MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Improving the Performance of an Active Noise Cancelling Headphones Prototype

Piero Rivera Benois[1], Patrick Nowak, Etienne Gerat, Muhammad Salman and Udo Zölzer[2]
Department of Signal Processing and Communications
Helmut Schmidt University / University of the Federal Armed Forces Hamburg, Germany

ABSTRACT

Headphones can be equipped with active noise cancelling for providing their users with the attenuation of the acoustical environmental noise that surrounds them. This technology is based on feedforward or feedback control schemes adapted for the control of sound. While feedback schemes can be implemented as fixed systems, the feedforward ones have to be implemented as adaptive systems. The latter because of their performance dependency on the frequency content and direction-of-arrival of the noise signal. In this paper, two aspects concerning a feedforward LMS-based control scheme are studied: First, what changes can be done to the time-domain FxLMS algorithm in order to improve its stability, convergence speed, and steady-state attenuation performance; and second, how the position of the inner microphone can be optimized in order to help the adaptation scheme to achieve these goals. Based on the outcomes of this study, an active noise cancelling headphones prototype is optimized and evaluated based on measurements done by means of two different dummy-heads.

Keywords: Active Noise Cancelling, Adaptive Feedforward Control, Headphones
I-INCE Classification of Subject Number: 36

1. INTRODUCTION

Active Noise Cancelling (ANC), also known as Active Noise Control, aims to reduce the sound pollution present in the environment by actively generating sound pressure waves that overlap destructively on a sweet-spot with the ones of the noise sources. Its origins can be dated back to Lueg's patent [1] on the control of acoustical tonal noise inside of a duct, while its particular application to headphones can be dated to the later work of Simshauser *et al.* in [2]. In order to achieve a perfect attenuation, the phase

[1]piero.rivera.benois@hsu-hh.de
[2]udo.zoelzer@hsu-hh.de

and the magnitude of the noise around the sweet-spot have to be matched. Different approaches can be used to achieve this, depending on the nature of the noise and the solution's context.

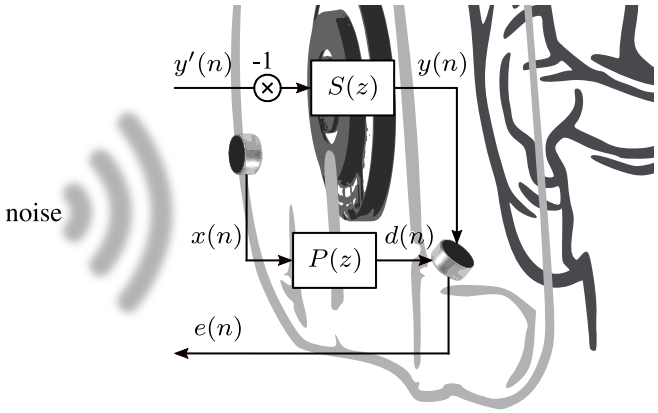


Figure 1: Signals and systems around ANC headphones

In the case of ANC headphones, the transducers and systems presented in Fig. 1 can be found. An external microphone measures a time-advanced reference of the incoming disturbance $x(n)$. This signal is measured again as $d(n)$, after it has entered the ear-cup and reached the so-called error microphone. Thus, $P(z)$, known as the primary path, is defined as the changes in magnitude and phase that the disturbance suffers by means of the headphone's construction materials and its angle of incidence. The control signal $y'(n)$ generated by the ANC system is phase-inverted, fed to the speaker and measured again by the error microphone as $y(n)$. Hence, the loudspeaker and the error microphone define the so-called secondary path $S(z)$, which considers the characteristics of both elements plus the acoustic path between them. At the end, the control signal $y(n)$ and the disturbance $d(n)$ overlap destructively at the error microphone's position, and the residual error $e(n)$ is generated. ANC approaches that use the residual error $e(n)$ for calculating $y'(n)$ are called feedback control schemes. On the other hand, the ones that use the time-advanced reference signal $x(n)$ for calculating $y'(n)$ are called feedforward or forward control schemes. Adaptive implementations of feedforward controllers also utilize $e(n)$ to solve the underlying optimization problem.

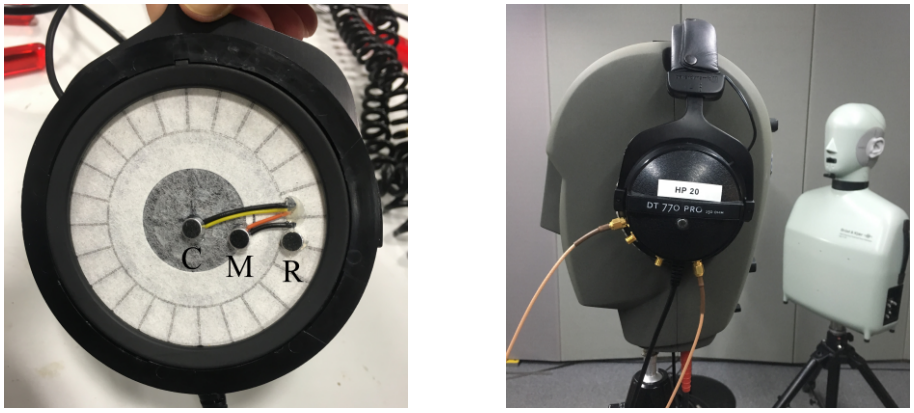
Between the adaptive feedforward control algorithms, the time-domain Filtered-x LMS (FxLMS) [3, 4] is preferred for real-time applications because of its simplicity. Nevertheless, its optimal parametrization has to be done under controlled laboratory conditions. In order to partially overcome this problem, two variations of the algorithm are studied in this work. The first one is the Normalized FxLMS (NFxLMS) with its time-dependent step-size by energy normalization, that provides a stable steady-state performance and a fast convergence speed under amplitude changing signals. The second one is the use of the Modified FxLMS (MFxLMS) [5, 6] together with the NFxLMS' time-dependent step-size in an MNFxLMS combination. This provides stability during the search for an optimal μ_0. Additionally, three different error microphone positions are evaluated, in order to improve the performance of the system.

In the following section, the three secondary paths related to the different microphone positions are measured and compared under the presence of two different dummy-heads. In Section 3 the variations of the FxLMS are presented and their improvements are exemplified by computer simulations. In Section 4, an ANC headphones prototype is

equipped with the MNFxLMS control algorithm, and the performance achieved with the three different microphone positions is compared based on measurements done by means of two dummy-heads. At the end, conclusions are drawn based on the observations done on the measurement results.

2. SYSTEM IDENTIFICATION

In this section, variations in the secondary path produced by different microphone positions are presented, based on impulse response measurements done on an ANC headphones prototype and two dummy-heads presented in Fig. 2. Additionally, the magnitude response of the transfer function between the headphones' loudspeaker and the microphone of the dummy-heads is measured, in order to establish how the magnitude of the control signal changes when reaching the user's ears.



(a) (b)

Figure 2: Measurement setup with (a) headphone prototype with error microphones at three different positions and (b) the ANC headphones prototype with the two dummy-heads.

In Fig. 2(a) the three microphone positions inside of the left earcup of the ANC headphones prototype are presented. The one in the center is labeled throughout this work with the letter C, the outer most with an R, and the one in the middle with an M. The device is also equipped with an external microphone, as shown in Fig. 2(b), for providing the control algorithm with the time-advanced reference signal. In order to include the influence of a human ear closing the air volume inside the earcups, two dummy-heads are used for the measurements as presented in Fig. 2(b). The first one is a Neumann KU-100, which has its transducer positioned at the entrance of the ear canal, and the second one is a Brüel & Kjær HATS type 5128, with an anatomically-correct ear canal at the end of which its transducer is located.

The systems are measured using the method described in [7], with a sinesweep signal of length 10 s at a sampling frequency $f_s = 48\,\text{kHz}$, start frequency $f_{start} = 5\,\text{Hz}$, and end frequency $f_{end} = 23.9\,\text{kHz}$. Additionally, the signal is faded in and out during 9600 samples, in order to avoid unwanted artifacts in the measurement results. The measured impulse responses are presented in Fig. 3, where Fig. 3(a) and Fig. 3(b) show the results achieved with the Neumann and Brüel & Kjær (BK) dummy-heads, respectively. By looking at the first samples of all six impulse responses, it can be seen that $s_{\text{R}}(n)$ has in both measurements one sample delay more than $s_{\text{C}}(n)$ and $s_{\text{M}}(n)$. This is basically because

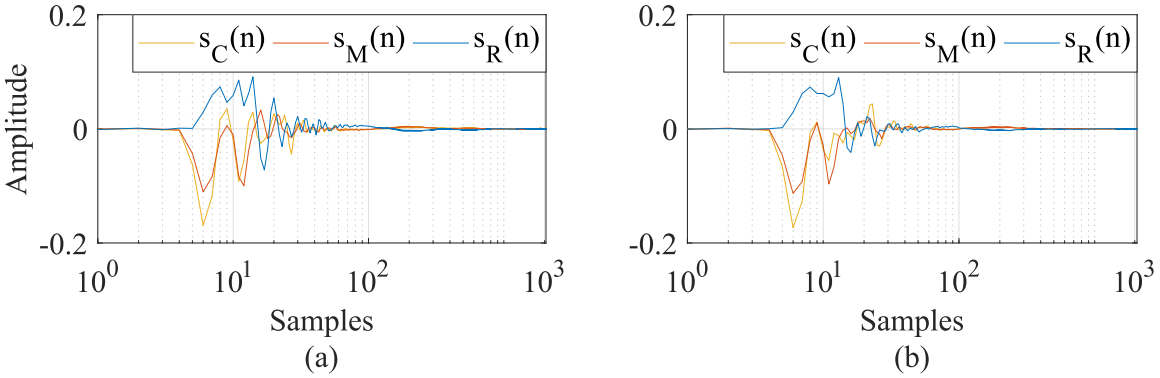


Figure 3: Impulse responses of secondary paths for (a) Neumann dummy-head and (b) BK dummy-head

it is not located directly in front of the loudspeaker. In the case of $s_{\mathrm{C}}(n)$ and $s_{\mathrm{M}}(n)$, it does not make a big difference to be placed in the middle or slightly shifted to the side of the cone. Looking at the starting amplitude of the impulse responses, it can be seen that the speaker produces a negative pressure around the microphones placed in front of it, as they are facing in the same direction. This is not the case of $s_{\mathrm{R}}(n)$, because its is not directly in front of the speaker and with its orientation it mainly measures the reflections coming from the ear and the air flow escaping through the small earcup's outlet at the side, behind the speaker.

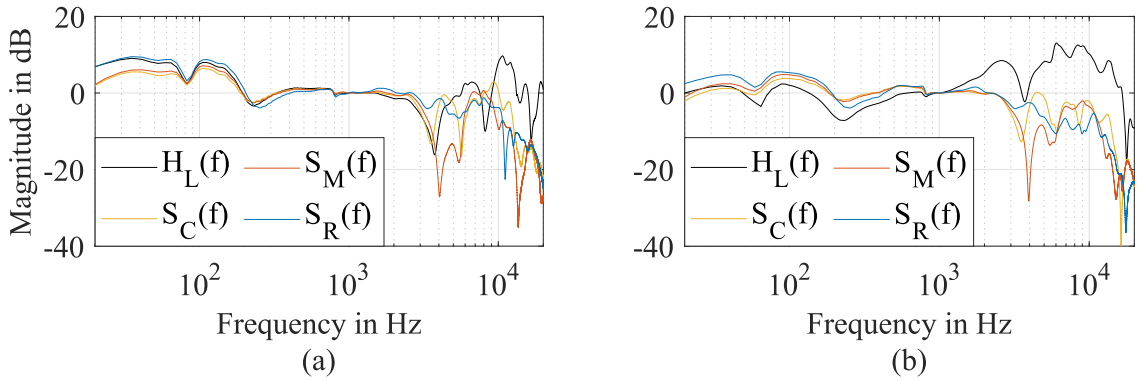


Figure 4: Frequency response of the three secondary paths (a) Neumann dummy-head and (b) BK dummy-head

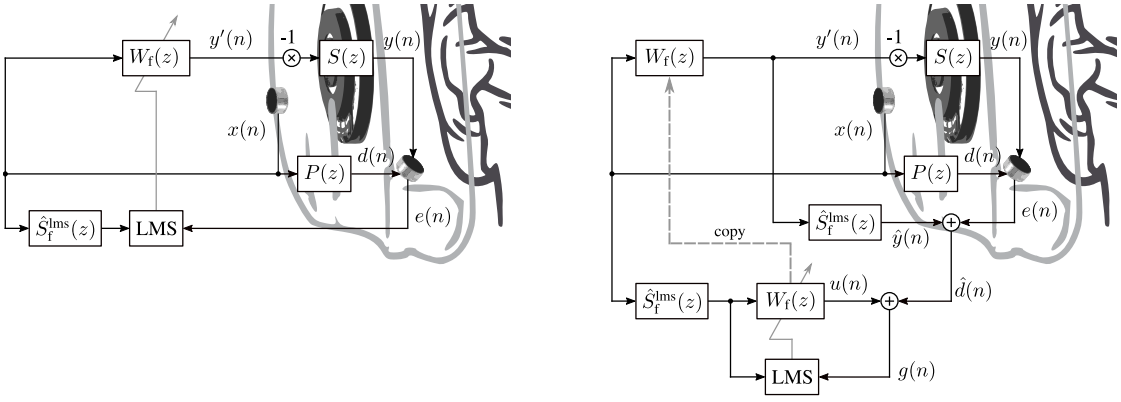
The frequency responses of the secondary paths and of the transfer function $H_{\mathrm{L}}(z)$ between the headphones' drive and the dummy-heads' left microphone are presented in Fig. 4. Fig. 4(a) presents the results achieved with the Neumann dummy-head, while Fig. 4(b) presents the ones achieved with the BK dummy-head. Because of the differences in sensitivity between all microphones, the curves have been normalized to match 0 dB at 1 kHz. By looking at the frequency responses presented in Fig. 4(a), one can see that below approximately 3 kHz all frequency responses show a very close behavior. Whereas beyond this frequency substantial differences can be seen, as the wavelengths start to be comparable with the dimensions of the earcup and the anthropometry of the ear. Interesting is to see that in the entrance of the ear canal an amplification around 10 kHz of 10 dB is produced, most probable because of the effect of the cavum concha. In Fig. 4(b) the secondary paths show a similar behavior to the ones already seen, with some deviations because of the differences between the dummy-heads' ears. However,

the frequency response H_{L} measured with the BK shows that the interaction between the ear canal and the closed earcup produces a significant amplification starting from 1 kHz.

All in all, $s_{\mathrm{C}}(n)$ and $s_{\mathrm{M}}(n)$ have one sample delay less than the $s_{\mathrm{R}}(n)$, which may be advantageous because of the underlying causality considerations that feedforward control requires. Because of the similarities between their frequency responses, it is expected that their performances below 3 kHz are very similar. However, the performance achieved at the error microphone position may differ from the one measured with help of a dummy-head. Moreover, stronger deviations are expected when measuring at the eardrum's position, if the noise coming from the outside of the earcup is not subject to the same amplification effect.

3. CONTROL ALGORITHM

In this section two improvements to the time-domain FxLMS are reviewed, in order to increase its stability and convergence speed. The first one is the time-varying step-size by means of energy normalization, and the second one is the use of the Modified FxLMS. The effects of this two changes are exemplified and discussed with simulations based on the impulse responses provided in [8].



(a) (b)

Figure 5: System diagram of (a) the Filtered-x LMS algorithm and (b) the Modified Filtered-x LMS algorithm.

The system diagram of an ANC headphone using the FxLMS algorithm is presented in Fig. 5(a). The time-advanced reference signal $x(n)$ is measured by the external microphone. The same signal reaches the position of the inner microphone by the primary path $P(z)$ as $d(n)$. The reference signal $x(n)$ is filtered by the controller $W_{\mathrm{f}}(z)$ to calculate the control signal $y'(n)$. This is then sign-inverted and fed to the secondary path $S(z)$, in order to destructively overlap with the disturbance signal $d(n)$ at the inner microphone's position. If the control algorithm is effective, then the residual error $e(n) = d(n) - y(n)$ is smaller than $d(n)$.

The FxLMS adaptation algorithm uses the reference signal $x(n)$ in a series connection with an estimated secondary path $\hat{S}_{\mathrm{f}}^{\mathrm{lms}}(z)$, in order to generate its filtered reference $x_{\mathrm{f}}(n)$. The coefficients are sample-by-sample updated following the equation

$$\mathbf{w}_{\mathrm{f}} = \mathbf{w}_{\mathrm{f}} + \mu \cdot \mathbf{x}_{\mathrm{f}} \cdot e(n), \tag{1}$$

where $\mathbf{w}_\mathrm{f}$ is the controller's coefficients vector of length L_w, $\mathbf{x}_\mathrm{f}$ is the buffer containing the last L_w samples of $x_\mathrm{f}(n)$, $e(n)$ is the current sample of the residual error, and μ is the adaptation step-size. In order to yield a stable adaptation, μ should be chosen according to

$$0 < \mu < \frac{2}{(L_\mathrm{w} + \Delta)\mathrm{E}[x_\mathrm{f}^2]}, \tag{2}$$

where Δ is the delay introduced by the secondary path and $\mathrm{E}[x_\mathrm{f}^2]$ the variance of the random signal $x_\mathrm{f}(n)$ (this can be replaced by the signal power P_{x_f}, if $x(n)$ is a deterministic signal). The maximum μ value in Eq. (2) is only a theoretical one, and if a fast adaptation is desired, an iterative search has to be performed with an $x(n)$ of stable $\mathrm{E}[x_\mathrm{f}^2]$. Alternatively, the search can be simplified, if a sample-by-sample estimation of $\mathrm{E}[x_\mathrm{f}^2]$ is included in the adaptation algorithm. This can be done based on the observations of x_f stored in the buffer $\mathbf{x}_\mathrm{f}$, as $E[x_\mathrm{f}^2] \approx (\mathbf{x}_\mathrm{f}^\mathrm{T} \cdot \mathbf{x}_\mathrm{f})/L_\mathrm{w}$. This yields the so-called Normalized FxLMS (NFxLMS), which coefficients update equation

$$\mathbf{w}_\mathrm{f} = \mathbf{w}_\mathrm{f} + \mu(n) \cdot \mathbf{x}_\mathrm{f} \cdot e(n) \tag{3}$$

contains a time-varying step-size $\mu(n)$ given by

$$\mu(n) = \frac{\mu_0}{\mathbf{x}_\mathrm{f}^\mathrm{T} \cdot \mathbf{x}_\mathrm{f}}, \tag{4}$$

where μ_0 is chosen as

$$0 < \mu_0 < \frac{2}{(1 + \Delta/L_\mathrm{w})}. \tag{5}$$

Thus, the search for the highest μ, now μ_0, can be be carried out in scenarios with changing $\mathrm{E}[x_\mathrm{f}]$ and the result will depend only on the delay of the secondary path Δ and the length of the controller L_w. In order to exemplify how the stability and convergence speed of two systems may contrast, if one uses an FxLMS instead of an NFxLMS, the system presented in Fig. 5(a) is simulated using the impulse responses provided in [8]. The systems $W_\mathrm{f}(z)$, $P(z)$, $S(z)$, and $\hat{S}_\mathrm{f}^{\mathrm{lms}}(z)$ are simulated as FIR filters with 128 taps, 256 taps, 128 taps, and 128 taps, respectively. The performance of both systems is compared based on the Average Noise Reduction

$$\mathrm{ANR}(n) = 20 \log_{10}\left(\frac{A_e(n)}{A_d(n)}\right), \tag{6}$$

where $A_e(n)$ and $A_d(n)$ are sample-by-sample first-order estimations of the residual error's and disturbance's magnitudes following $A_e(n) = \lambda \cdot A_e(n-1) + (1-\lambda) \cdot |e(n)|$ and $A_d(n) = \lambda \cdot A_d(n-1) + (1-\lambda) \cdot |d(n)|$, with $\lambda = 0.994$.

For the first and second scenario, the one second long uniformly distributed random reference signals $x(n)$ with changing amplitudes as presented in Fig. 6(a) and Fig. 6(b) are respectively used. In both cases an iterative search with the FxLMS for the fastest μ that yields a stable adaptation is performed. For the search a signal that matches the initial characteristics of the respective $x(n)$ is utilized. Consecutively, a μ_0 that yields a comparable performance is chosen for the NFxLMS. In Fig. 6(c), the performance of the FxLMS and NFxLMS can be compared in cases in which the $\mathrm{E}[x_\mathrm{f}^2]$ increases beyond the one used for estimating μ. It can be observed that the FxLMS performance degrades significantly when $\mathrm{E}[x_\mathrm{f}^2]$ is higher than the one used for finding μ. In contrast to this behavior, the performance of the NFxLMS remains unaltered during these periods of

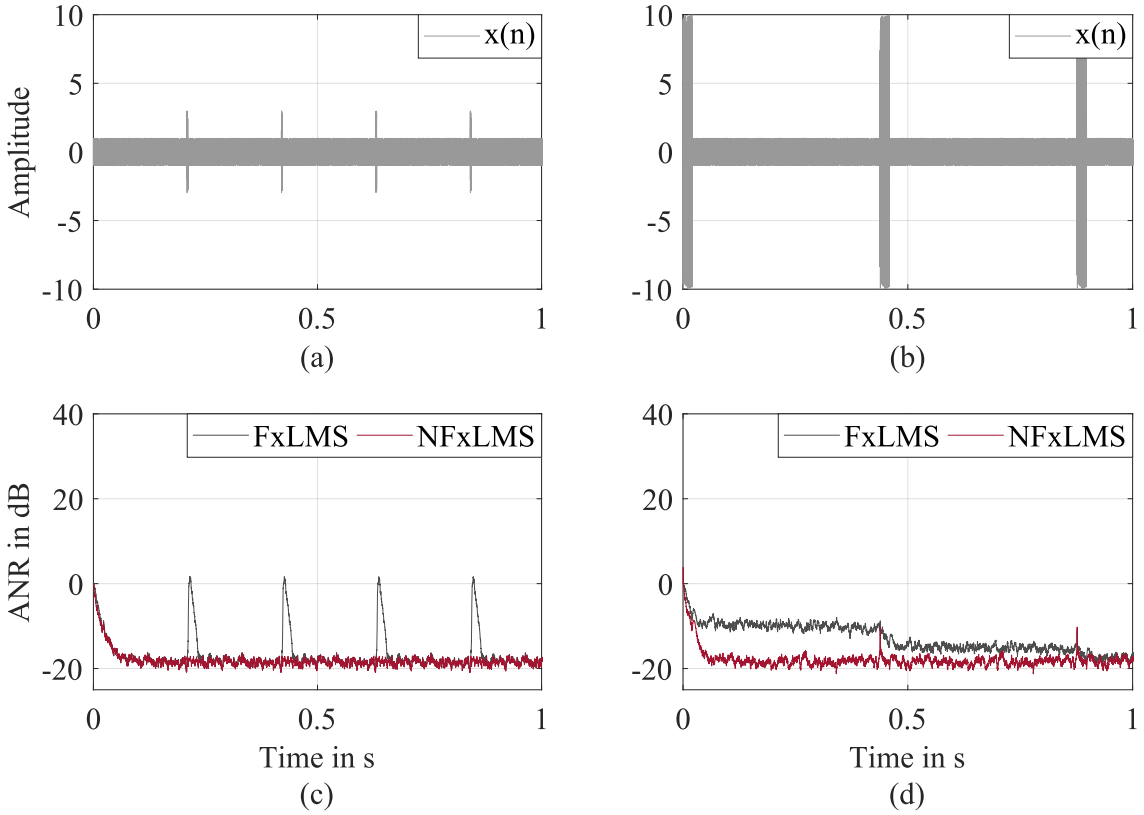


Figure 6: Convergence speed and stability of the Filtered-x LMS and of the Normalized Filtered-x LMS for uniformly distributed random reference signals with changing amplitudes.

time. In Fig. 6(d), the performance of both algorithms is compared in cases in which the $\mathrm{E}[x_\mathrm{f}^2]$ decreases below the one used for estimating μ. It can be observed, that during the beginning both algorithms converge similarly fast, as the characteristics of the signals match the one used for estimating μ. Nevertheless, as soon as the amplitude of the signal drops, the convergence speed of the FxLMS decreases while the one of the NFxLMS remains unaltered.

As it can be seen from Eq. (5), μ_0 is independent from the variance or power of the filtered reference $x_\mathrm{f}(n)$. This means that a search for its optimum value can be performed under more realistic circumstances. However, during this search the system may run into instability, if $\hat{S}_\mathrm{f}^{\mathrm{lms}}(z)$ deviates from $S(z)$ or upon small increments in μ_0. Thus, the search should not be carried out in the presence of a human subject. An example of this search is presented in Fig. 7(a). The starting μ_0 has been chosen to be 0.1. As the μ_0 is incremented by 300%, the convergence speed gets higher without loosing in steady-state ANR. A relatively small increment of approx. 10% shows some sporadic unstable ANR performance through time. If μ is incremented again by approx. 10%, then the system becomes unstable. In practical implementations, such a scenario produces clipping of the digital signals and generates harmonic distortion throughout the signal processing chain, such that the only solution to recuperate the control of the system is to reset the controller coefficients.

A variation of the NFxLMS called the Modified NFxLMS (MNFxLMS) simplifies the search for an optimum μ_0. Its system diagram is presented in Fig. 5(b). As it can be seen, an estimated secondary path $\hat{S}_\mathrm{f}^{\mathrm{lms}}(z)$ is used to estimate the disturbance signal $d(n)$ from

the measured error signal $e(n)$. The estimated $\hat{d}(n)$ is used in combination with the filtered $x(n)$ to isolate a copy of the controller and adapt it using

$$\mathbf{w}_\mathrm{f} = \mathbf{w}_\mathrm{f} + \mu(n) \cdot \mathbf{x}_\mathrm{f} \cdot \left(\mathbf{w}_\mathrm{f}^\mathrm{T} \cdot \mathbf{x}_\mathrm{f} - \hat{d}(n)\right), \tag{7}$$

i.e. the LMS algorithm. By doing this, the controller coefficients are corrected using a local error estimation, independently from the delays and current state of the secondary path. This changes the theoretical boundaries for stable μ_0 to

$$0 < \mu_0 < 2. \tag{8}$$

A simulation of the search for μ_0 using the MNFxLMS is presented in Fig. 7(b). It can be seen for $\mu_0 = 0.1$ and $\mu_0 = 0.4$, that the convergence speed and steady-state ANR are the same as for the NFxLMS. Nevertheless, if a higher μ_0 is chosen, the system looses in steady-state ANR, instead of becoming unstable. Thus, μ_0 can now be optimized by monitoring the steady-state ANR and the convergence curve, without risking the stability of the system. However, it has to be mentioned that, just like the NFxLMS, the MNFxLMS may still become unstable if the phase of $\hat{S}_\mathrm{f}^\mathrm{lms}(z)$ deviates by more than 90° from the one of the $S(z)$.

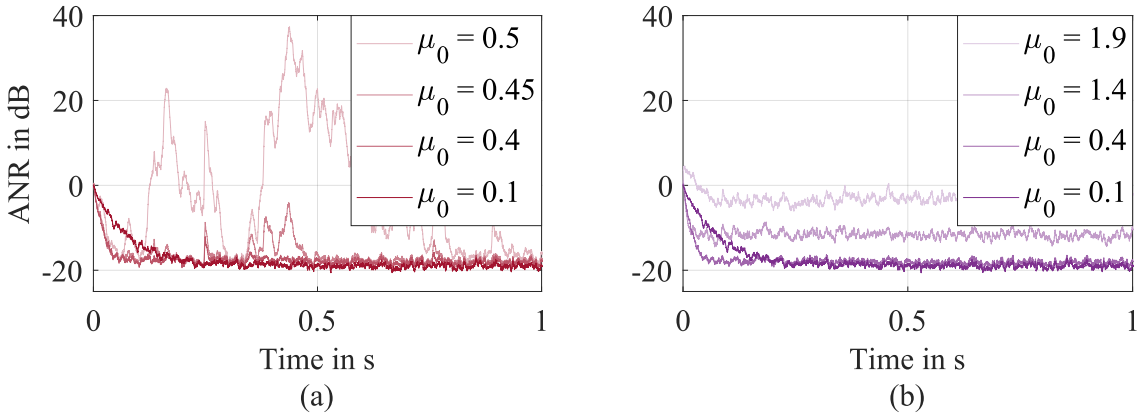


Figure 7: Example of a search for the optimal μ_0 when using (a) the Normalized Filtered-x LMS and (b) the Modified Normalized Filtered-x LMS.

4. MEASUREMENT RESULTS

In this section, the three error microphone positions presented in Fig. 2(a) are evaluated in terms of active noise control performance. For that, the prototype presented in Fig. 2(b) is equipped with the MNFxLMS adaptation algorithm and two measurement sets are performed, i.e. one with the Neumann dummy-head and another one with the BK dummy-head. A measurement setup overview is presented in Fig. 8. A Genelec loudspeaker is placed 90 cm at the left side of the dummy-head wearing the prototype, connected to an RME Fireface UCX, through which the uniformly distributed white noise excitation signal is generated. The prototype's microphones are connected through a pre-amplifier to a MicroLabBox from dSpace, where the MNFxLMS is implemented on its FPGA. The MicroLabBox's output is connected to a PreSonus HP4 headphones amplifier, which drives the headphones' left speaker. The dummy-head is connected through a LAN-XI from Brüel & Kjær to the computer for recording purposes. A MNFxLMS with a

controller length of $L_w = 512$ taps and an estimated secondary path of length 2048 taps are used. A first run is performed in order to find a μ_0 that yields the best steady-state result. The resulting $\mu_0 = 5 \cdot 10^{-4}$ is used in all performed measurements. A measurement of the magnitude over frequency of the reference signal $X(f)$, disturbance signal $D(f)$, and dummy-head signal $L_{OFF}(f)$ is done for later comparison, while the ANC system is turned off. Afterwards, the system is connected to its corresponding microphone, it adapts for 10 s, after which the adaptation is stopped, and a 20 s recording is done by means of the dummy-head. The magnitude over frequency at the error microphone $E_{C/M/R}(f)$ is calculated in the MicroLabBox and exported through Control Desk. Additionally, the magnitude over frequency at the dummy-head $L_{C/M/R}(f)$ is calculated from the 20 s recordings.

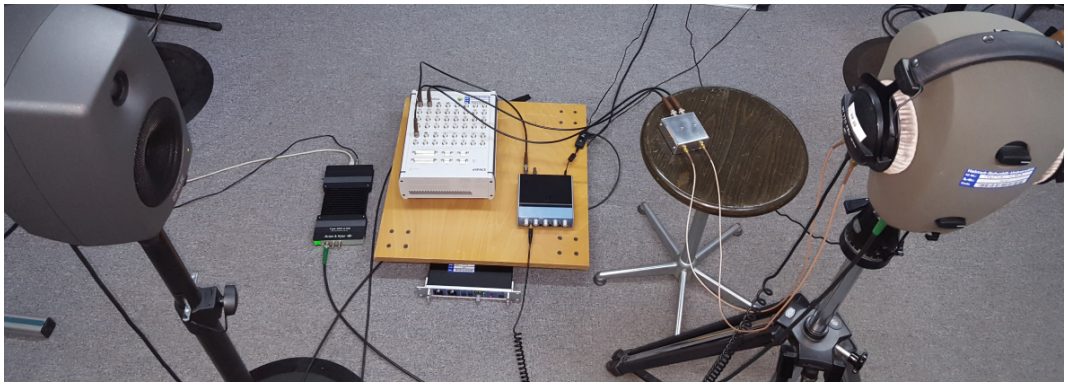


Figure 8: Measurement setup overview

The results measured with the Neumann dummy-head are presented in Fig. 9. The left column presents the signals measured at the control system, while the right column presents the ones measured at the dummy-head. The first row presents the case with the microphone at the C position, the second row the one at the M position, and the last one the case with the microphone at the R position. It can be seen from the left column, that the results at the three different error microphone positions are very similar for frequencies below 3 kHz, as expected from the similarities found in the secondary paths' frequency responses. Beyond this frequency, the proximity of the microphone to the center of the earcup improves the attenuation performance and the attenuation bandwidth that the system can achieve. This is probably produced by the distance to the speaker's center and the acoustic delay related to it. Comparing the attenuation level achieved at the error microphone with the respective one at the Neumann ear canal entrance, it can be seen that for frequencies below 3 kHz the residual error frequency content suffers a small coloration, but an attenuation can still be achieved. However, for frequencies beyond 4 kHz the attenuation performance degrades significantly and causes small amplifications in $L_M(f)$ and $L_R(f)$ at approximately 10 kHz. This is produced by the differences in this frequency region between the secondary paths and the transfer function until the entrance of the ear canal in Fig. 4(a), which is combined with the different disturbance spectrum at the microphone $D(f)$ and at the dummy-head's ear $L_{OFF}(f)$.

In Fig. 10 the results obtained with the BK dummy-head are presented. Here, similar results are obtained at the position of the error microphones, i.e. the proximity to the center of the earcup increases the attenuation performance and its bandwidth beyond the 10 kHz. However, the achieved attenuation degrades significantly, when it is measured at the position of the eardrum. This produces an amplification for frequencies higher

than 4 kHz. This is to be expected, since the control signal suffers a strong amplification when going through the transfer function $H_{\mathrm{L}}(z)$ in Fig. 4(b). This amplification has not been seen at the disturbance, when comparing $L_{\mathrm{OFF}}(f)$ with $D(f)$ in Fig. 10. Thus, the amplitude of the control signal is higher than the one of the disturbance, and the control signal arrives at the eardrum with a high gain.

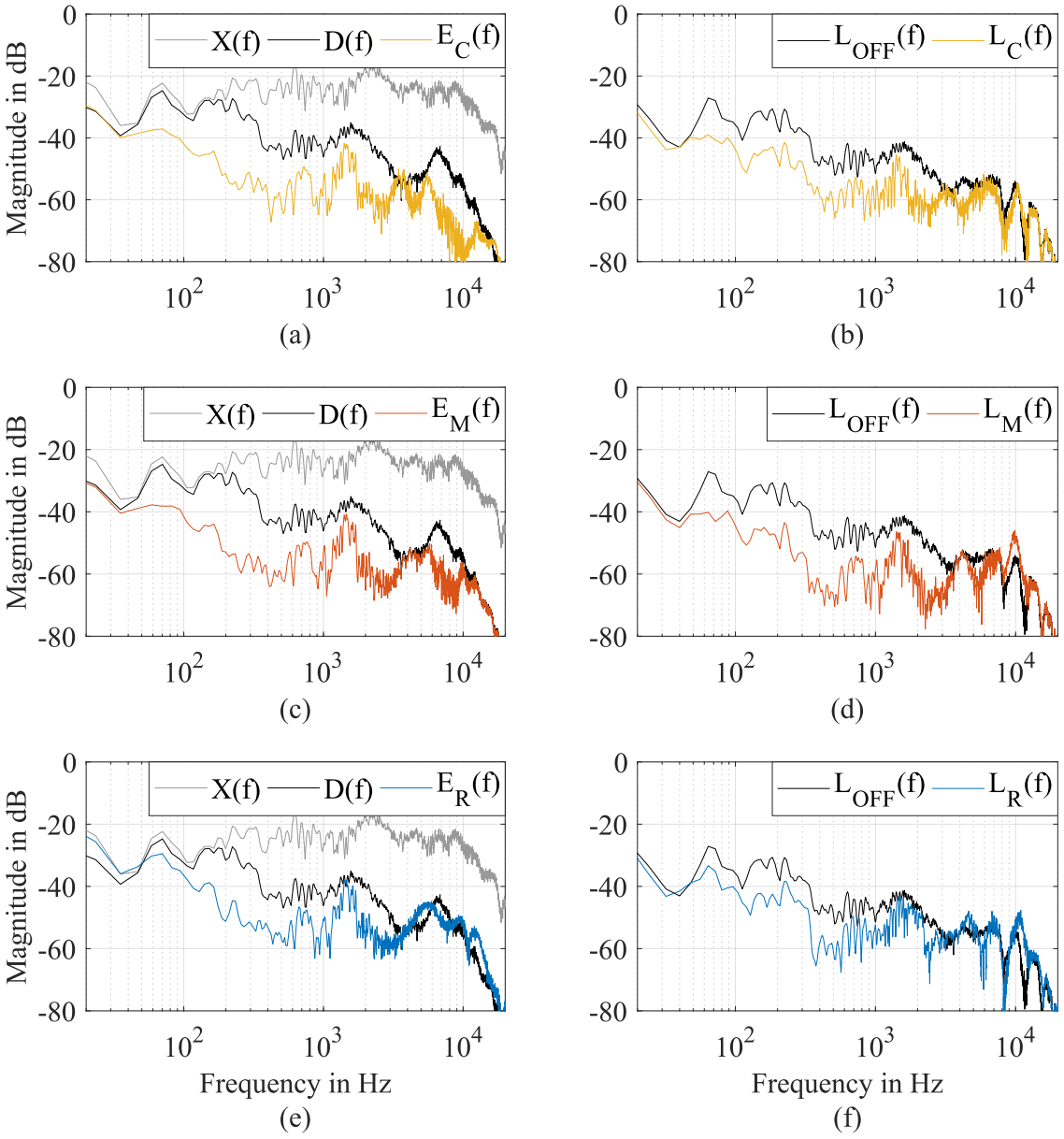


Figure 9: Attenuation results in presence of the Neumann dummy-head: The left column presents the signals measured at the headphones prototype microphones and the right column presents the signals measured at the dummy-head microphone.

5. CONCLUSIONS

In this paper three error microphone positions have been evaluated to improve the attenuation performance of a Modified Normalized FxLMS implemented on an ANC prototype. The attenuation performance measured at the error microphones has been contrasted to the ones measured at the ear canal's entrance by means of a Neumann

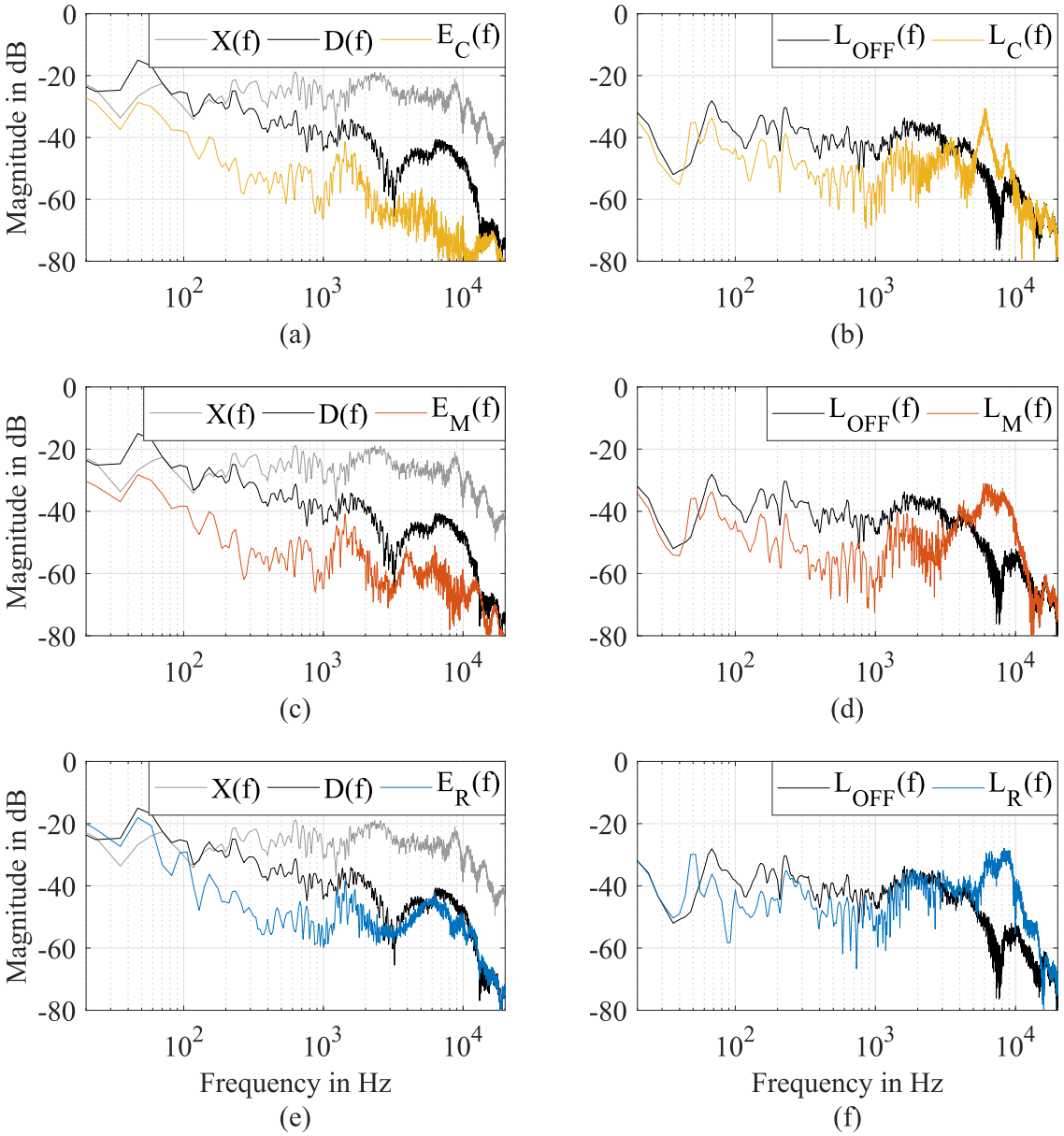


Figure 10: Attenuation results in presence of the BK dummy-head: The left column presents the signals measured at the headphones prototype microphones and the right column presents the signals measured at the dummy-head microphone.

KU-100 dummy-head, and at the eardrum by a Brüel & Kjær HATS type 5128. The measurements showed that the nearer the microphone to the center of the speaker is, the better are the results. However, the results obtained at frequencies higher than 3 kHz degrade, when measured at the ear canal's entrance. Moreover, with measurements done at the eardrum, it could be seen that the control signal and the disturbance are subject to different coloration and, therefore, do not overlap destructively anymore at frequencies higher than 4 kHz. All in all, if over-ear headphones are going to be used, it would be better to make use of a sub-band ANC approach, in order to avoid the undesired effects for frequencies higher than 4 kHz. Alternatively, an in-ear headphone construction may solve the problem, as its error microphone would be closer to the eardrum than the one of the present prototype.

6. REFERENCES

[1] Paul Lueg. Process of silencing sound oscillations, June 3 1936. US Patent 2,043,416.

[2] Elvin D. Simshauser and Mones E. Hawley. The noise cancelling headset—an active ear defender. *The Journal of the Acoustical Society of America*, 27(1):207–207, 1955.

[3] J. C. Burgess. Active adaptive sound control in a duct - A computer simulation. *Acoustical Society of America Journal*, 70:715–726, September 1981.

[4] B. Widrow, D. Shur, and S. Shaffer. On adaptive inverse control. In *15th Asilomar Conference on Circuits, Systems and Computers*, pages 185–189, Nov 1981.

[5] Chaoeng Bao, Paul Sas, and Hendrik Van Brussel. A novel filtered-x LMS algorithm and its application to active noise control of sound. In *1992 European Signal Processing Conference (EUSIPCO)*, pages 1709–1712, Dec 1992.

[6] Bjarnason E. Active noise cancellation using a modified form of the filtered-x LMS algorithm. In *1992 European Signal Processing Conference (EUSIPCO)*, pages 1053–1056, Dec 1992.

[7] Piero Rivera Benois, Purbaditya Bhattacharya, and Udo Zölzer. Derivation Technique for Headphone Transfer Functions Based on Sine Sweeps and Least Squares Minimization. *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 253(4), 2016.

[8] Kong-Aik Lee, Woon-Seng Gan, and Sen M. Kuo. *Subband Adaptive Filtering: Theory and Implementation*. Wiley Publishing, 2009.