Audio Engineering Society

Convention Express Paper 239

Presented at the AES 156th Convention
2024 June 15–17, Madrid, Spain

This Express Paper was selected on the basis of a submitted synopsis that has been peer-reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This Express Paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Personalized Equalization of Sound Pressure at Eardrum With Insert Earbuds

Andri Bezzola[1], Adrian Celestinos[1], and Ema Souza Blanes[1]

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

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

ABSTRACT

Ear-canal geometries can vary significantly from human to human. The differences in geometry (length, width, etc.) can affect the resulting pressure at the eardrum. Headphones are often designed to match a target pressure-response curve at the drum reference point of some ear simulator. However, due to the individual differences the pressure response from an insert earbud at the eardrum of individual human ears can differ by several dB compared to the response of the simulator. In addition to ear canal geometry, the quality of the fit between the earbud tip and ear canal influences the pressure response at low frequencies.

In this work, we introduce an equalization method that uses measurements at a near-field microphone in the earbud tip to estimate the personalized transfer function from near-field microphone to the eardrum. Knowledge of this personalized transfer function allows for personalized equalization of the sound pressure at the eardrum. The presented method takes into account the personal ear canal geometry as well as the quality of the seal between earbud tip and ear canal.

We conducted a listening test to subjectively compare the personalized equalization algorithm against equalization using a mannequin-derived transfer function, and against the default voicing of the earbuds without equalization. The results show that the personalized equalization was favored on average with statistical significance.

1 Introduction

In a previous paper [1] we demonstrated that the ear canal geometries of individual people are different enough to result in significant variations of pressure levels at the drum reference point (DRP). The resulting consequence of these differences in ear canal geometries means insert headphones that are tuned to a fixed equalization (EQ) on a mannequin will result in a different timbre for the individual users of the headphones. It has been suggested that a more personalized tuning of headphones can result in a better listening experience [2, 3, 4]. In order to achieve a more controlled pressure response at DRP, physiological differences of each ear need to be taken into account [5, 6]. Our first

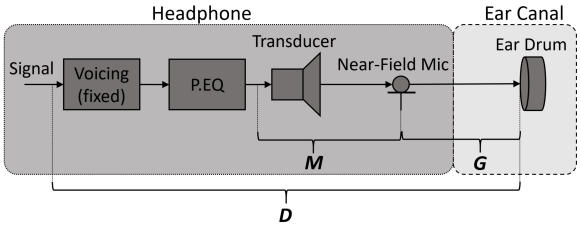


Fig. 1: A block diagram of the signal chain. P.EQ is the personalized EQ to control the transfer function from signal voltage to pressure at the eardrum. D, M, and G denote the transfer functions used throughout this work.

paper had shown the differences in sound pressure levels at DRP by means of finite element method (FEM) simulations using ear canal geometries obtained via magnetic resonance imaging (MRI) scans [7]. From these simulations we were able to extract personalized transfer functions $G_p(f)$ that correlate the sound pressure level (SPL) per volt measured in dB at a near-field microphone (NFM) $M_p(f)$ to the SPL per volt at DRP $D_p(f)$:

$$G_p(f) = D_p(f) - M_p(f) \tag{1}$$

where the subscript p means *personalized*, and f is the frequency in Hz. A block diagram of the signal chain is shown in Figure 1.

With knowledge of $G_p(f)$ it becomes possible to predict $D_p(f)$ by means of a transfer function measurement at the NFM. This is trivial enough in simulations, but it becomes a great challenge in real human subjects where $D_p(f)$ is not easily accessible [8]. This work presents a method to estimate $G_p(f)$ - and by extension $D_p(f)$ - by means of measuring $M_p(f)$ and extracting some key value from it. In a first section we explain how we are able to correct for differences in ear canal geometries. In a second section we explain how non-perfect fits can affect the low-frequency response at the eardrum and how to correct for it.

2 Compensating for Differences in Ear Canal Geometries

2.1 Estimating $G_p(f)$ from Measurements of $M_p(f)$ at NFM

In order to be able to control the SPL at DRP, it is imperative to have a good estimate of the transfer function $G_p(f)$. In this section we introduce a method to estimate $G_p(f)$ from measurements at NFM $M_p(f)$.

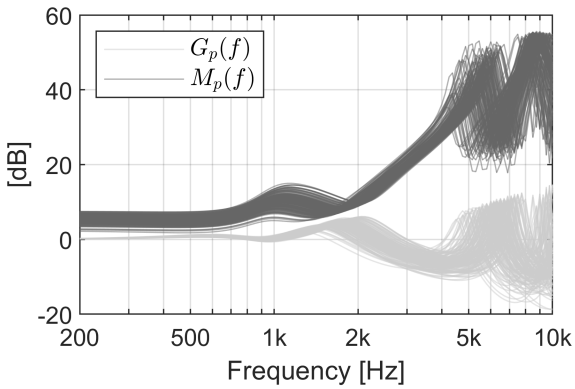


Fig. 2: $G_p(f)$ and $M_p(f)$ simulated with perfectly sealing insert earbud on 31 ear scans with five different insert positions each.

To model a sample of the population we obtained Computer Tomography (CT) scans of 31 ears from Samsung Medical Center in Korea. The sample of the 31 ears included scans of the right ear from 18 females and 13 males between the ages of 11 and 68. The mean age of the subjects was 37.3 years old, and the median age was 36 years old. We simulated the acoustic response of the Galaxy Buds2 Pro earbud inserted into the ear canal geometries at five different positions per ear, resulting in an ensemble of 155 simulated cases. The simulation setup was analogous to the simulation procedure in our previous paper [1] using the software COMSOL Multiphysics [9]. The simulation results for $G_p(f)$ and $M_p(f)$ are plotted in Figure 2. We like to note that the curves for $M_p(f)$ are not representative of real earbuds in ears, because in the simulation we assumed a constant displacement of the diaphragm as a function of frequency. The transfer function of the transducer is not captured in the simulations.

From the simulation results it becomes apparent that $M_p(f)$ and $G_p(f)$ are relatively smooth up to about 4 kHz. At higher frequencies the resonance peaks and dips are increasingly sharp and would be very hard to equalize. For this reason, we limited our attempts to achieve a personalized equalization of the earbuds below 4 kHz. We like to note that the Samsung Medical Center CT data yielded results that were very much in line with the results obtained on the geometries of the IHA database [7] published in [1].

From the simulation data shown in Figure 2, we can see that $G_p(f)$ has main peaks at frequencies f_G^{max} between

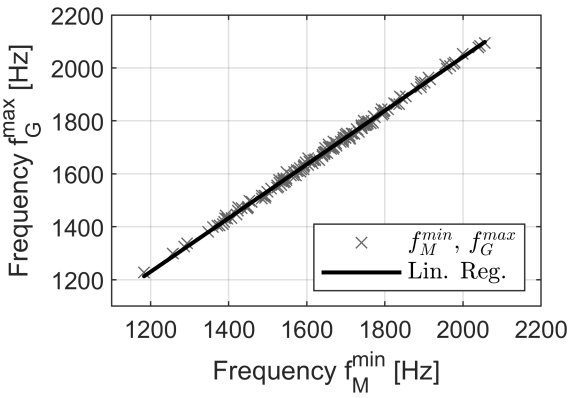


Fig. 3: Correlation between f_G^{max} with f_m^{min}. The linear regression has a zero-intercept of -8.3 and a slope of 0.98.

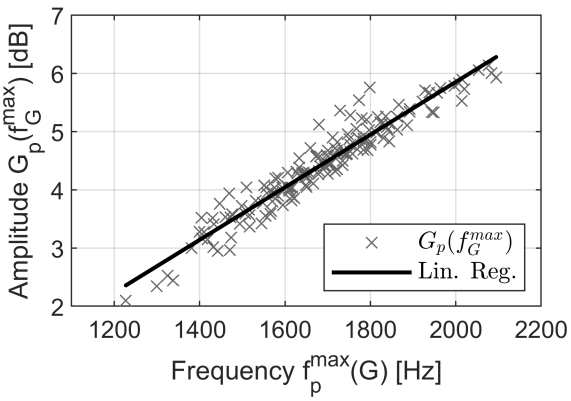


Fig. 4: Correlation between max gain of $G_p(f)$ and f_G^{max}. The linear regression has a zero-intercept of -3.2 and a slope of 4.5×10^{-3}.

1200 Hz and 2400 Hz. Furthermore, we can also see that while $M_p(f)$ also has pronounced peaks between 800 Hz and 1600 Hz, these peaks do not appear to correlate with any major peaks or dips in $G_p(f)$. However, there is also a pronounced dip in $M_p(f)$ at frequencies f_M^{min} between 1200 Hz and 2400 Hz which do appear to correlate with the peaks in $G_p(f)$. If we plot the frequencies of the dips in $M_p(f)$ against the frequencies of the peaks in $G_p(f)$, then it appears that the two frequencies are indeed very much linearly correlated as shown in Figure 3.

Additionally, the amplitude of $G_p(f)$ at the first peak also appear to follow a linear trend. The relationship between G_p^{max} and f_G^{max} is plotted in Figure 4. While the linearity of the relationship is not as strong as the relationship between $G_p(f_p^{max})$ and f_M^{min}, the largest error between $G_p(f_p^{max})$ and the linear regression is only 0.82 dB.

With the above regressions, we can now approximate $G_p(f)$ by an estimate $\widehat{G_p}(f)$ via measuring the personal $M_p(f)$ and finding f_M^{min}. In the next sections we will describe how to efficiently approximate $G_p(f)$ by using a parametric, generic function $G_0(f)$ and adjusting the parameters based on f_M^{min}.

2.2 Parametric Transfer Function $G_0(f)$

The measurements on the G.R.A.S. KEMAR mannequin $G_K(f)$ as described in [1] provided the basis to define a parametric base function $G_0(f)$. We chose to use four biquad filters to approximate $G_K(f)$. Figure 5 shows the $G_K(f)$ along with its parametric approximation $G_0(f)$ that was constructed using four second-order biquadratic filters. The center frequencies $f_c^{(j)}$, gains $g^{(j)}$, and Q-factors $Q^{(j)}$ for each of the four filters was found empirically and are listed in Table 1.

Table 1: Center frequencies f_c, gains g, and Q-factors Q for the 4 biquad filters of $G_0(f)$.

j	$f_c^{(j)}$ [Hz]	$g^{(j)}$ [dB]	$Q^{(j)}$ [-]
1	670	0.8	0.5
2	1650	4.7	1.1
3	5850	-11.0	0.5
4	7870	8.0	3.5

Biquad 2 is the main filter that controls the peak in $G_0(f)$ around 1600 Hz. By keeping the center frequency and the gain of this biquad variable, we can now construct a personalized estimate $\widehat{G_p}(f)$ from $G_0(f)$ by measuring f_M^{min}. From Figure 2 it can be seen that the individual $G_p(f)$ appear shifted along the frequency axis. We must thus also correct for the center frequencies of biquads 3 and 4, along with the center frequency and gain of biquad 2. We have established the following rules to estimate $\widehat{G_p}(f)$:

- estimate $f_G{}^{max} = 8.3 + 0.98 f_M^{min}$
- estimate $g_p{}^{(2)} = -3.2 + 4.5 \times 10^{-3} f_G{}^{max}$
- $f_c^{(2)} = 1.1 f_M^{min}$
- $g^{(2)} = g_p{}^{(2)} + 4.0$

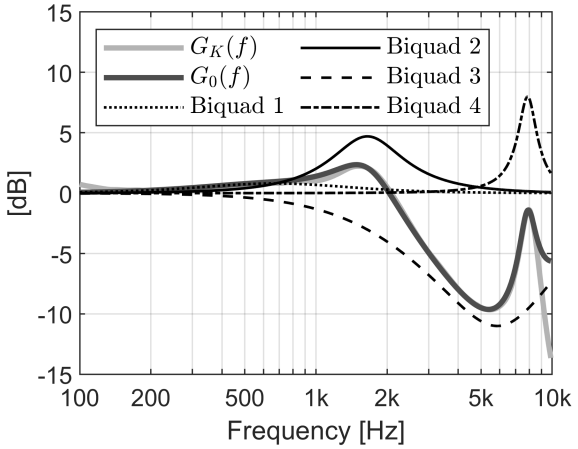


Fig. 5: $G_K(f)$ and its parametric version $G_0(f)$, along with the four biquad filters that were chosen to construct $G_0(f)$.

- $f_c^{(3)} = 2.7 f_M^{min}$
- $f_c^{(4)} = 3.9 f_M^{min}$

The correction factors in the steps above were found empirically for an optimal fit. The factors of 1.1 in $f_c^{(2)}$ and 4.0 in $g^{(2)}$ are used to correct for the effect of the strong notch caused by the third biquad. The other two factors for $f_c^{(3)}$ and $f_c^{(4)}$ are used to ensure that the center frequencies of the third and fourth biquad scale with f_M^{min}, i.e. they are shifted proportionally to f_M^{min} in the log-frequency plot. The remaining parameters for $f_c^{(1)}$, $g^{(1)}$, $g^{(3)}$, $g^{(4)}$, and all $Q^{(j)}$ are kept constant.

2.3 Results

Two examples of the personal true $G_p(f)$ and their estimates $\widehat{G_p}(f)$ based on the measurement of f_M^{min} are shown in Figure 6. The first example has an f_p^{min} of 1220 Hz and the second example has an f_p^{min} of 2039 Hz. From these examples of a very short and a very long ear canal we can see that the approximations $\widehat{G_p}(f)$ are within ± 2 dB from the true transfer functions $G_p(f)$ below 4 kHz.

Figure 7 shows error between $G_p(f)$ and their estimate $\widehat{G_p}(f)$ and compares them against the errors that would result from using the mannequin transfer function $G_K(f)$ instead of the personalized estimate. As previously presented in [1], the error made in the assumption of the transfer function will directly translate

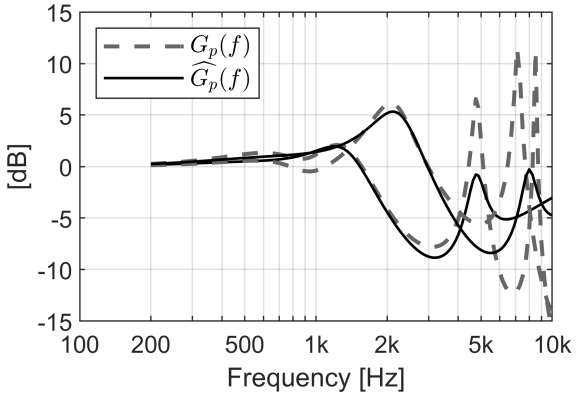


Fig. 6: Two examples of true $G_p(f)$ (from FEM simulation), and their estimated $\widehat{G_p}(f)$ based on measurement of f_p^{min}.

to the error of the pressure estimate at the eardrum. Between about 1 kHz and 4 kHz the error is significantly lower when using a personalized estimate of the transfer function rather than a mannequin's transfer function. Above 4 kHz the personalized estimate transfer function does not provide any benefit. The high-Q resonances above 4 kHz do not lend themselves for equalization, so we propose to limit the equalization to frequencies below 4 kHz.

3 Compensating for Low-Frequency Leakage Due to Imperfect Fits

3.1 Low-Frequency Loss Due to Non-Perfect Seal Between Earbud and Ear Canal

Less-than-perfect fits of the earbud inside the ear canal result in a loss of low-frequency energy at DRP. This loss is also registered at the NFM, but the amount of energy loss is not equal at both points. To estimate the amount of loss of energy at DRP, we again need to infer it from a measurement at NFM.

We measured the earbud in the KEMAR mannequin with 14 more-or-less-than-ideal fits between earbud and ear canal. The resulting responses at NFM and DRP along with the individual transfer functions are shown in Figure 8. The fits ranged from tight and well-sealing to rather loose. Below 600 Hz the levels of $M_p(f)$ and $D_p(f)$ dropped with decreasing fit quality. However, $D_p(f)$ did not drop by the same amount as $M_p(f)$. With increasing leakage $D_p(f)$ dropped more

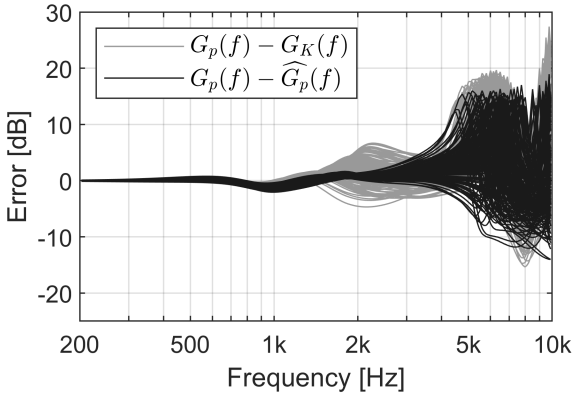


Fig. 7: Differences between true $G_p(f)$ and their personalized estimates $\widehat{G_p}(f)$ compared to the differences between $G_p(f)$ and the mannequin transfer function $G_K(f)$.

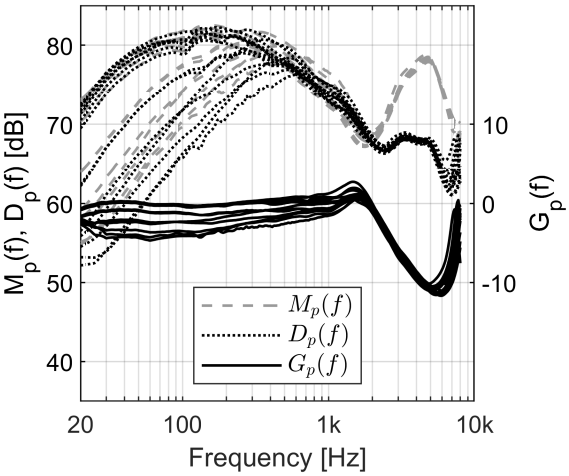


Fig. 8: 14 measurements of $M_p(f)$, $D_p(f)$, and $G_p(f)$ on the KEMAR mannequin outfitted with a Galaxy Buds2 Pro earbud. Losses at low frequencies increased with decreasing fit quality.

than $M_p(f)$, which resulted in a drop off of $G_p(f)$ below 0 dB at lower frequencies. This difference will of course impact the personalized equalization and will need to be accounted for.

3.2 Estimating Low-Frequency Drop of $G_p(f)$ Due to Imperfect Fits

To estimate the amount of low-frequency drop of $G_p(f)$ with less-than-ideal fits we decided to take the measurements of $M_p(f)$ at two frequencies: f_1 = 46.8 Hz and f_2 = 200 Hz. We then plotted ΔG_{LF}, the mean drop of $G_p(f)$ between 40 Hz and 100 Hz against the SPL-difference of $\Delta M_{2,1} = M_p(f_2) - M_p(f_1)$. The results are shown in Figure 9. The results suggest that the relationship between $\Delta M_{2,1}$ and ΔG_{LF} can be approximated by a quadratic function. A curve-fit of the data to a quadratic function yielded a prediction $\widehat{\Delta G_{LF}}$ of

$$\widehat{\Delta G_{LF}} = -0.044\Delta M_{2,1}^2 - 0.24\Delta M_{2,1} - 0.40. \quad (2)$$

Finally, we implemented the correction at low frequencies for $\widehat{G_p}$ with a biquad shelving filter with gain set to $\widehat{\Delta G_{LF}}$. A center frequency of $f_c^{(LF)}$ = 1493 Hz and a Q-factor of $Q^{(LF)}$ = 0.246 resulted in a minimal root-mean-square error between the predicted and measured transfer function for the mannequin in the band between 20 Hz and 2000 Hz.

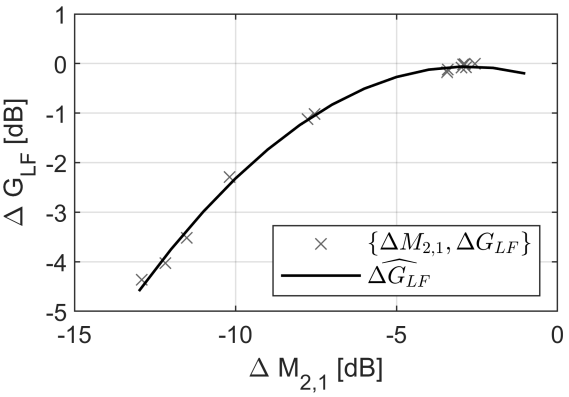


Fig. 9: ΔG_{LF} plotted versus $\Delta M_{2,1}$, and the quadratic approximation $\widehat{\Delta G_{LF}}$.

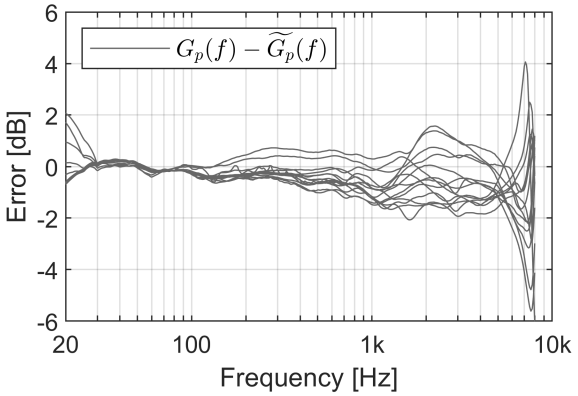


Fig. 10: Error between the estimated $\widetilde{G_p}(f)$ and the measured $G_p(f)$ on the KEMAR mannequin with imperfect fits of the earbud.

3.3 Results

The new modified transfer function estimate $\widetilde{G_p}(f)$ that takes into account the low-frequency leaks can be calculated as

$$\widetilde{G_p}(f) = \widehat{G_p}(f) + LSHV(f_c^{(LF)}, \widehat{\Delta G_{LF}}, Q^{(LF)}), \quad (3)$$

where $LSHV(f_c^{(LF)}, \widehat{\Delta G_{LF}}, Q^{(LF)})$ is the amplitude of the frequency response of the low-shelving filter in dB.

The error between the estimated $\widetilde{G_p}(f)$ and the measured $G_p(f)$ for the 14 fittings is shown in Figure 10. The resulting error stayed below ± 2 dB between 30 Hz and 6000 Hz. The slight increase in error below 30 Hz is due to poor signal-to-noise ratio.

This estimation of low-frequency loss of SPL at DRP due to imperfect fits between earbud and ear canal has been shown to work well for typical insert positions of the earbud into the mannequin. If the fit becomes too loose and the low-frequency leak becomes so severe that it will affect $M_p(f)$ around f_p^{min}, the personalization algorithm falls apart. Hence we suggest this personalized equalization algorithm only for cases with good to moderate seal between the earbud tip and the ear canal.

4 Full Personalized Equalization Algorithm.

With a better estimate of the transfer function of the pressure at the near-field microphone to the pressure at the eardrum at frequencies between 20 Hz and 4 kHz, the personalized equalization procedure can be summarized as follows:

1. Establish a desired pressure response target at the eardrum $T_{DRP}(f)$.
2. Insert earbud and determine f_p^{min} based on measurement at near-field microphone $M_p(f)$.
3. Generate personalized transfer function estimate $\widehat{G_p}(f)$ based on measurement of f_p^{min} as detailed in Section 2.
4. Calculate the leakage loss factor $\Delta M_{2,1}$ by measuring the SPL difference at two frequencies.
5. Estimate transfer function $\widetilde{G_p}(f)$ that takes into account leakage as detailed in Section 3.
6. Calculate personalized target response function at near-field microphone
$$T_p^{NFM}(f) = T_{DRP}(f) - \widetilde{G_p}(f). \quad (4)$$
7. Calculate equalization function
$$P.EQ(f) = T_p^{NFM}(f) - M_p(f). \quad (5)$$
8. Employ equalization method of choice.

5 Listening Test

A listening test was conducted to assess if equalization using the personalized estimate $\widetilde{G_p}(f)$ is preferred over equalizaton using a generic $G_K(f)$ as outlined in [10] is preferred. Four audio samples (two music tracks, one speech sample and one film excerpt) were used for assessors to compare the following test conditions: *No-P.EQ*, *P.EQ with* G_K, and *P.EQ with* $\widetilde{G_p}$. The test stimuli were presented on insert earbuds using a custom designed Max/MSP interface.

The *No-P.EQ* condition consisted of the sound samples played back directly through the earbuds, while equalization was applied to the other two conditions using personalized and KEMAR-derived G functions, respectively. The earbuds used in the test had wire-leads attached to the transducer terminals as well as to the NFM to record $H_p(f)$. The personalized estimate $\widetilde{G_p}(f)$ was obtained using a custom MATLAB script

reading $M_p(f)$ from the earbuds' NFM. The personalized equalization was performed using a finite impulse response (FIR) filter of length 2048.

Listeners were instructed to insert the earbuds into their ears with a fit that would be slightly loose, but stable enough to remain the same for the duration of the test. The importance of staying still to avoid changing the fit of the earbuds was communicated by the test administrator. The measurement was then run once using a pink-noise signal. Personalized filter coefficients were then calculated and applied in the Max/MSP software.

The test comprised eight trials, where listeners compared and rated the test conditions according to their preference for sound equalization for all four samples, with one replicate, on a 0-100 rating scale. The Max/MSP interface allowed for randomization of the order of presentation of the test samples, as well as the arrangement of the conditions (labeled A, B and C) for each trial and each listener. Ratings were given on sliders and listeners were able to leave comments. Assessors were also able to control the playback and loop segments of the audio.

16 participants, including 11 experienced listeners, completed the preference rating task. Results show that both conditions with equalization were rated higher than the *No-P.EQ* condition, and that the *P.EQ with* $\widetilde{G_p}(f)$ condition was rated higher on average than the *P.EQ with* G_K condition, as shown in Figure 11.

A two-way ANOVA confirmed that there was a significant effect of the condition ($p < 0.0001$) and no significant effect of the audio sample ($p = 0.128$) or sample/condition interaction ($p = 0.745$) factors. Pairwise two-sided t-tests revealed that the difference between mean scores was statistically significant for each pair of condition: $p = 0.00015$ for the condition of *No-P.EQ* vs *P.EQ with* G_K; $p = 1.9\text{e-}09$ for the condition of *No-P.EQ* vs *P.EQ with* $\widetilde{G_p}$; and $p = 0.00010$ for the condition of *P.EQ with* $\widetilde{G_p}$ vs *P.EQ with* G_K. This result confirms that the use of a personalized equalization using $\widetilde{G_p}$ was preferred on average by the panel of assessors.

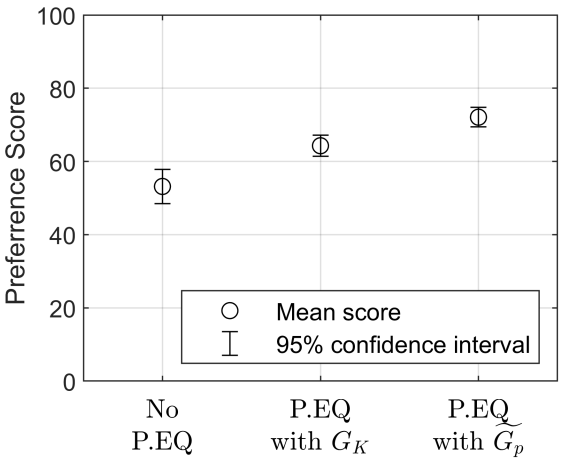


Fig. 11: Listening test results showing mean preferrence scores and their 95% confidence intervals.

6 Conclusion

In summary, we have introduced a method to improve the accuracy of control over the sound pressure response at the ear drum below 4 kHz when using insert earbuds with moderate to good sealing against the ear canal. The errors between 1 kHz and 4 kHz are reduced significantly by using a approach that estimates the transfer function from pressure at NFM inside the earbud to the pressure at DRP. This estimate is based on the observation that individual transfer functions have a strong relationship to the frequency of a minimum in the pressure response function at the near-field microphone. A parametric approach was employed to generate a better estimate of the personal transfer function that compensates for the user's personal ear canal geometry.

The loss of low-frequency energy with imperfect fits of the earbud tip against the ear canal was counteracted with measurements at NFM at two frequencies. The ratio of the pressure at those two frequencies informed a wide-band low frequency loss that was modeled with a low shelving biquad filter. The full personalized equalization algorithm reduces the error of the prediction of the SPL at the eardrum by approximating the transfer function of pressure at NFM to pressure at DRP.

We tested the proposed algorithm against an equalization method using a generic ear-simulator-derived transfer function, and on average, listeners preferred the personalized equalization with statistical significance.

7 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] Bezzola, A., "In-Ear Headphones on Ear Canal Simulator vs Real Human Ear Geometries: Quantifying the Differences with Simulations," in *Audio Engineering Society Convention 155*, Audio Engineering Society, New York City, 2023.

[2] Vogl, S. and Blau, M., "Individualized prediction of the sound pressure at the eardrum for an earpiece with integrated receivers and microphones," *The Journal of the Acoustical Society of America*, 145(2), pp. 917–930, 2019, ISSN 0001-4966, doi:10.1121/1.5089219.

[3] 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.

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

[5] Stinson, M. R. and Lawton, B. W., "Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution," *Journal of the Acoustical Society of America*, 85(6), pp. 2492–2503, 1989, ISSN NA, doi:10.1121/1.397744.

[6] 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.

[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] Chan, J. C. K. and Geisler, C. D., "Estimation of eardrum acoustic pressure and of ear canal length from remote points in the canal," *The Journal of the Acoustical Society of America*, 87(3), pp. 1237–1247, 1990, ISSN 0001-4966, doi:10.1121/1.398799.

[9] COMSOL, A., "COMSOL Mulitphysics v6.2," 2024.

[10] Celestinos, A., "Personalized headphone equalization," U.S. Patent 11 206 003 B2, Dec. 21, 2021.