^2) a_n(\tau) = f_n(\tau). \quad (7)$$

$$\begin{aligned} K_1 &= 4u_{dc}u_{ac} \left| \frac{Q}{kQ+\omega(i-Q\omega)-4Qu_{dc}^2} + \frac{Q^2((kQ+\omega(i-Q\omega))(3kQ+\omega(i+3Q\omega)))+4Q(3kQ+\omega(i-3Q\omega))}{(kQ+\omega(i-Q\omega)-4Qu_{dc}^2)(kQ+\omega(i+Q\omega)-4Qu_{dc}^2)} u_{ac}^2 \right|, \\ K_2 &= 0, \\ K_3 &= 4u_{dc}u_{ac}^3 \left| \frac{Q^2(kQ+\omega(i-Q\omega))(kQ+\omega(i-Q\omega)-4Qu_{dc}^2)}{(kQ+3\omega(i-3Q\omega)-4Qu_{dc}^2)(kQ+\omega(i-Q\omega)-4Qu_{dc}^2)^3} \right|. \end{aligned} \quad (10)$$

The driving forces $f_{2n}(\tau)$ with even indices vanish, as expected,

$$f_{2n}(\tau) = 0 \quad n \in \mathbf{N}. \quad (8)$$

The driving forces for odd indices are computed to be,

$$\begin{aligned} f_1(\tau) &= 4u_{dc} \cos(\omega\tau), \\ f_3(\tau) &= 4\cos^2(\omega\tau)a_1(\tau) + 12u_{dc} \cos(\omega\tau)a_1(\tau)^2 + 8u_{dc}^2 a_1(\tau)^3, \\ f_5(\tau) &= 8\cos^2(\omega\tau)a_1(\tau)^3 + 20u_{dc} \cos(\omega\tau)a_1(\tau)^4 + 12u_{dc}^2 a_1(\tau)^5 \\ &\quad + 24u_{dc}^2 a_1(\tau)^2 a_3(\tau) + 24u_{dc} \cos(\omega\tau)a_1(\tau)a_3(\tau) \\ &\quad + 4\cos^2(\omega\tau)a_3(\tau), \\ &\text{etc..} \end{aligned} \quad (9)$$

This hierarchy of equations can be solved recursively, beginning at $a_1(\tau)$ and progressing from one index to the next, higher index. Note that the force term at any level of the recursion only depends on coefficient functions already known from previous steps. At each step only a linear equation must be solved, which can be readily done by purely algebraic means, based on the usual Laplace transformation argument. Note in particular that $a_1(\tau)$ only contains the first harmonic. Using this result we are able to compute the explicit form of the driving force $f_3(\tau)$. Reducing the involved trigonometric polynomials to simple harmonics, we find that this driving force has only odd harmonics components up to 3rd order. This implies that $a_3(\tau)$ also only contains odd harmonics up to 3rd order. The argument can be repeated for any higher order, proving the claim that a perfectly balanced push-pull actuator entirely suppresses all even harmonic distortions, and in particular the second harmonic distortion (K2). Therefore, a perfectly balanced push-pull actuator only produces odd harmonics, predominantly the third harmonic (K3), which should lead to a substantially better performance compared to asymmetric driving schemes²⁸.

The procedure outlined above can easily be used to actually compute the THD, as a power series in u_{ac} . The first three harmonic components K_n of the THD, are given for a perfectly balanced actuator up to order $O[u_{ac}]^5$ and for $\theta = 1$ by

The results for non-trivial θ are obtained by scaling all voltages by a factor of $\sqrt{\theta}$. These formulae nicely fit numerical simulations. More importantly, the efficient algorithm sketched here serves well as a solid base for analyzing the impact

of imperfect balancing. As is well known from the theory of Volterra series^{34–36}, the radius of convergence of Eq. (5) extends to the point where the magnitude of the signal amplitude causes a dynamic instability. This situation is reached when the frequency response curve develops a Duffing-type multivaluedness^{31,37}. We will not dwell on the details of such analysis here, but we share the results in additional publications^{32,38}. For comparison with our experimental findings, we also need to add a suitable acoustic load. This is done in the network model used below.

Fabrication

The linear bending actuators were fabricated on a single chip design. The basic wafer material was bonded silicon on insulator wafers with a phosphor dopant concentration of 10^{18} cm^{-3} for the device layer, resulting in a specific electrical resistance of 0.01–0.025 $\Omega \text{ cm}$. The main process step for the bulk micro-machining was deep reactive ion etching. Layer deposition for trench filling as well as cover wafer bonding complete the technological processes (Fig. 6). A 200 mm diameter wafer had 197 devices arranged on it. Details of the technology and processes used are supplied in the following.

The linear bending actuators were fabricated starting with bare bonded silicon on an insulator wafer. The first step was the creation of the through-silicon vias. The about 2.4 μm wide holes were etched by deep reactive ion etching (DRIE), subsequently followed by an etch step to remove the buried oxide otherwise acting as an electrical insulator. The holes or trenches were afterwards filled with polysilicon using low

pressure chemical vapor deposition. The deposition was split into two steps of about 1000 nm deposition layer thickness. The intermediate doping of the polysilicon was achieved by applying an atmosphere of phosphorochloride (POCl_3) at 950 $^\circ\text{C}$ for about 20 min. This sequence was repeated for the second layer of polysilicon. Another set of trenches was etched, again utilizing DRIE and subsequently filled with alumina using atomic layer deposition (1800 nm, 300 $^\circ\text{C}$). The alumina-filled trenches formed the spacers between the electrodes in the linear bending actuator. Finally, device layer etching was completed by etching open trenches again to build the actuator shape, sidewalls and air chambers. Note that this etch step is crucial for the actuators to work properly. Etch insufficiencies in this step are likely inducing two main failure causes. First, conduction paths may remain between electrodes inducing a short cut, thus inhibiting the creation of an electrical field and the genesis of electrostatic forces respectively. Second, electrodes are not etched homogeneously into the depth, i.e. exhibiting a thinning of the electrode with regard to the depth. Following, such thinned electrodes and actuators respectively may not sustain voltages and forces as high as their neighboring actuators causing them to fail early likely because of pull-in. Subsequently openings in the handle wafer were DRI-etched from the backside through 650 μm of silicon to create the air inlets and outlets in the bottom wafer. Finally, part of the oxide between device and handle layer was removed by exposing the wafer to several cycles of hydrogen flouride (HF) which released the actuators and ensures movement capabilities. In parallel, the cover wafer was prepared by

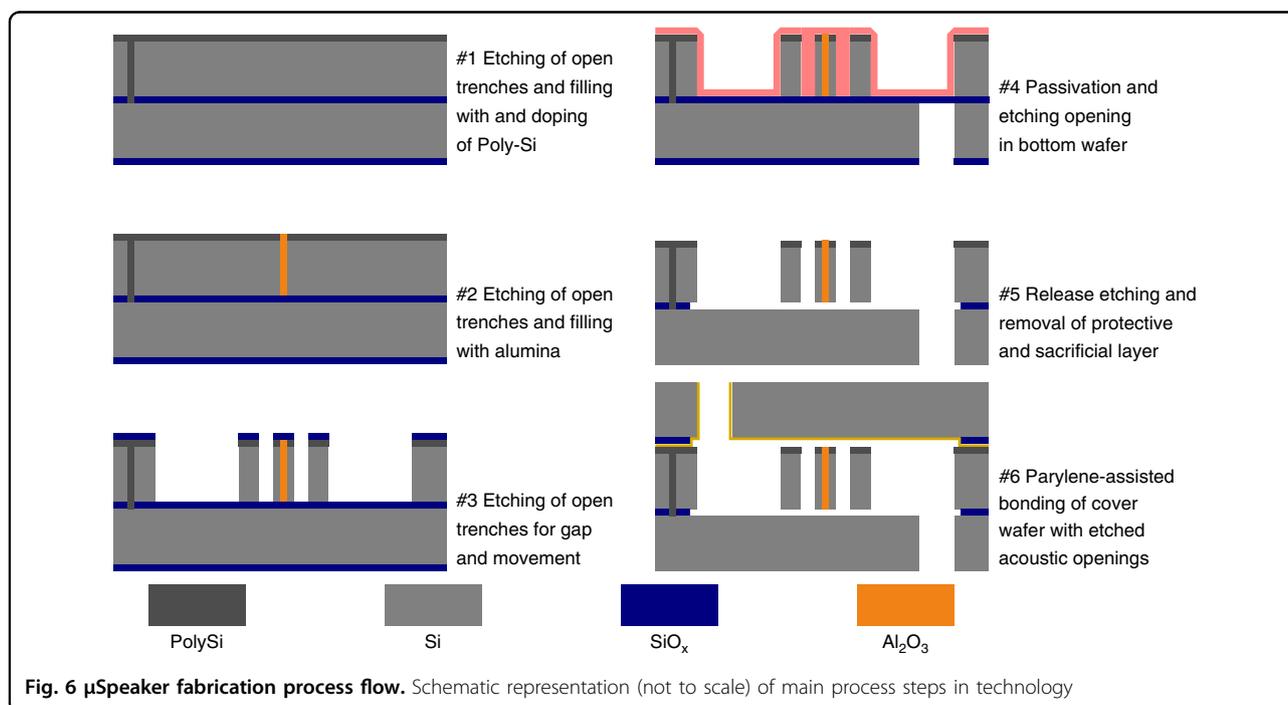


Fig. 6 $\mu\text{Speaker}$ fabrication process flow. Schematic representation (not to scale) of main process steps in technology

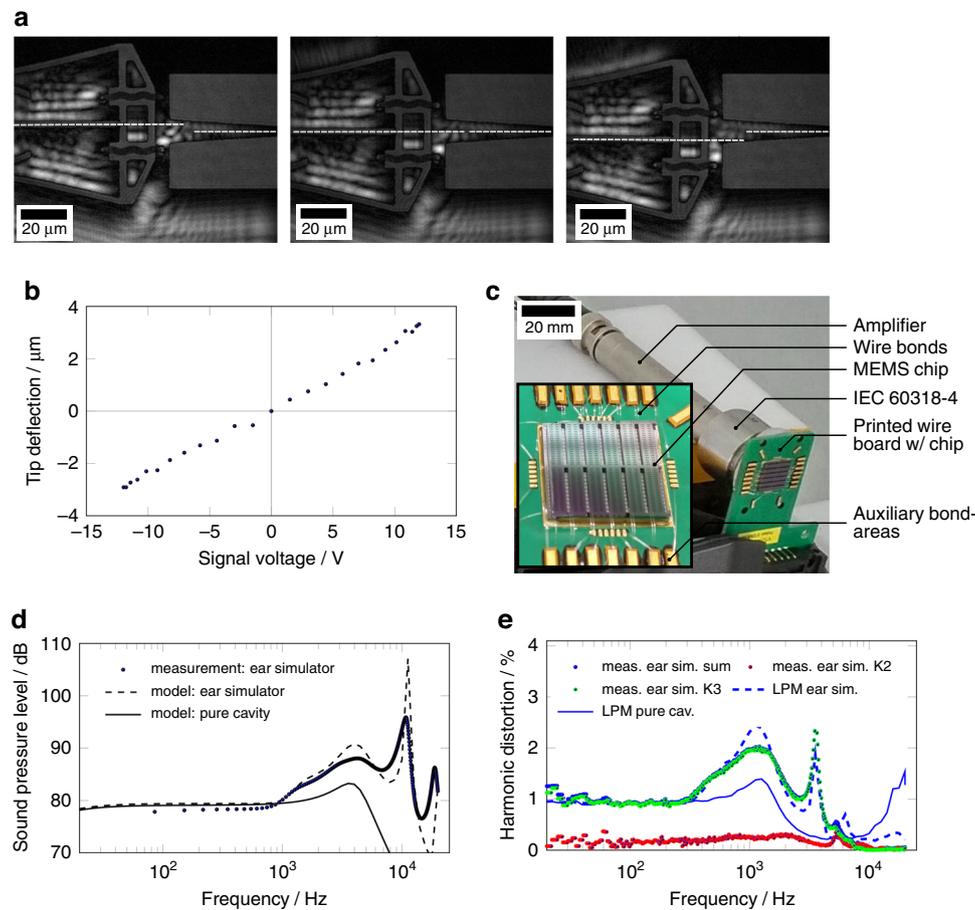


Fig. 7 MEMS specimen setup and characterization. **a** deflection states of actuator for both deflection directions, dashed lines indicate deflection. **b** in-plane deflection measurement data obtained with edge tracking (sinusoidal excitation at $50 \text{ Hz} \pm 12 \text{ V}_{\text{dc}}$) **c** MEMS device on carrier board including wire bonds. **d** sound pressure level measurement data ($\pm 10 \text{ V}_{\text{dc}}, \pm 6 \text{ V}_{\text{p}}$). **e** THD measurement data ($\pm 10 \text{ V}_{\text{dc}}, \pm 6 \text{ V}_{\text{p}}$), THD and K3 measured data overlap, while K2 fraction remains very low

etching a silicon step of $1 \mu\text{m}$ and subsequently etching the inlets and outlets through $400 \mu\text{m}$ of silicon, both by using DRIE. The wafer stack was completed by gluing the cover wafer using parylene to assist wafer bonding³⁹. A layer of 500 nm of parylene was deposited on the cover wafer. Afterwards the device wafer is aligned and clamped to the cover wafer and bonded at $300 \text{ }^\circ\text{C}$ and 2000 mbar of uniform pressure. The parylene will remain as baked glue holding both wafers tightly together. Subsequently, the wafer is separated into individual devices using stealth dicing. An infrared laser weakens the silicon between the chips during multiple passes with different focus depths. Finally, actual cleavage is achieved by expansion of the carrier foil introducing breakage at the respective weakened paths within the silicon.

Measurement setup and results

To optically access the bending actuators, the devices were examined before the cover wafer was bonded on top.

The characterization of the linear actuator deflection was carried out using a digital holographic microscope, see Fig. 7a. The edge tracking function of the associated software allowed resolution down to 163 nm . The electrical contact was established by tungsten needles. These were placed directly onto the surfaces for all three voltages (two constant voltages $\pm 10 \text{ V}_{\text{dc}}$ and audio signal 6 V_{p}).

The optically recorded deflection characteristics in Fig. 7b clearly show the targeted inversion symmetry required in Eq. (3) to suppress the second harmonic distortion K2. This also aligns with the deflection spectra of the same type of actuator measured at its tip showing the second harmonic well below the third harmonic⁴⁰. These measurements were taken by tracking the actuator's tip in both deflection directions as shown in Fig. 7a. The optical measurements reveal a fairly linear deflection up to amplitudes of $\pm 3.1 \mu\text{m}$ ($\pm 12 \text{ V}_{\text{dc}}, \pm 12 \text{ V}_{\text{p}}$). Following the linear relations at the actuator the peak-to-peak deflection available for audio reproduction was $2.6 \mu\text{m}$ ($\pm 10 \text{ V}_{\text{dc}}, \pm 6 \text{ V}_{\text{p}}$).

For acoustic measurements, the chips were glued to a printed wire board (Fig. 7d), using conductive glue, thus providing electrical contact to the bottom wafer. Wire bonds (20 μm diameter, aluminum wires) were attached to the bond pads on the bare silicon surface. Before the acoustic test, we ran a simple electrical test using voltage sweeps up to the designed maximums. This allowed us to identify groups of actuators that were short-circuited or suffered from pull-in at low voltages. Subsequently, such groups of actors were separated during the test. The reasons for such problems lie in the technology, which will be further optimized in the future to increase yield. Finally, only 60 actuator pairs out of 276 (total active chip area 19.8 mm^2) were used in the acoustic measurements due to electrical yield.

The acoustic performance was characterized by measuring the harmonic distortion and the sound pressure level. The ear simulator (IEC 60318-4) was placed at the backside (bottom wafer openings) of the MEMS $\mu\text{Speaker}$ device. The front side was left open to the environment. A small amount of circuitry was used to create both constant voltages ($\pm 10 V_{\text{dc}}$) with a virtual ground. A voltage amplifier (Krohn-Hite Model 7602M) was used to amplify the driving signal of the audio analyzer (NTI Flexus 100). The main signal pattern was a sinusoidal excitation ($6 V_{\text{p}}$) while sweeping the frequency.

The sound pressure level, as displayed in Fig. 7d, shows an almost constant plateau of 79 dB (logarithmic reference pressure is 20 μPa , $\pm 10 V_{\text{dc}}$ and $6 V_{\text{p}}$) for frequencies up to 1 kHz, revealing the quasistatic transduction behavior of the linear bending actuators. No significant drop in SPL is observed below 30 Hz. The referential sound pressure at 500 Hz was ascertained as 78.4 dB. At this frequency, no influence from either dynamical effects or the ear simulator is assumed. Between 1 kHz and 7 kHz, an overall increase in SPL, as well as a peak of 88 dB at a frequency of about 4.2 kHz, can be identified. The peak is clearly related to the first mode bending resonance of the actuators, as expected from electromechanical measurements. The overall increase as well as the steep increase in SPL for frequencies above 7 kHz, is attributed to the resonance of the ear simulator. This resonance leads to an absolute peak in the SPL of more than 101 dB at a frequency of about 9.1 kHz. This peak is not related to the transducer and should be expected when using the ear simulator according to the standard IEC 60318-4.

The THD, displayed in Fig. 7e, shows a constant level of never more than 1.26% for frequencies between 20 Hz and 400 Hz, including a constant plateau at around 1.0%. At frequencies above 400 Hz, we see the consequences of a dominant K3 fraction in the harmonics, i.e. a increasing THD at a frequency that is one third of that at which the SPL increases. The THD reaches a peak value of 2% at a frequency of 1.4 kHz, matching the third harmonic of the

forementioned actuator resonance. Beyond 10 kHz the ear simulator is not specified and contributes to distortion significantly. In accordance with the theory presented above, the K2 component does not dominate the THD. Instead, the THD is dominated by the third harmonic distortion K3.

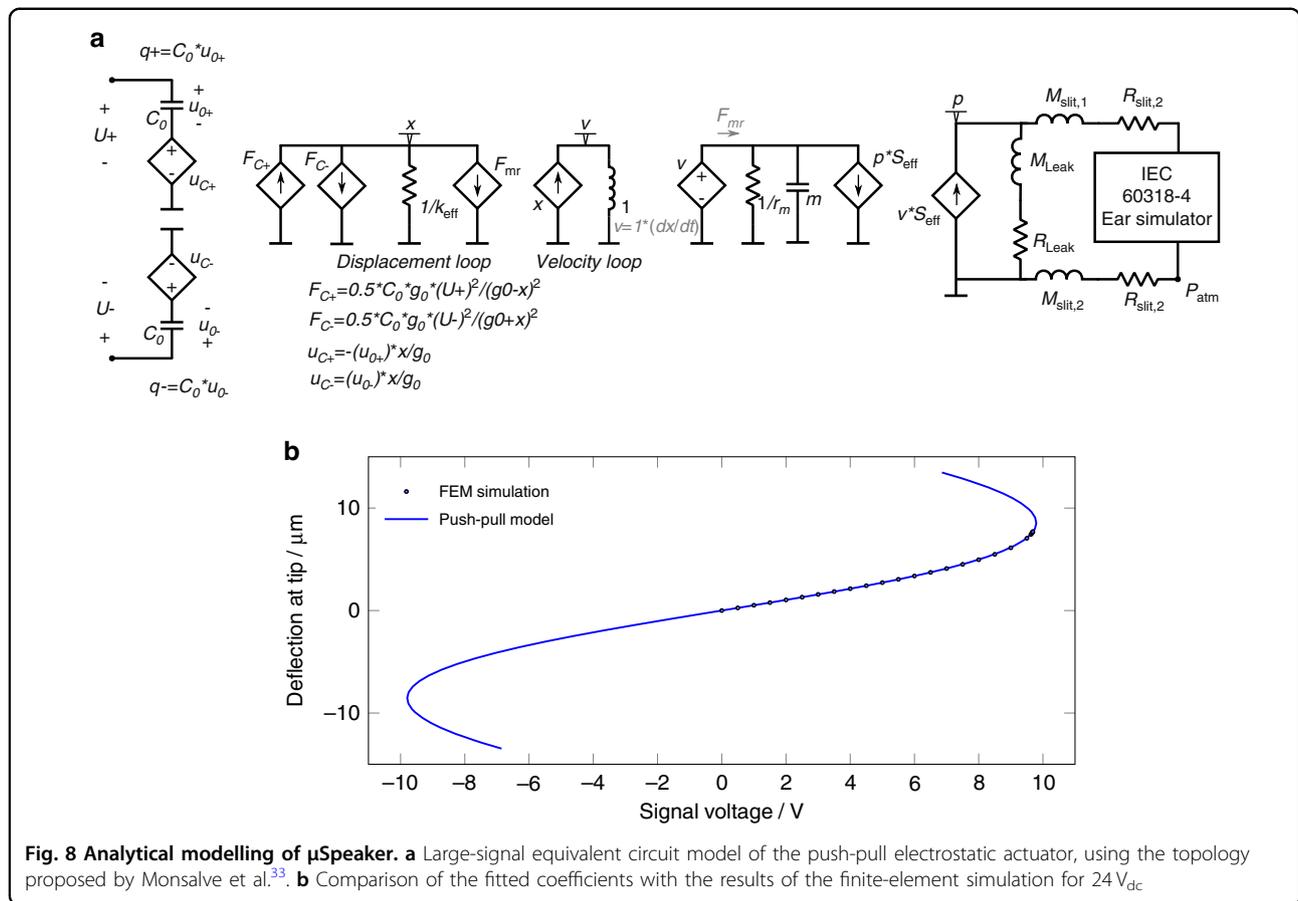
The 60 actuator pairs were measured as having an approximate electrical capacitance of 260 pF using an impedance spectrometer (Keysight E4990A). The total chip capacitance includes some parasitic capacitance originating from stray fields within the chip as well as, e.g., the bond pads. The sensitivity may be calculated assuming a 40% efficiency of the driving circuitry for the pure reactive load. A value of 97.4 dB/mW is achieved at 500 Hz, assuming 1 mW in an IEC 60318-4 ear simulator.

Discussion

A novel, electrostatic push-pull bending actuator for low-voltage operation has been theoretically described, fabricated and characterized. Our new actuator shows a maximum deflection amplitude of 3.1 μm for a driving voltage of $\pm 12 V_{\text{dc}}$ and $12 V_{\text{p}}$. The deflection amplitude exceeds previously reported quasistatic deflection values for electrostatic bending actuators and is normalized to the driving voltage, by a factor of 7.2 (1.5 μm at voltages of $40 V_{\text{dc}}$ and $5 V_{\text{p}}$ ²⁶). This rise is partially attributed to the specific suspension bridge like design of the electrodes. A deviation from a linear behavior may be identified for higher voltages, visible as an increasing slope. This has to be expected due to increased effect of electrostatic and mechanical nonlinearities (ref. Fig. 8b). We demonstrated that our implementation of the push-pull principle yields a deflection, that is both, linear and symmetrical in either direction.

By implementing our novel push-pull actuators in a MEMS $\mu\text{Speaker}$ we demonstrated in particular the viability of a commercially highly attractive application. Our first balanced $\mu\text{Speaker}$ design exhibited excellent audio reproduction over a broad frequency range of more than nine octaves (10 Hz to 6.3 kHz) with a THD below 1.2%. A comparison of the measurements shown in Fig. 1 (our previous work) and in Fig. 7e shows both a significant diminution of the THD and, particularly, the reduction of the K2 component. In other words, the reduction of the THD is attributed to the diminishing K2 component following the symmetry in deflection behavior of the actuator. No signal preconditioning was required to achieve this linearity, giving rise to the argument of a possibly very efficient drive. Additionally this is supported by the electrical capacitance as low as 260 pF keeping potential reactive currents for driving low.

The chip size is $10 \times 10 \text{ mm}^2$. Calculation of the area-specific sound pressure level gives a value of about 53.8 dB/ mm^2 . This is 4.2 dB above our previously reported value and must be highlighted in combination with the



significantly reduced harmonic distortion. This is partly achieved thanks to the denser packing of the actuators within the device layer. In the specific case of our presented device one may consider all actuator pairs (276 instead of 60) being active. Theoretically, an increase in SPL of about 13 dB could be anticipated. Note that this would be the case without increasing neither driving voltages nor the THD. With ongoing development of the fabrication technology it should be possible to achieve these improvements.

The models presented in comparison to the measured values show a good approximation. Deviations are attributed to the tuning of the models parameters. Fabrication imperfections not only during etching but also for wafer bonding may explain the perceived differences. Especially fluidic damping of the moving structures is quite sensitive to the curves. Following the peak heights the actuators resonance and the ear simulators resonance differ in both the SPL and THD measurements. As revealed in the THD measurement, the K3 fraction is dominating. Seemingly stochastic measurement variations for low-frequency THD values are attributed to limits of the microphone setup, since 1% translates to -40 dB in measuring value.

The demonstration of the linear bending actuators showed their application in a MEMS $\mu\text{Speaker}$. The unique bending capabilities not only allowed relevant performance indicators to be significantly enhanced, but also could potentially allow for reduced device sizes. This is of major interest for the commercialization of any MEMS device. In this context, it is important to mention that the silicon fabrication technology is compatible with complementary metal-oxide semiconductor technology.

Perspective

Further work will focus on technology development to decrease the minimum possible gap distances and increase area utilization at the same time. One major challenge is to keep the electrical capacitance low, not only for the actuators but also for sources of additional parasitic capacitances.

In the field of MEMS-based speakers, quite impressive piezoelectric speakers are at this moment in time probably the most advanced (e.g. USound). Other approaches (e.g. AudioPixels) have, to the best knowledge of the authors, so far had no or only very little experimental data relevant to practical applications published. When compared to accessible data, the $\mu\text{Speakers}$ presented in this

paper have an attractive performance profile. The electrostatic actuation principle does not encounter a piezoelectric hysteresis or intrinsic power dissipation mechanisms²⁸, resulting in a THD < 1% without any further acoustic design or signal preprocessing. The linearization strategies which furnish macroscopic electrostatic speakers with their spectacular performance have successfully been translated to the low-voltage MEMS world.

The electrostatic μ Speakers presented in this paper also promise major progress regarding power consumption and peak current draw. The batteries of advanced in-ear devices are tiny (typical 60 mAh). Most of the battery budget is reserved for smart functions such as speech recognition and wireless connectivity. This confines the power available to the audio reproduction system to only a small, single-digit number of milliwatts. MEMS μ Speakers must beat this target to be competitive with classical electrodynamic or balanced armature speakers. The relevant key parameter for the battery drain caused by the transducer is the electrical capacity. The capacity determines the reactive peak current and the reactive power that must be handled by the inevitably power-dissipating drive electronics. The μ Speakers presented here feature a total electric capacity of less than 1 nF. In comparison, capacity values in excess of 20 nF or even 150 nF have been published for piezoelectric systems (USound, xMEMS). The beneficial combination of a comparatively low signal voltage and a low actuator capacity allows the presented μ Speaker to be driven with small, fully integrated charge pumps connected to a small lithium polymer or zinc-air battery.

While a fair comparison should be between μ Speakers that are all able to deliver peak SPLs of say about 120 dB across the relevant parts of the audio spectrum, an assessment of the adiabatic work required to generate 120 dB in the ear channel in relation to the energy content of the electrical field available from the electrostatic cells reveals that it should be possible to design an electrostatic 120 dB MEMS μ Speaker with a capacitance far below 1 nF.

Materials and methods

Finite element analysis

In the main paper body we gave the relevant principles, which lead to the constitution of the design of the actuator and specifically of the electrodes shape. Starting from a three-electrode basis having three straight electrodes, it is beneficial to have a suspension bridge arrangement because of its mechanical susceptibility to bending. Thus, the general shape is set. Parameters which need to be defined are: the length of the electrodes parallel to the longitudinal axis of the actuator, the electrodes distances or gap, their thicknesses and the height of the electrodes anchor. Due to symmetry reasons there are just two electrode thicknesses, i.e. middle and outer electrode, and one

gap thickness to be defined. All these parameters mentioned before are dependent from each other. For example, a long electrode (e.g. >600 μ m) would require a bigger gap (e.g. >4 μ m) for a fixed voltage to not to be susceptible to pull-in. Conversely, longer electrodes are softer, thus enabling higher deflections with the same force. To address all these dependencies and include actual features of a possible fabrication design, finite element analysis seem suited to define the parameters.

As a method, we implemented the Covariance Matrix Adaptation Evolution Strategy⁴¹. A method that searches for beneficial sets of parameters by calculating several individuals and adapt the search over several generations with then newly created individuals. This method seems suitable because the behavior of the actuator based on the dependencies is noisy. In total, several thousands of individuals have been calculated.

We used a 2D model which implements both the mechanical and the electrostatic domain. Calculations were carried out using ANSYS Software and by implementing plane223 elements which are capable of both domains mentioned before. The list of results with the parameter sets and their respective maximum deflection was then sorted regarding this variable of merit. This parameter set was implemented in the layout process for fabrication. Table 1 shows the parameter set chosen for the layout.

Lumped-parameter model

If a push-pull, spring-capacitor system can be used to describe the motion of the actuator, there will be a set of parameters (i.e. capacitance and spring constant) for which the S-curve obtained by finite-element simulations can be reproduced analytically. The electrostatic gap is set by design (2.5 μ m), and the zero-voltage capacitance can be obtained by geometrical considerations. In order to describe the mechanical amplification achieved by the NED mechanism, we have introduced a lever factor θ that couples the electrode displacement and the motion of the tip, as was proposed by Spitz et al.³¹. This lever factor and the equivalent spring constant can be obtained by

Table 1 Parameter set chosen for the layout for fabrication

Parameter	Value	Unit
Electrode length	430	μ m
Middle electrode thickness	5	μ m
Outer electrode thickness	2	μ m
Electrode distance/gap	2.5	μ m
Height of electrode anchor	34	μ m

Table 2 Parameters used in the equivalent circuit model

Parameter	Fitting against FEM	Adjustment to experiments
Spring constant, k_{eff}	97.8 N/m	77.8 N/m
NED lever factor, θ	10.8	9.81
Effective mass, m_{eff}	1.25×10^{-7} kg	9.96×10^{-8} kg
Quality factor, Q	–	2.5

performing a least-squares fit using equation (11), inserting the simulated values of the deflection for the corresponding AC and DC voltages. The obtained parameters are shown in Table 2 and the comparison of the analytical curve and the finite-element simulation is shown in Fig. 8b. It can be confirmed that the analytical S-curve describes the motion of the actuator very well. Given that fabrication tolerances have an impact on the actual geometry of the device, these parameters were slightly adjusted when comparing them to the measurements of the speakers. The Q-factor was obtained by trial and error, in particular by comparing with the THD curve, which is very sensitive to this value.

$$k_{eff} \left(\frac{w}{\theta} \right) = \frac{C_0}{2g_0} \left[\frac{(u_{dc} + u_{ac})^2}{\left(1 - \frac{w}{\theta g_0}\right)^2} - \frac{(u_{dc} - u_{ac})^2}{\left(1 + \frac{w}{\theta g_0}\right)^2} \right] \quad (11)$$

In order to simulate this lumped-parameter model as an equivalent circuit, the procedure proposed by Monsalve et al.³³ was implemented for the case of the push-pull actuator. This work describes how the large-signal behavior of a spring-capacitor system can be fully represented by an equivalent circuit, which allows its harmonic distortion and the interaction with the acoustical load (IEC 60318-4 ear simulator) to be simulated. This equivalent circuit (Fig. 8a) is composed of cross-coupled loops that represent the electrical, mechanical and acoustic domains. The electrical domain (left-hand side) contains the two variable capacitors connected in series. The mechanical domain portrays the force balance (Eq. (11)), with the two opposing electrostatic forces. The acoustic load is found on the right-hand side, where air is pumped by the actuator through the slits (including some leakage) towards the ear simulator. This lumped-parameter model is able to predict the behavior of the μ Speaker device over a broad range of frequencies.

Acknowledgements

This work was supported by Fraunhofer Zukunftsstiftung and by the State of Brandenburg and the European Union under contract PROFIT/EFRE No. 80256335.

Author details

¹Fraunhofer Institute for Photonic Microsystems IPMS, Dresden, Germany. ²Arioso Systems GmbH; since Sep. 1st 2022: BOSCH Sensortec GmbH, Dresden, Germany. ³Chair of Micro and Nano Systems, Brandenburg University of Technology Cottbus-Senftenberg, Cottbus, Germany

Author contributions

B.K. developed the practical actuator design and created the layout of the MEMS device, accompanied fabrication and coordinated characterization. H.A.G.S. invented and modeled the concept of implementing the electrostatic push-pull principle in an NED-based μ Speaker for low-voltage operation. Lutz Ehrig performed the acoustic characterization, including the design of a suitable experimental setup. His measurements provided the empirical basis for parameter extraction from circuit models. F.W. validated the theories by H.A.G.S. and investigated the parameter dependency of the non-linearity system performance in depth. S.L. was responsible for the fabrication in the IPMS cleanroom facilities and consulted on the layout. M.S. performed the mechanical and electrical measurements. J.M.M.G. developed and applied the non-linear equivalent circuit model for simulation and parameter extraction. A.M. performed extensive FEM simulations of the non-linear actuator performance. D.S. optimized the NED actuator geometry by extensive FEM simulations. H.C. and H.S. proposed the concept of NED μ Speakers and revised the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Conflict of interest

The authors declare no competing interests.

Received: 26 May 2022 Revised: 30 August 2022 Accepted: 28 September 2022

Published online: 30 November 2022

References

- Lee, S. S., Ried, R. & White, R. Piezoelectric cantilever microphone and microspeaker. *J. Microelectromech. Syst.* **5**, 238–242 (1996).
- Han, C.-H. & Kim, E. S. Polyethylene-diaphragm piezoelectric acoustic transducers. In *Proc IEEE Thirteenth Annual International Conference on Micro Electro Mechanical Systems* (Cat. No.00CH36308), 148–152 (IEEE, 2000).
- Ko, S. C., Kim, Y. C., Lee, S. S., Choi, S. H. & Kim, S. R. Micromachined piezoelectric membrane acoustic device. *Sens. Actuators A Phys.* **103**, 130–134 (2003).
- Yi, S., Yoon, M. & Ur, S. Piezoelectric microspeakers with high compressive zno film and floating electrode. *J. Electroceram.* **23**, 295 (2009).
- Casset, F. et al. A 256 mems membrane digital loudspeaker array based on pzt actuators. *Procedia Eng.* **120**, 49–52 (2015).
- Arevalo, A., Conchouso, D., Castro, D., Kosel, J. & Foulds, I. G. Piezoelectric transducer array microspeaker. In *Proc 2016 IEEE 11th Annual International Conference on Nano/Micro Engineered and Molecular Systems (NEMS)*, 180–183 (IEEE, 2016).
- Tseng, S.-H. et al. Sound pressure and low frequency enhancement using novel pzt mems microspeaker design. In *Proc 2020 IEEE 33rd International Conference on Micro Electro Mechanical Systems (MEMS)*, 546–549 (IEEE, 2020).
- Wang, H., Chen, Z. & Xie, H. A high-spl piezoelectric mems loud speaker based on thin ceramic pzt. *Sens. Actuators A Phys.* **309**, 112018 (2020).
- Diamond, B., Neumann, J. & Gabriel, K. Digital sound reconstruction using arrays of cmos-mems microspeakers. In *Technical Digest. MEMS 2002 IEEE International Conference. Fifteenth IEEE International Conference on Micro Electro Mechanical Systems* (Cat. No.02CH37266), 292–295 (IEEE, 2002).
- Kim, H. et al. Bi-directional electrostatic microspeaker with two large-deflection flexible membranes actuated by single/dual electrodes. In *IEEE Sensors*, 89–92 (IEEE, 2005).
- Zhou, Q. & Zettl, A. Electrostatic graphene loudspeaker. *Appl. Phys. Lett.* **102**, 223109 (2013).
- Glaser, C., Dehé, A., Tumpold, D. & Laur, R. Silicon microspeaker with out-of-plane displacement. In *Proc the 9th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (NEMS)*, 12–16 (IEEE, 2014).

13. Murarka, A., Lang, J. H. & Bulovic, V. Printed membrane electrostatic mems microspeakers. In *Proc 2016 IEEE 29th International Conference on Micro Electro Mechanical Systems (MEMS)*, 1118–1121 (IEEE, 2016).
14. Arevalo, A., Castro, D., Conchouso, D., Kosel, J. & Foulds, I. G. Digital electrostatic acoustic transducer array. In *Proc 2016 IEEE 11th Annual International Conference on Nano/Micro Engineered and Molecular Systems (NEMS)*, 225–228 (2016).
15. Sano, C., Ataka, M., Hashiguchi, G. & Toshiyoshi, H. An electret-augmented low-voltage mems electrostatic out-of-plane actuator for acoustic transducer applications. *Micromachines* **11** <https://www.mdpi.com/2072-666X/11/3/267> (2020).
16. Tian-Ling, R., Lin-Tao, Z., Li-Tian, L. & Zhi-Jian, L. Design optimization of beam-like ferroelectrics-silicon microphone and microspeaker. *IEEE Trans Ultrasonics Ferroelectrics Frequency Control* **49**, 266–270 (2002).
17. Chen, Y. C., Liu, W.-T., Chao, T.-Y. & Cheng, Y. T. An optimized cu-ni nanocomposite coil for low-power electromagnetic microspeaker fabrication. In *TRANSDUCERS 2009 - 2009 International Solid-State Sensors, Actuators and Microsystems Conference*, 25–28 (IEEE, 2009).
18. Chen, Y. C. & Cheng, Y. T. A low-power milliwatt electromagnetic microspeaker using a pdms membrane for hearing aids application. In *Proc 2011 IEEE 24th International Conference on Micro Electro Mechanical Systems*, 1213–1216 (IEEE, 2011).
19. Shahosseini, I. et al. Electromagnetic mems microspeaker for portable electronic devices. *Microsyst. Technol.* **19**, 879–886 (2013).
20. Majlis, B. Y., Sugandi, G. & Noor, M. M. Compact electrodynamic mems-speaker. In *Proc 2017 China Semiconductor Technology International Conference (CSTIC)*, 1–3 (IEEE, 2017).
21. Suk, J. W., Kirk, K., Hao, Y., Hall, N. A. & Ruoff, R. S. Thermoacoustic sound generation from monolayer graphene for transparent and flexible sound sources. *Adv. Mater.* **24**, 6342–6347 (2012).
22. Fei, W., Zhou, J. & Guo, W. Low-voltage driven graphene foam thermoacoustic speaker. *Small* **11**, 2252–2256 (2015).
23. Wang, H. et al. Review of recent development of mems speakers. *Micromachines* **12**, 1257 (2021).
24. Sugimoto, T. et al. A lightweight push-pull acoustic transducer composed of a pair of dielectric elastomer films. *J. Acoust. Soc. Am.* **134**, EL432–7 (2013).
25. Chiang, H.-Y. & Huang, Y.-H. Experimental modeling and application of push-pull electrostatic speakers. *J. Acoust. Soc. Am.* **146**, 2619 (2019).
26. Kaiser, B. et al. Concept and proof for an all-silicon mems micro speaker utilizing air chambers. *Microsyst. Nanoeng.* **5**, 43 (2019).
27. Marxer, C., Manzardo, O., Herzig, H. P., Dändliker, R. & de Rooij, N. F. An electrostatic actuator with large dynamic range and linear displacement - voltage behaviour for a miniature spectrometer (poster). In *Proc Transducers'99 10th Internat. Conf. on Solid-State Sensors and Actuators*, 786–789 (1999).
28. Hunt, F. V. *Electroacoustics: The Analysis of Transduction, and Its Historical Background*, vol. 5 of *Harvard Monographs in Applied Science* (Harvard University Press, s.l., 1954).
29. Conrad, H. et al. A small-gap electrostatic micro-actuator for large deflections. *Nat. Commun.* **6**, 10078 (2015).
30. Sennheiser. Sennheiser he 1: Reshaping excellence – the legend reborn. <https://en-us.sennheiser.com/sennheiser-he-1> (2021).
31. Spitz, B. et al. (eds.). *Audio-Transducer for In-Ear-Applications Based on CMOS-Compatible Electrostatic Actuators* (IEEE, 2019).
32. Melnikov, A. et al. Coulomb-actuated microbeams revisited: experimental and numerical modal decomposition of the saddle-node bifurcation. *Microsyst. Nanoeng.* **7** (2021).
33. Monsalve, J. M. et al. Large-signal equivalent-circuit model of asymmetric electrostatic transducers. *IEEE/ASME Trans. Mechatron.* **27**, 2612–2622 (2021).
34. Rugh, W. J. *Nonlinear system theory: the volterra/wiener approach*. In *Johns Hopkins Series in Information Sciences and Systems* (Hopkins Univ. Press, Baltimore, Md, 1981).
35. Boyd, S., Chua, L. O. & Desoer, C. A. Analytical foundations of volterra series. *IMA J. Math. Control Inf.* **1**, 243–282 (1984).
36. Wambacq, P. & Sansen, W. *Distortion Analysis of Analog Integrated Circuits*, vol. 451 of *The Kluwer International Series in Engineering and Computer Science, Analog Circuits and Signal Processing* (Springer, Boston, MA, 1998).
37. Kovacic, I. (ed.) *The Duffing Equation: Nonlinear Oscillators and their Behaviour* (Wiley, Chichester, 2011).
38. Schenk, H. A. et al. Electrically actuated microbeams: An explicit calculation of the coulomb integral in the entire stable and unstable regimes using a chebyshev-edgeworth approach. *Phys. Rev. Appl.* **18**, 014059 (2022).
39. Kim, H. & Najafi, K. Characterization of low-temperature wafer bonding using thin-film parylene. *J. Microelectromech. Syst.* **14**, 1347–1355 (2005).
40. Melnikov, A. et al. Nonlinearity of balanced mems loudspeakers: Optical experiments and numerical modeling using time-harmonic signals. In *Proc 27th International Congress on Sound and Vibration, Annual Congress of International Institute of Acoustics and Vibration (IIAV)* (Silesian University Press, 2021).
41. Hansen, N. & Ostermeier, A. Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation. In *Proc IEEE International Conference on Evolutionary Computation*, 312–317 (IEEE, 1996).