PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

Time-variant acoustic front-end measurements of active noise cancellation headphones

Johannes FABRY, David HILKERT, Stefan LIEBICH, Peter JAX

Institute of Communication Systems, RWTH Aachen University, Germany, {fabry, hilkert, liebich, jax}@iks.rwth-aachen.de

Abstract

A robust design of active noise cancellation (ANC) headphones by means of digital signal processing requires a deep understanding of the underlying acoustic front-end. Particularly knowledge about the primary path, which is commonly defined as the transfer path between the outer and inner microphone of the headphone, and the secondary path, which is the transfer path between the loudspeaker and the inner microphone, is required to designing an ANC system. These paths may vary due to e.g. fitting of the headphone, the physiology of the user or the direction of arrival of ambient noise.

In this contribution, we present the results of a measurement series. The objective is to examine acoustic paths for their intra-person variances considering jaw movement, head rotation and refitting of the headphone, as well as inter-person variance between different subjects. The measurements were conducted in an anechoic chamber with 25 participants from age 21 to 61. Furthermore, the implications on the performance for time-invariant ANC solutions will be considered and the benefit of calibrating the secondary path will be investigated.

Keywords: Active noise control, device related transfer functions

1 INTRODUCTION

These days, a lot of applications and products based on active noise cancellation (ANC) technology are emerging. Especially ANC headphones have risen in popularity over the last few years, due to the ever-growing amount of environmental noise [12].

An implementation of complex adaptive algorithms is still not feasible, due to strict power and form-factor constraints of consumer ANC headphones. Hence, most designs rely on time-invariant digital feed-forward (FF) filters and feed-back (FB) controllers. The optimal solution to feed-forward filters for ANC is given by the causal Wiener filter, which is a well-studied topic [7, 3, 5, 4]. Also, numerous publications cover the topic of feed-back controller design for ANC applications [11, 8]. In order to obtain a robust feed-back system, the uncertainties of the secondary path of the ANC headphone need to be considered.

Hence, to obtain a well-performing and robust design for ANC headphones a good understanding of the electro-acoustic front-end is required. To receive this deep understanding, several measurement series have previously been conducted by the authors, which focus, e.g., on the direction of arrival dependency of ANC headphones [10] and the occlusion effect [9]. However, these measurements did not specifically analyze the influence of the user and use-cases on the transfer functions of an ANC headphone.

This contribution presents the results of a measurement series to identify how knowledge of individual device related transfer functions (DRTFs) can improve the performance of ANC headphones. We measured time-variant DRTFs of 25 female and male participants in the age of 21 to 61 years. To obtain a database that covers a wide spectrum of transfer functions, the participants were asked to go through a scripted sequence of different jaw positions, head rotations and headphone fittings.

This article is structured as follows. In Sec. 2 we describe the functional principle of ANC headphones. We then elaborate on the measurement setup, execution and post-processing in Sec. 3. In Sec. 4 we evaluate the measurements. We conclude the paper by discussing the implications on the performance of ANC headphones in Sec. 4.2.

2 ACTIVE NOISE CANCELLATION

The goal of active noise cancellation (ANC) is to create a zone of silence in which ambient noise is attenuated. This is achieved by the principle of destructive interference of sound waves. In the case of ANC headphones, the zone of silence is defined as the volume inside the auditory canal. To create an anti-noise, we require knowledge about the ambient noise. For that reason, ANC headphones are equipped with at least one microphone that measures the ambient noise. For the measurement series presented in this paper, we used the electro-acoustic front-end of the Bose QC20 headphones, without the ANC electronics. The left half of Fig. 1 illustrates the headphone. It contains the reference microphone $\mathrm{M_{ref}}$, directed to the opposing side of the head, the error microphone $\mathrm{M_{err}}$, directed towards the eardrum, as well as the internal loudspeaker $\mathrm{S_{int}}$ that emits the anti-noise and can also be used to playback multimedia signals. The behavior of the headphone is characterized by two acoustic transfer functions. The primary path $P(z)$ describes the transfer function between $\mathrm{M_{ref}}$ and $\mathrm{M_{err}}$, whereas the secondary path $S(z)$ describes the transfer function between $\mathrm{M_{err}}$ and $\mathrm{S_{int}}$. An acoustic feedback from $\mathrm{S_{int}}$ to $\mathrm{M_{ref}}$ can be neglected for sealed headphones due to a very low magnitude of the transfer function.

The anti-noise is generated by applying a digital filter $W(z)$ to the reference microphone signal $x(n)$ in the case of feed-forward ANC and by applying a digital controller $K(z)$ to the error microphone signal $e(n)$ in the case of feed-back ANC. Hence, the overall transfer function of a feed-forward ANC system $H_{\mathrm{FF}}(z)$ is

$$H_{\mathrm{FF}}(z) = P(z) - W(z)S(z), \tag{1}$$

whereas $P(z)$ and $S(z)$ also comprise the characteristics of the analog-to-digital and digital-to-analog converters. According to the Wiener criterion, we find the impulse response $\boldsymbol{w} \in \mathbb{R}^L$ of length L of $W(z)$ as

$$\boldsymbol{w} = \boldsymbol{\Psi}_{ss}^{-1}\boldsymbol{\varphi}_{ps}. \tag{2}$$

Here, $\boldsymbol{\Psi}_{ss}$ denotes the auto-correlation matrix of the secondary path and $\boldsymbol{\varphi}_{ps}$ denotes the cross-correlation vector between the primary and secondary path.

The transfer function $H_{\mathrm{FB}}(z)$ of a feed-back ANC system as described in [11, 8] is given by

$$H_{\mathrm{FB}}(z) = \frac{1}{1 + S(z)K(z)}. \tag{3}$$

In order to obtain a robust system, the uncertainties of the plant, here $S(z)$, need to be considered. The multiplicative uncertainty model is a common choice. It describes the plant by means of a multiplicative deviation from a nominal path $\overline{S}(z)$

$$S(z) = \overline{S}(z)\left[1 + \Delta(z)U(z)\right], \tag{4}$$

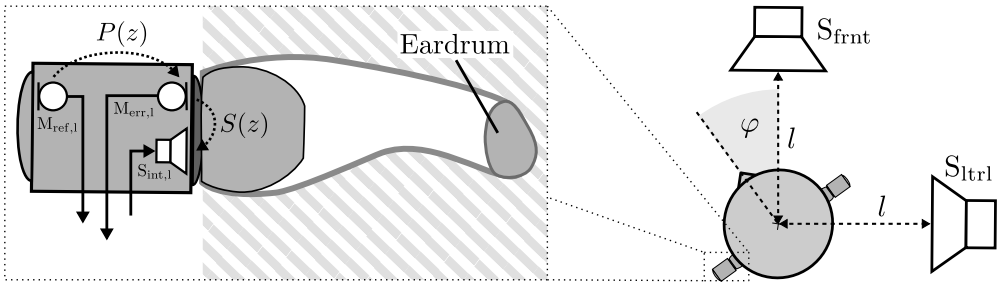

Figure 1. Measurement setup with participant, frontal and lateral external speakers in distance l of center as well as an illustration of the ANC headphone that was used.

with $|\Delta(z)| < 1$. We choose the nominal path as the average over a set of J measured plant transfer functions and calculate the uncertainty so that all plants from this set can be obtained from (4)

$$\overline{S}(z) = \frac{1}{J} \sum_{j=0}^{J-1} S_j(z) \tag{5}$$

$$U(z) = \max_{j \in [0..J-1]} \left| \frac{S_j(z)}{\overline{S}(z)} - 1 \right|. \tag{6}$$

Rafaely shows in [11] that the plant uncertainty has strong implications on the performance of a feed-back controller. We will have a closer look at the plant uncertainty in Sec. 4.2.

3 TIME-VARIANT MEASUREMENT

The following sub-sections elaborate on the acquisition of the data by means of time-variant measurements. We first describe the measurement setup. Then, we shortly introduce the topic of time-variant system identification and finally give some insights into the required post-processing steps.

3.1 Setup

The measurements were conducted in a chamber that is anechoic down to a frequency of 200 Hz. The right half of Fig. 1 illustrates the measurement setup. The participants were placed on a rotating chair in the center of the chamber, facing the frontal external speaker $\mathrm{S}_{\mathrm{frnt}}$ ($\varphi = 0$). Two Neumann KH120 speaker were positioned at a distance of $l = 90\,\mathrm{cm}$ in front and right of the center of the head. The participants wore a pair of Bose QC20 headphones. Fig. 2 shows the connection of the hardware components. We connected a RME Fireface UCX soundcard via USB to the host PC. A PreSonus DigiMax DP88 was connected to the soundcard via ADAT and served as a pre-amplifier for the microphone signals.

As a ground truth, the acoustic transfer functions of each participant were first measured using an exponential sweep. Afterwards, the time-variant measurement began and the participants were instructed to go through a scripted set of movements, as outlined in Fig. 3. The instructions were displayed on a TV screen. The movements involve different jaw positions, head rotations and headphone fittings. The head rotation is $\varphi = 0$ by default and increases to $\varphi = 90°$ within the last segment, so that all directions of arrival are covered with the two external speakers. Furthermore, the headphones were removed and refitted four times, twice instructed as a loose fit and twice as a normal fit. The headphones were initially inserted with a normal fitting.

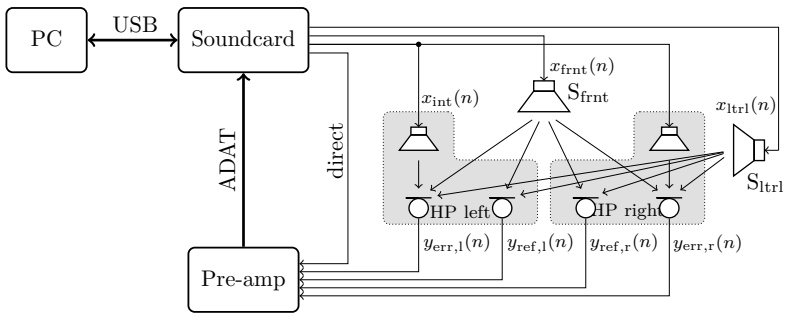

Figure 2. Digital and electro-acoustic connection between computer, soundcard, pre-amp as well as the loudspeakers and microphones

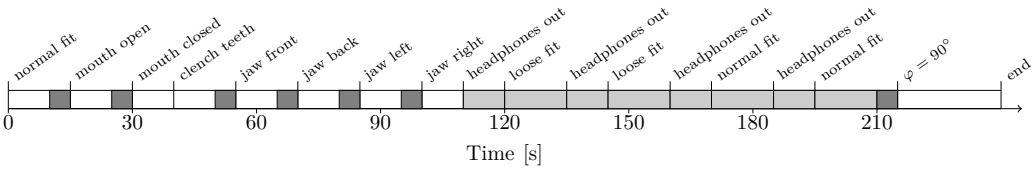

Figure 3. Sequence of positions instructed to the participants. Here, ▭ marks time-variant segments as transitions between different positions and ▭ marks segments where the headphones are refitted.

3.2 Time-Variant System Identification

Fig. 2 indicates that we measure a total of 10 time-variant impulse responses. We obtain one impulse response from each of the two external speakers to each of the four microphones as well as one impulse response from each of the internal speakers to the respective error microphone. As an example, the block diagram in Fig. 4 shows the multi-channel system model to identify the transfer functions from the external speakers to the left reference microphone. The measured response $y_{\mathrm{ref,l}}(n)$ is given as a superposition of the excitation signals convoluted with the respective impulse responses and a noise term $v_{\mathrm{ref,l}}(n)$

$$y_{\mathrm{ref,l}}(n) = \boldsymbol{h}_{\mathrm{ltrl,ref,l}}^{\mathrm{T}}(n)\boldsymbol{x}_{\mathrm{ltrl}}(n) + \boldsymbol{h}_{\mathrm{frnt,ref,l}}^{\mathrm{T}}(n)\boldsymbol{x}_{\mathrm{frnt}}(n) + v_{\mathrm{ref,l}}(n), \tag{7}$$

with $\boldsymbol{h}(n) = [h_0(n), h_1(n), \ldots, h_{L-1}(n)]^{\mathrm{T}}$ and $\boldsymbol{x}(n) = [x(n), x(n-1), \ldots, x(n-L+1)]^{\mathrm{T}}$. We use the normalized least mean squares (NLMS) algorithm to estimate the time-variant impulse responses as described in [1]. As an example, the update rule for the adaptive filter from the lateral external speaker to the left reference microphone is given by

$$\hat{\boldsymbol{h}}_{\mathrm{ltrl,ref,l}}(n+1) = \hat{\boldsymbol{h}}_{\mathrm{ltrl,ref,l}}(n) + \mu \frac{e_{\mathrm{ref,l}}(n)\boldsymbol{x}_{\mathrm{ltrl}}(n)}{\|\boldsymbol{x}_{\mathrm{ltrl}}(n)\|^2 + \|\boldsymbol{x}_{\mathrm{frnt}}(n)\|^2}, \tag{8}$$

with the error signal

$$\begin{aligned} e_{\mathrm{ref,l}}(n) &= y_{\mathrm{ref,l}}(n) - \hat{\boldsymbol{h}}_{\mathrm{ltrl,ref,l}}^{\mathrm{T}}(n)\boldsymbol{x}_{\mathrm{ltrl}}(n) - \hat{\boldsymbol{h}}_{\mathrm{frnt,ref,l}}^{\mathrm{T}}(n)\boldsymbol{x}_{\mathrm{frnt}}(n) \\ &= \left[\boldsymbol{h}_{\mathrm{ltrl,ref,l}}(n) - \hat{\boldsymbol{h}}_{\mathrm{ltrl,ref,l}}(n)\right]^{\mathrm{T}} \boldsymbol{x}_{\mathrm{ltrl}}(n) + \left[\boldsymbol{h}_{\mathrm{frnt,ref,l}}(n) - \hat{\boldsymbol{h}}_{\mathrm{frnt,ref,l}}(n)\right]^{\mathrm{T}} \boldsymbol{x}_{\mathrm{frnt}}(n) + v_{\mathrm{ref,l}}(n). \end{aligned} \tag{9}$$

To obtain a good tracking behavior of the time-variant impulse responses, we need to choose a suitable excitation signal. Perfect sequences (PSEQs) have proven to be a superior excitation signal for tracking

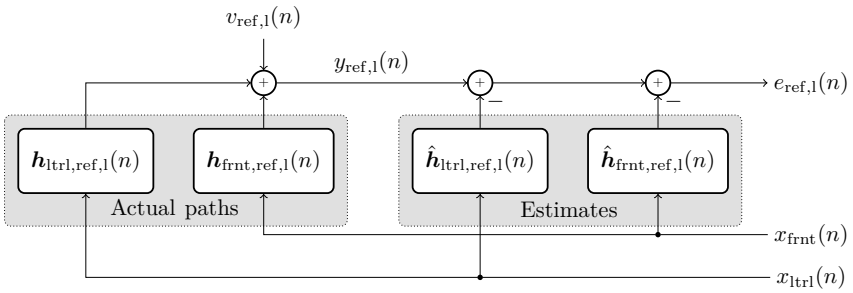

Figure 4. System identification block diagram at the example of the transfer functions to the left reference microphone.

time-variant impulse responses [1]. This is due to their perfectly flat magnitude spectrum. The auto-correlation function $\varphi_{xx}(\lambda)$ of a perfect sequence of length M is equal to one for $\lambda = iM$ with $i \in \mathbb{N}$ and zero otherwise. The tracking behavior of PSEQs has been analyzed extensively in [6]. Furthermore, a perfect sequence has been introduced as a noise signal, which has a uniformly distributed group delay and a flat magnitude spectrum. We have a maximum of three signals that arrive at the same microphone. To prevent the three excitation signals from interfering, we design a perfect noise signal of the length $3L$. The signal is circularly shifted by L so that the mutual shift for the frontal, lateral and internal loudspeaker signal is L. This way, the first L values of all cross-correlation functions of the excitation signals are zero.

3.3 Post-Processing

To compensate for remaining flaws and to structure the measured data during evaluation, we introduce some processing steps. The Neumann KH120 speakers have a frequency range starting from 52 Hz, and also the anechoic chamber is not reflection free for such low frequencies. Furthermore, the pre-amps might add a DC offset. Therefore, we apply a Butterworth high-pass filter with a passband frequency of 50 Hz to all microphone signals $y(n)$ as well as the respective excitation signals $x(n)$ used for the NLMS algorithm as in (9).

As we are solely interested in the acoustic transfer functions of the primary path $P(z)$ and the secondary path $S(z)$ without the influence of the electronic back-end, we need to compensate the transfer functions of the back-end. These transfer functions were measured beforehand, as indicated by the direct path in Fig. 2. They are used to deconvolute the microphone signals by means of spectral division. Now, to obtain $P(z)$ and $S(z)$, we use the time-variant estimates from the NLMS algorithm. The secondary path $S(z,n) = H_{\text{int,err}}(z,n)$ at time n is given by the transfer function from the left or right internal speaker to the respective error microphone. As the primary path is the transfer function from M_{ref} to M_{err}, we need to perform a spectral division. The primary path for, e.g., the lateral external speaker is given by

$$P(z,n) = \frac{H_{\text{ltrl,err}}(z,n)}{H_{\text{ltrl,ref}}(z,n)}. \tag{10}$$

If the denominator transfer function has a notch, this will lead to an unnatural amplification at certain frequencies in the primary path. When such primary paths are used for the design of a feed-forward filter, the amplifications result in an annoying resonating sound [2]. We compensate these notches by analyzing the resulting primary path $p(n)$, as Fig. 5 shows. In case of a notch, the primary path's impulse response does not decay. We analyze the latter part of the primary path using a zero-padded Hann window $q(n)$, and we use the phase of the denominator to generate the compensation spectrum

$$C(z) = \alpha \left|\mathcal{F}\left\{p(n) \cdot q(n)\right\}\right| e^{j\angle H_{\text{ltrl,ref}}(z,n)}, \tag{11}$$

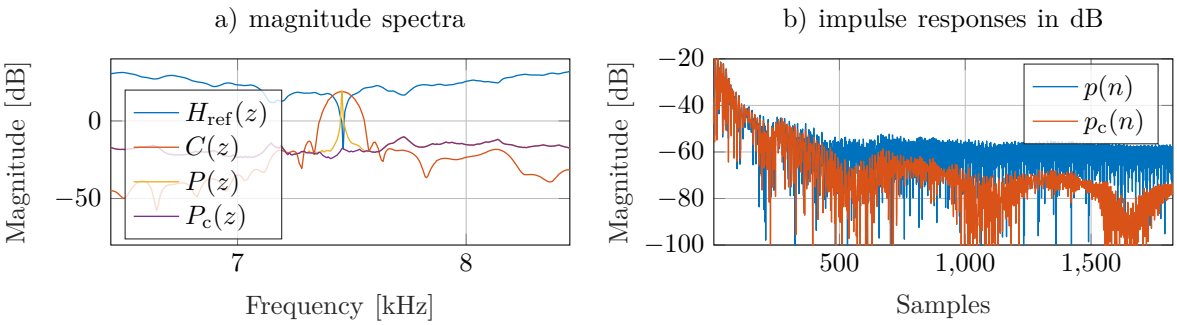

Figure 5. Notch compensation by analyzing $p(n)$ with a Hann window.

with the Fourier transform $\mathcal{F}\{\cdot\}$, the scaling factor $\alpha = \sqrt{M/\|\boldsymbol{q}\|^2}$ and the FFT length M. We then obtain the transfer function of the corrected primary path as

$$P_{\mathrm{c}}(z,n) = \frac{H_{\mathrm{ltrl,err}}(z,n)}{H_{\mathrm{ltrl,ref}}(z,n) + C(z)}. \tag{12}$$

In order to save only the transfer functions that convey new and reliable information, we define a set of rules which must be fulfilled before saving a set of impulse responses. The short-time power of the error signals needs to be below $-50\,\mathrm{dB}$. Furthermore, for the time-invariant segments from Fig. 3 we save only a single set of transfer functions where the short-time power of the error signal is smallest. For the time-variant and refitting segments we save a new set of transfer functions if the Euclidean distance to any respective transfer function that was saved last exceeds $-40\,\mathrm{dB}$.

4 MEASUREMENT RESULTS

In this section, we discuss the results of the time-variant measurement series. The following plots are based on over $10,000$ individual paths for 25 participants. All paths are calculated with a filter length $L = 2048$ at a sample rate of $f_{\mathrm{s}} = 48\,\mathrm{kHz}$.

4.1 Nominal Paths

Fig. 6 shows the nominal secondary paths in a) as the average over all segments with normal fitting for each subject and in b) as the average over all subjects for each use-case. It is evident that $S(z)$ strongly depends on the geometry of the subjects ear canal. We further see that nominal path changes notably for the segments *headphones out*, where the headphones are not inserted into the ears, and *loose fit*, where the headphones are inserted loosely into the ears. The drop in magnitude at frequencies below 1 kHz for the use-case *headphones out* is due to the lower acoustic impedance for which the internal headphone driver does not provide enough power. On the other hand, the nominal secondary path is independent of the yaw position.

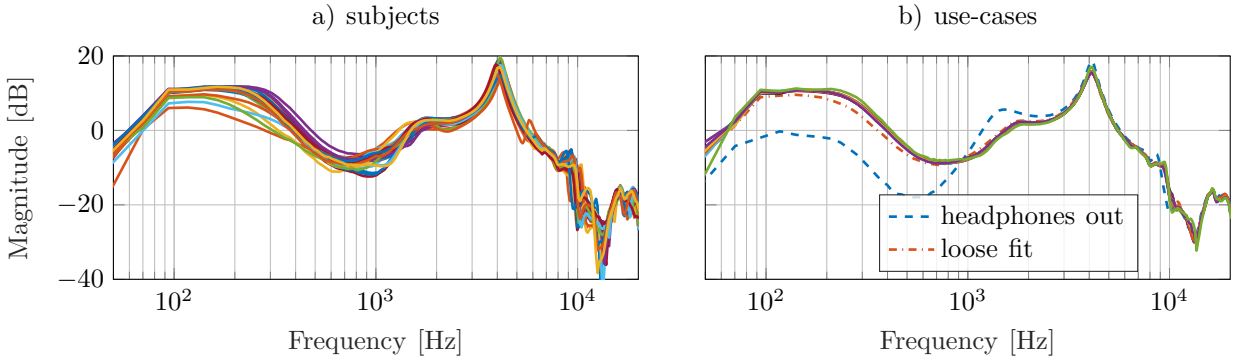

Figure 6. Nominal secondary paths $\overline{S}(z)$ in a) averaged over all segments with instructed normal fitting for different subjects and in b) averaged over all subjects for different segments.

Similar results can be observed when evaluating the primary path. Fig. 7 shows the averaged magnitude spectra of the primary paths of the right headphone and the lateral external loudspeaker, again averaged over use-cases and subjects. We see that the use-cases *headphones out* and *loose fit* result in a boost at frequencies between 400 Hz and 4 kHz. This is due to the reduced passive insulation. For $\varphi = 90°$ we experience additional attenuation of frequencies above 500 Hz and especially the peak at 4 kHz is reduced. We notice that different yaw positions do also not affect the primary path.

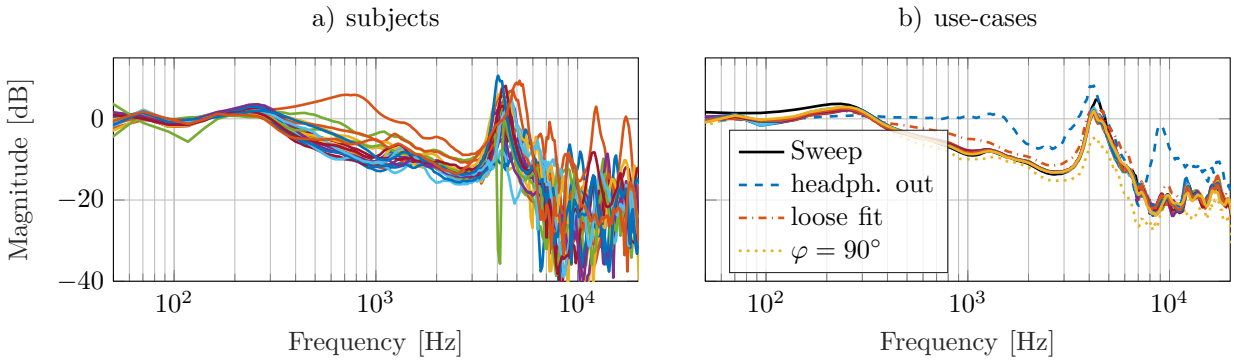

Figure 7. Averaged magnitude spectra of the primary paths $\overline{P}(z)$ for the right headphone and the lateral external loudspeaker in a) averaged over all use-cases with normal fitting for different subjects and in b) averaged over all subjects for different use-cases.

4.2 Implications on the Performance of Active Noise Cancellation Headphones

Rafaely derived an explicit relationship between the uncertainty of the plant in a feedback ANC system and the possible performance [11]. From Sec. 4.1 we know that the secondary path varies strongly for different users and not so much for different use-cases. Fig. 8 reinforces this statement. It shows the percentiles of the multiplicative uncertainty according to (6) in steps of 10% from 10% to 100% for a) the nominal path based on multiple subjects and b) an individual nominal path for each subject. We see that the maximum uncertainty for both cases is similar, nonetheless, the 90% percentile for the intra-subject uncertainty is already much smaller than for the inter-subject uncertainty. Hence, calibrating $S(z)$ once for a user may result in a very accurate representation of the actual path and can drastically improve the performance of both feed-back and feed-forward ANC [11, 4]. Also, calibrating the secondary path can simply be done by using the internal loudspeaker. The primary path on the other hand varies quite strongly, especially for different directions-of-arrival, which has been examined in more detail in [10]. Calibrating $P(z)$ for multiple discrete angles is feasible but requires a dedicated measurement setup. With a good primary path measurement the performance of feed-forward ANC systems can be drastically improved as shown in [4].

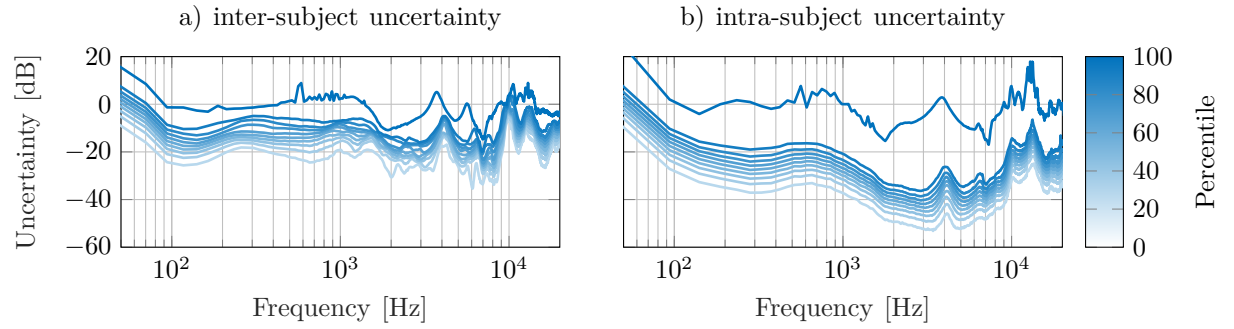

Figure 8. Multiplicative uncertainty of the secondary path in decibels for a) the nominal path based on all use-cases and subjects and b) individual nominal paths for each subject based on all use-cases.

5 CONCLUSION

In this contribution we present the results of a measurement series of time-variant device-related impulse responses of an active noise cancellation (ANC) headphone. The focus of this study is to see how the relevant acoustic transfer functions change for different subjects and use-cases. The measurements were conducted using perfect sequences (PSEQs) and the normalized least mean squares (NLMS) algorithm. We further propose a method for compensating notches when calculating impulse responses based on spectral division. The results based on over $10,000$ acoustic transfer functions show that an individualization of the primary and secondary path is a promising measure to improve the performance of time-invariant ANC systems. This is because the intra-person variance of the acoustic transfer functions is much lower compared to the inter-person variance. Furthermore, we investigate the multiplicative uncertainty, which plays an important role in designing robust feed-back systems, to underline this statement.

REFERENCES

[1] C. Antweiler. Multi-channel system identification with perfect sequences–Theory and applications–. In R. Martin, U. Heute, and C. Antweiler, editors, *Advances in Digital Speech Transmission*, chapter 7, pages 171–198. John Wiley & Sons, Ltd., Jan. 2008.

[2] S. Cecchi, A. Carini, and S. Spors. Room response equalization—a review. *Applied Sciences*, 8(1):16, 2018.

[3] S. Elliott. *Signal Processing for Active Control*. Elsevier, Amsterdam, Netherlands, 2000.

[4] J. Fabry, F. König, S. Liebich, and P. Jax. Acoustic equalization for headphones using a fixed feed-forward filter. IEEE, IEEE, May 2019.

[5] C. Hansen, S. Snyder, X. Qiu, L. Brooks, and D. Moreau. *Active Control of Noise and Vibration*. CRC Press, Boca Raton, 2012.

[6] S. Kühl, S. Nagel, T. Kabzinski, C. Antweiler, and P. Jax. Tracking of time-variant linear systems: Influence of group delay for different excitation signals. In *Proceedings of International Workshop on Acoustic Signal Enhancement (IWAENC)*, pages 131–135. IEEE, Sept. 2018.

[7] S. M. Kuo and D. R. Morgan. *Active Noise Control Systems: Algorithms and DSP Implementations*. Wiley, Hoboken, 1996.

[8] S. Liebich, C. Anemüller, P. Vary, P. Jax, D. Rüschen, and S. Leonhardt. Active noise cancellation in headphones by digital robust feedback control. In *2016 24th European Signal Processing Conference (EUSIPCO)*, pages 1843–1847. IEEE, 2016.

[9] S. Liebich, P. Jax, and P. Vary. Active cancellation of the occlusion effect in hearing aids by time invariant robust feedback. In *Speech Communication; 12. ITG Symposium*, pages 1–5. VDE, 2016.

[10] S. Liebich, J.-G. Richter, J. Fabry, C. Durand, J. Fels, and P. Jax. Direction-of-arrival dependency of active noise cancellation headphones. In *ASME 2018 Noise Control and Acoustics Division Session presented at INTERNOISE 2018*, pages V001T08A003–V001T08A003. American Society of Mechanical Engineers, 2018.

[11] B. Rafaely. *Feedback control of sound*. PhD thesis, Unicersity of Southhampton (United Kingdom), 1997.

[12] World Health Organization (WHO). Burden of disease from environmental noise: Quantification of healthy life years lost in Europe. 2011.