Audio Engineering Society

Convention Express Paper 123

Presented at the 155th Convention
2023 October 25–27, New York, USA

In-Ear Headphones on Ear Canal Simulator vs Real Human Ear Geometries: Quantifying the Differences with Simulations

Andri Bezzola[1]

[1] *Samsung Research America, DMS Audio, Valencia CA 91355*

Correspondence should be addressed to Andri Bezzola (andri.b@samsung.com)

ABSTRACT

There are several challenges in predicting the pressure response of a sound source at the eardrum. Placing microphones at or closely near the eardrum is challenging and not feasible for daily use. In-ear headphones outfitted with microphones can provide an estimation of the pressure at eardrum by measuring the pressure at the headphone microphone and by using a transfer function of the pressure at the microphone to the eardrum. The accuracy of the estimation thus depends on the accuracy of the transfer function. We have used Finite Element simulations to calculate an in-ear headphone in twenty ear canal geometries of ten different people and compared the results with measurements and simulation of the same headphone in a mannequin outfitted with ear canal simulators. The results show that using the mannequin's transfer function to predict personalized pressure responses at the ear drum will result in error of up to 7.8 dB below 5 kHz and even greater above 5 kHz. Two simulation studies with different targets at the ear drum showcase the need for better estimates of the transfer function from in-device microphone to eardrum when attempting to control the pressure response at the eardrum.

1 Introduction

Objective headphone and earbud evaluation has recently centered around measuring them on a mannequin or ear simulator. Designers typically have a target curve that defines their optimal frequency response of the device. For stereo reproduction, the preferred headphone target approximates the in-room response of a an accurate loudspeaker in a semi-reflective room. A good summary of the methodologies to establish such target curves is described in [1] and the references therein. This paper does not advocate for or against any specific target curve. Instead it focuses on how the individual shapes of ear canals influence the sound pressure level at the eardrum.

Measurements of the pressure at the ear Drum Reference Point (DRP) in humans are extremely difficult and bear a high risk of injury to the eardrum and ear canal. In order to circumvent this challenge, ear simulators have been developed to predict a population average of the pressure response at DRP, i.e. the results of the sim-

ulator should mimic the acoustic behavior of an average human ear. However, these simulators do not predict the range and variance of the pressure responses at DRP, only their average. Recent research by Olive et al. suggests that there is significant variance in frequency response of over-ear headphones, and they emphasize the need for personalized headphone solutions [2].

Another goal of a headphone could be to sound maximally transparent [3]. This can be particularly true for hearing aids or headphones with pass-through function. An external microphone records the surrounding sounds at the ear canal entrance and the device reproduces the sound inside the ear canal. A headphone would not sound transparent if the spectral content of the reproduced sound at DRP is changed from what the spectral content would have been if it had been able to reach the eardrum unobstructed by the device itself. Models for the transfer functions from external microphone to internal transducer and finally to the DRP are used to predict the pressure response at DRP, but the predictions are only as good as the estimates of the transfer functions.

The pressure response at DRP for in-ear headphones outfitted with microphones can be estimated by means of a transfer function $G(f)$ that relates the pressure at DRP $D(f)$ with the pressure at the Near-Field Microphone (NFM) inside the earbud $M(f)$. If an accurate estimate of $G(f)$ can be obtained for an individual ear, then we can also accurately predict $D(f)$ for that ear. The absolute value of the pressure at DRP can be calculated in dB as

$$|D(f)| = |M(f)| + |G(f)| \tag{1}$$

This approach can be used to provide a personalized equalization of earbuds, for example by equalizing the earbud to a target $M^*(f)$ that is defined at the headphone microphone [4].

A reasonable estimate of such a transfer function $G(f)$ could be obtained by measuring the headphone in an ear simulator and measuring the pressure responses at NFM and at DRP of the simulator. The question that remains is how the pressure at DRP will differ in a real human ear compared to the simulator. We present two studies that use Finite Elements (FEM) to simulate the pressure response in human ear geometries that were obtained by Magnetic Resonance Imaging (MRI) scans and translated into CAD geometries. The first study compares these simulation results with simulation and measurement of a Head and Torso Simulator (HATS) and we quantify the differences between human ears and HATS-based results for the transfer functions $G(f)$. A second study expands these results with the inclusion of simulations of a far-field source in front of the heads. Results suggest that a maximally transparent headphone on a mannequin does not directly translate to a maximally flat headphone for an individual.

2 Methods

2.1 Modeling and Measurement of Earbud in KEMAR Mannequin Ear With 711 Coupler

2.1.1 Modeling Setup

We modeled the anthropometric pinna of the GRAS 45BB KEMAR dummy head coupled to a standard 711 coupler. The geometry of the coupler was taken from COMSOL's tutorial model [5] and closely corresponds to the Brüel & Kjær Simulator Type 4157. The model of the 711 coupler was calibrated to match the transfer impedance as function of frequency within the tolerances of standard IEC 60318-4 [6] up to 10 kHz.

The CAD geometry of a Samsung Galaxy Earbud 2 Pro was placed in the CAD geometry of the left anthropometric pinna of the GRAS 45BB KEMAR dummy head such that a tight coupling between earbud tip and ear canal would be ensured. Furthermore, the insert distance was chosen to allow a natural fit of the earbud with the shape of the outer ear. Figure 1 shows how the earbud was positioned inside the dummy head ear and connected to the 711 coupler and the resulting air volumes used in the simulation.

With exception of the transducer and coupler microphone surface, the surfaces of the 711 coupler and the earbud were treated as rigid boundaries, and the surfaces of the ear canal were assigned the surface impedance of skin that is included in the Acoustics Module of COMSOL Multiphysics. The surface of the coupler microphone was assigned a lumped acoustic RCL impedance with the following elements in series: acoustic resistance $R_{ac} = 119 \times 10^6\,\mathrm{N\,s/m^5}$, acoustic capacitance $C_{ac} = 62 \times 10^{-15}\,\mathrm{m^5/N}$, and acoustic inductance $L_{ac} = 710\,\mathrm{kg/m^4}$.
The narrow connectors between the central cylinder and the outer resonance rings of the coupler were modeled with the appropriate narrow region acoustics to account for the acoustic damping. All details of the

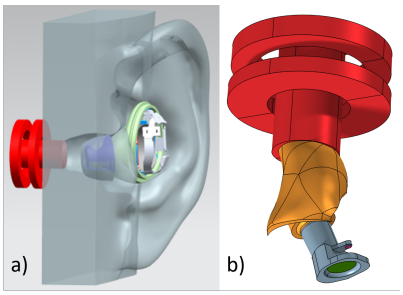

Fig. 1: a) Positioning of Samsung Galaxy Earbud 2 Pro in the pinna model of the GRAS head simulator coupled to a generic 711 coupler. **b)** Air volumes used in simulation, including 711 coupler (red), part of the ear canal (orange), air volume in the earbud (grey), surface of the earbud transducer (green), and the near-field microphone surface (pink).

implementation of the acoustics in the coupler can be found in [5].

The specific acoustic surface impedance of the skin used to model the ear canal is shown in Figure 2. The real part is held constant at 0.06 MPa s/m, but the imaginary part is frequency dependent.

The earbud has a screen and wax guard at its entrance to avoid contamination with ear wax and other debris. These elements also act as acoustic dampers, and must thus be taken into account when modeling. The wax guard is metal grill of 0.1 mm thickness, with square holes of 0.1 mm width, and open-area ratio of 0.444. The resulting real and imaginary parts of the grill transfer impedance are shown in Figure 3. In the COMSOL model, the wax guard has been implemented as a "Interior Perforated Plate" feature. The screen's supplier has specified an acoustic impedance of 8 Pa s/m for the screen covering the wax guard, which has been implemented as an "Interior Impedance" feature in the COMSOL model.

For excitation, we prescribed the displacement of the diaphragm H by a Butterworth third-order low pass filter with center frequency $f_c = 3700$ Hz, which can be defined as

$$H = \frac{H_0}{\sqrt{1 + \left(\frac{j\omega}{j\omega_c}\right)^{2n}}} \tag{2}$$

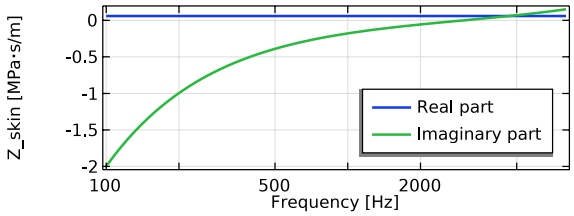

Fig. 2: Real and imaginary part of the skin surface impedance used to model the ear canal.

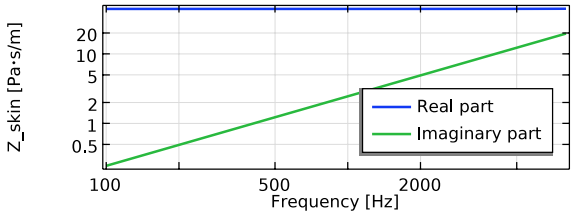

Fig. 3: Real and imaginary part of ear wax grill transfer impedance.

with H0 = 1.2 μm, $\omega = 2\pi f$, and $n = 3$. The simulation then calculated the resulting pressure distribution in the earbud, ear canal and coupler at frequencies between 100 Hz and 8 kHz. The resulting pressure at NFM and DRP are shown in Figure 4.

2.1.2 Results

The COMSOL model of the Galaxy Buds 2 Pro in the GRAS KEMAR mannequin was validated by measurements of the pressure at NFM and DRP using a modified earbud that was placed in a real GRAS KEMAR mannequin. The earbuds were modified to allow direct access to the voltage of the NFM microphone in order to get real-time measurements of the pressure at NFM. The earbud was measured with 11 repeated inserts into the pinna and ear canal entrance positions.

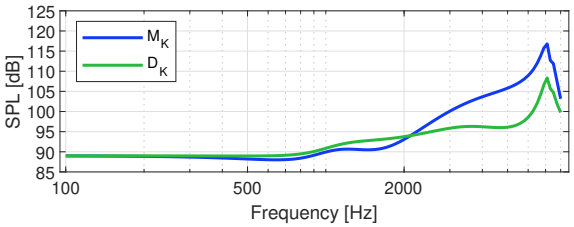

Fig. 4: Simulated SPL at DRP and NFM locations in the test setup with the KEMAR mannequin.

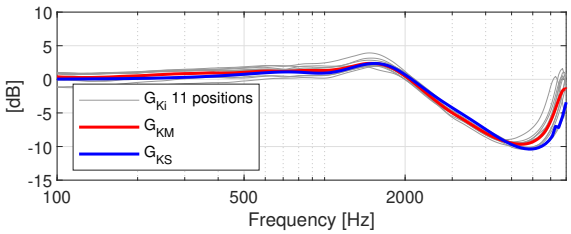

Fig. 5: Measured transfer functions G_{Ki} and their mean G_{KM} as well as simulated G_{KS} of the pressure at NFM to pressure at DRP in the KEMAR mannequin.

Using equation (1) and the pressure at NFM and at DRP we can now easily calculate the magnitude of the transfer function $G(f)$:

$$|G(f)| = |D(f)| - |M(f)| \tag{3}$$

With knowledge of $G(f)$ it is possible to predict the frequency response of SPL at DRP even in cases where the response at DRP is not directly accessible.

Figure 5 shows the frequency responses for $G_{K_i}(f)$ obtained from measurements with the KEMAR mannequin and 711 coupler at 11 different insertions. We the calculated the db-mean, $G_{KM}(f)$ of these measurements. The calculated transfer function for the FEM simulation, $G_{KS}(f)$, corresponds excellently with $G_{KM}(f)$, indicating a high confidence that the impedance values for the wax guard and mesh are valid.

2.2 Modeling of Earbud Response at DRP in Individual Ears

.

2.2.1 Modeling Setup

The "IHA database of human geometries including torso, head and complete outer ears for acoustic research" is an open-source database that contains the geometries of upper torso, head, and outer ear (including the ear canals) of 10 individuals [7, 8]. Torso, head, and ear shapes are provided in STL format and can be imported into CAD or simulation software packages. A selection of geometries is shown in Figure 6.

To model the pressure response at DRP we combined the CAD data of IHA models of ear canals with the

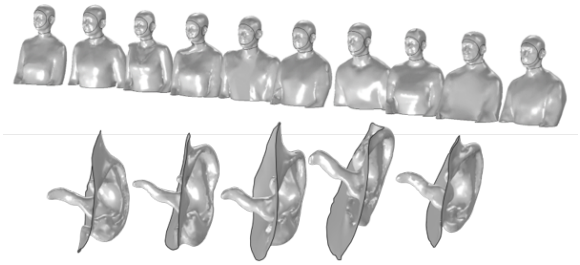

Fig. 6: Torso, head, and outer ear geometries of some of the individuals provided in the IHA database.

CAD model model of an earbud. The earbud was implemented in the same way as in Section 2.1. The ear drum was modeled as a circular area with a diameter of 5 mm on the surface of the ear canal. The impedance of the ear drum was taken as the COMSOL-implented version of the ear drum impedance published by Hudde and Engel in [9]. The remainder of the ear canal was assigned the acoustic surface impedance of human skin shown in Figure 2. The earbud was positioned in the individual ear canals in a way that prevented excessive protursion of the main body of the earbud through the surface of pinna, concha and ear canal. Each ear canal was modeled with the earbud at 5 different insert depths within a range of 5 mm, resulting in 100 different configurations (10 humans, 2 ears, 5 positions each). In each configuration the diaphragm of the earbud was assigned the excursion from (2).

2.2.2 Results

The range of the resulting pressure responses at NFM and DRP of the individual configurations, along with the responses in the KEMAR mannequin, is shown in Figure 7. The measurement of D_K in the mannequin does not appear to be a great predictor for the individual responses D_i at the ear drum for the sample population of the 10 people used in the IHA database. The difference between D_i and D_K is plotted in Figure 8.

In an attempt to reduce the errors at DRP, the earbud can employ an equalization Q_i to control the pressure response at NFM. Assuming a target pressure response at the ear drum, D^*, is the design goal, then the target at near-field microphone, M^*, can be calculated as

$$M^* = D^* - G_K. \tag{4}$$

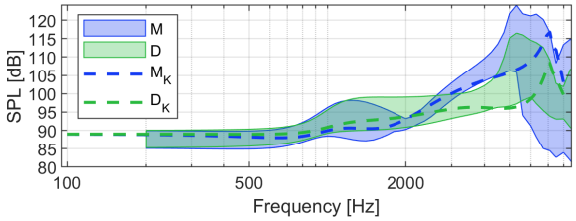

Fig. 7: Spread of simulated NFM and DRP pressure responses for IHA geometries (shaded areas) and corresponding responses from simulation in KEMAR mannequin (dashed lines).

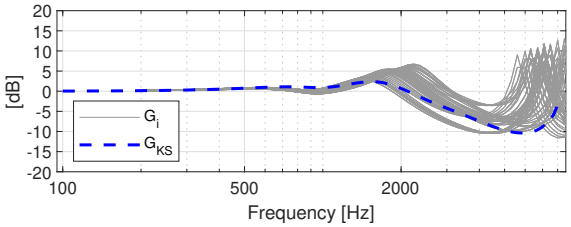

Fig. 9: Simulated individual pressure transfer functions G_i from NFM to DRP in IHA database models (gray) compared to simulated G_S in KEMAR mannequin (dashed blue).

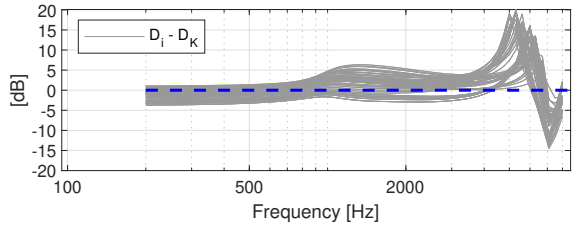

Fig. 8: Difference between individual responses at DRP and response in KEMAR mannequin.

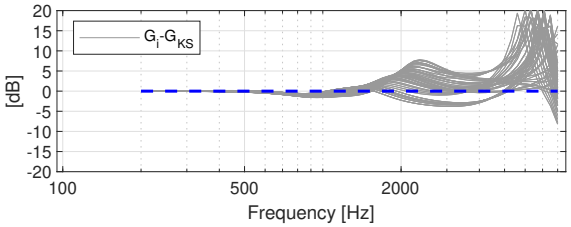

Fig. 10: Error between the individual transfer functions G_i and the transfer function simulated in the KEMAR mannequin G_{KS}.

During use of the earbud, the personalized equalization Q_i will control $M_i(f)$ to be equal to $M^*(f)$

$$Q_i = M^* - M_i \tag{5}$$

The resulting pressure response at DRP after equalization D_i^Q can thus be calculated as

$$D_i^Q = M^* + G_i = D^* + (G_i - G_K) \tag{6}$$

The error estimate at the eardrum ε_D can now be calculated as the difference between the individual transfer function G_i and the mannequin's transfer function G_K, irrespective of the actual shape of the target function D^*:

$$\varepsilon_{D_i} = D_i^Q - D^* = G_i - G_K \tag{7}$$

To get an idea of the remaining error we have plotted G_i and G_{KS} in Figure 9. The error between individual transfer functions G_i and the transfer function in the mannequin G_{KS} is shown in Figure 10. Below 1500 Hz, the error between G_i and G_{KS} is less than 2 dB, but it can be as high as 7.8 dB around 2350 Hz. Above 5 kHz the error further increases to more than 20 dB in some cases. These results suggest that if the goal is to achieve a target SPL response at DRP $D^*(f)$, a personalized equalization at NFM only significantly reduces the error below about 1800 Hz. In order to correct at higher frequencies, a better estimate of the individual transfer function from NFM to DRP will need to be calculated.

2.2.3 Conclusion

This result shows that people who use the same earbud with the same voicing will experience very different pressure responses at their DRP. An earbud voiced on a mannequin to achieve a certain target response at that mannequin's DRP will produce a response at an individual's DRP that is different by up to 7.8 dB below 5 kHz and even higher above that frequency.

If the goal is to achieve the same frequency response at each individual's DRP, then a personalized correction based on equalization at NFM can be employed, but it is only effective at reducing the error at DRP below 1800 Hz to within 3 dB. To correct for the response at DRP one will have account for differences in the transfer functions G_i. How to implement such

a personalized correction is beyond the subject of this paper. The goal of this work is to provide insight into how large the errors could be when approximating an individual's transfer function with that obtained on a mannequin.

2.3 Modeling of Far-Field-Source Response at DRP in Individual Ears

2.3.1 Modeling Setup

We next modeled the entire heads including pinnae and ear canals for the 10 individuals in the IHA database. In order to keep computational cost low we limited the geometry to the head and part of neck only and did not include the torso. As a source, we modeled a point source at 3 m directly in front of the heads. The strength of the point source was chosen to produce a pressure amplitude of 1 Pa at 1m distance for all frequencies. The head was assumed to be in anechoic conditions, i.e. no room effects such as reflections from floor, walls, or ceiling was included. The acoustic property of the entire head was set to have a surface impedance of skin (see Figure 2), with exception of ear drum which was again given the ear drum impedance from [9] implemented in the COMSOL Multiphysics software. The simulation was run in 12th-octave steps between 100 Hz and 8 kHz.

We also modeled the head and neck of the KEMAR mannequin outfitted with 711 couplers at the ear canal ends. For this model, we assumed the that the anthropometric pinnae have the same surface impedance as human skin. The rest of the head was modeled as rigid surfaces. The details of the 711 coupler remained the same as for the simulations with the earbud discussed above.

2.3.2 Results

The pressure distribution around the individual heads was calculated and we extracted $F_i(f)$, the individual SPL responses at DRPs due to the flat frequency response of the source. The SPL fields on the surfaces of the KEMAR mannequin and one individual in the IHA database is shown in Figure 11, and the SPL responses at DRP are plotted in Figure 12.

Rather than trying to achieve a given fixed SPL target curve at DRP like in Section 2.2, we now want to voice the earbud to be maximally transparent, i.e. we attempt

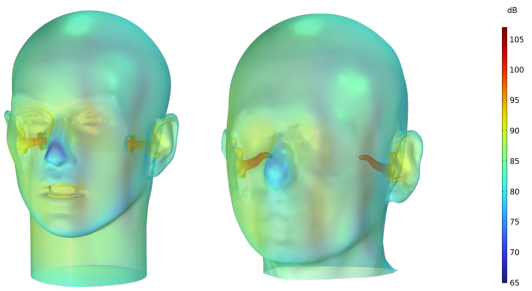

Fig. 11: SPL distribution around the head of the Kemar mannequin (left) and an individual in the IHA database (right) resulting from a source located at 3 m in front of the head.

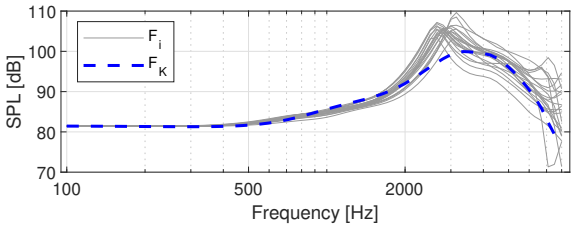

Fig. 12: Responses F_i at DRP due to source in far field 3 m in front of the heads. Gray lines correspond to individuals in IHA database, dashed blue line is for the KEMAR mannequin.

to achieve $D_i(f)$ at DRP that is equivalent to the SPL produced by the far-field source $F_i(f)$. We can again use the mannequin setup to determine a target $M_F^*(f)$ at NFM based on the $F_K(f)$, the pressure response at DRP due to the far-field source.

$$M_F^* = F_K - G_K \tag{8}$$

If the same personal equalization

$$Q_{F_i} = M_F^* - M_i \tag{9}$$

is employed as in Section 2.2, then the SPL at DRP in an individual's ear will again be

$$D^{Q_{F_i}} = M_F^* + G_i \tag{10}$$

which can now be written as

$$D^{Q_{F_i}} = F_K - G_K + G_i \tag{11}$$

However, if we now want to estimate the error between $D_i^{Q_F}$ and F_i, we end up with an error estimate of

$$\varepsilon_F = (F_i - F_K) + (G_K - G_i), \tag{12}$$

which is not only dependent on the differences of the far-field responses at DRP F, but also on the difference in transfer functions G_i and G_K. The question remains how the differences are related. The difference in F_K and F_i to some degree due to different head-related transfer functions (HRTF), but also due to differences in geometry of the ear canals. The differences in G are entirely due to differences in geometry of the ear canals.

If $(F_i - F_K)$ is equal to $(G_i - G_K)$, i.e. if the spectral differences between KEMAR mannequin and individuals are the same in the far-field scenario and the in-ear earbud scenario, then errors would cancel themselves and a perfect replication of the far-field response could be achieved for each individual without knowledge of the individuals' HRTF functions or transfer functions G_i. Unfortunately, head and ear canal geometries are not really correlated, so the errors get compounded. This was expected and demonstrated in measurements by Hammershøi et al.[10].

To get an idea of how strongly HRTF's and ear canal geometry influence the pressure response at DRP, we have plotted ε_F for the 100 different setups in Figure 13.

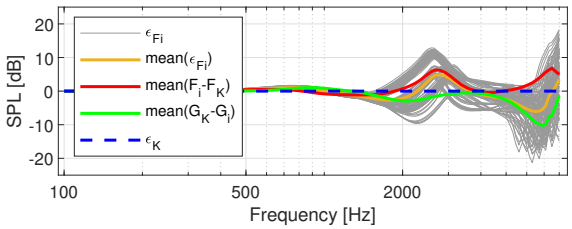

Fig. 13: Error of SPL at DRP compared to individuals' own far-field response. Earbuds were voiced on KEMAR mannequin to match its far-field response at DRP, resulting in a target response at NFM. Indiviual inserts were then equalized to match that response at NFM. In addition, we also plotted the means of the two contributions to the total error, $(F_i - F_K)$ and $(G_K - G_i)$.

2.3.3 Conclusion

Attempting to recreate a SPL response of a far-field source using estimates based on mannequin measurements results in large errors above about 1.5 kHz. Between 1.5 kHz and 4 kHz there are large differences between F_i and F_K, as well as between G_i and G_K. While the means of those differences have opposing signs, their shape is not exactly symmetric. While the difference of the means could be in the order of ± 5 dB up to 5 kHz, for some individuals the error can reach 12 dB. Above 5 kHz the errors are even bigger due to large differences of G_i compared to G_k.

Adding in larger sections of the torso in the simulations would not change this result. In fact, the sum of the differences could be even bigger, since it can be argued that the relation between between F_i and G_i gets less correlated the larger the included torso geometry for the far-field calculations. It is conceivable that head size are somewhat correlated to ear canal length and shape, but it is not unreasonable to assume that body size and ear canal shape have a much weaker correlation.

In order to truly create individually transparent headphones, it becomes imperative to have a better estimation of the individual HRTF as well as a better estimation of $G_i(f)$.

3 Summary

Using publicly accessible data on ear and head geometry of real-world individuals and FEM simulations,

we have demonstrated that the differences in ear canal geometry can lead to relatively large differences in SPL responses at the ear drum while wearing in-ear headphones. In a first study we showed the spread of the individual transfer functions $G_i(f)$ of the pressure at a microphone inside an in-ear headphone to the pressure at the eardrum. Comparing these transfer functions to the transfer function of the same earbud in a KEMAR mannequin, we showed that significant differences of up to 7.8 dB can exist in the band between 1.5 kHz and 5 kHz. Above 5 kHz the differences can reach 20 dB and become intractable. If the goal is to control the sound pressure at the eardrum, then a better estimation of the personal transfer functions will be needed. Even if the sound pressure is corrected to a target at the microphone inside the earbud, the estimate of the pressure at the eardrum will be off by the error of the estimated transfer function. In a second study we showed that creating transparent headphones based on in-ear measurements also requires a better estimate of the transfer function, in addition to improved estimates of the HRTF for each individual. Creating an in-earbud target pressure response based on transparency for a mannequin setup will result in errors of up to 12 dB below 5 kHz for individuals whose heads and ear canal geometries do not closely resemble those of the mannequin.

Both studies clearly suggest that better estimates of the transfer functions from in-earbud microphone to eardrum are necessary to have better control over the pressure at the eardrum. Simulations can be used to generate better insight into those transfer functions because the pressure at the eardrum is directly accessible. Future work will focus on how to generate such better estimates of the transfer functions in order to provide more precise control of the pressure at the eardrum regardless of the individual's ear canal shape.

4 Acknowledgements

Samsung Electronics and Samsung Research America supported this work. The authors would like to thank the entire staff of Samsung's US Audio Lab who helped with all aspects of this research, offered insightful suggestions, and contributed to this work.

References

[1] Olive, S. E., "The Perception and Measurement of Headphone Sound Quality: What Do Listeners Prefer?" *Acoustics Today*, 18(1), p. 58, 2021, ISSN 15570215, doi:10.1121/at.2022.18.1.58.

[2] Olive, S. E., "The Need for Headphone Personalization," in *ALTI Audio & Loudspeaker Technologies International*, 2023.

[3] Hoffmann, P. F., Christensen, F., and Hammershøi, D., "Insert earphone calibration for hear-through options," in *Proceedings of the AES International Conference*, Helsinki Finland, 2013, ISBN 9781629933283.

[4] Celestinos, A., "Personalized headphone equalization," U.S. Patent Application US11206003B2, 2020.

[5] Multiphysics, C., "Generic 711 Coupler: An Occluded Ear-Canal Simulator," *https://www.comsol.com/model/generic-711-coupler-8212-an-occluded-ear-canal-simulator-12227*, pp. 1–30, 1981, ISSN 0308-0110.

[6] International Electrotechnical Commission, *IEC 60318-4:2010 (Simulators of human head and ear – Part 4: Occluded-ear simulator for the measurement of earphones coupled to the ear by means of ear inserts)*, 1.0 edition, 2010, ISBN 9782889101214.

[7] Roden, R. and Blau, M., "The IHA database of human geometries including torso, head and complete outer ears for acoustic research," in *Proceedings of 2020 International Congress on Noise Control Engineering, INTER-NOISE 2020*, pp. 4226–4237, Seoul, Korea, 2020, ISBN 9788994021362.

[8] Roden, R. and Blau, M., "The IHA database of human geometries including torso, head and complete outer ears for acoustic research," 2020, doi: 10.5281/zenodo.5528766.

[9] Hudde, H. and Engel, A., "Measuring and Modeling Basic Properties of the Human Middle Ear and Ear Canal. Part III: Eardrum Impedances, Transfer Functions and Model Calculations," *Acustica*, 84(6), pp. 1091–1109, 1998, ISSN 00017884.

[10] Hammershøi, D. and Møller, H., "Sound transmission to and within the human ear canal," *The Journal of the Acoustical Society of America*, 100(1), pp. 408–427, 1996, ISSN 0001-4966, doi: 10.1121/1.415856.