New Approaches for Active Noise Control Headphones

Markus Guldenschuh

New Approaches for Active Noise Control Headphones

PhD Thesis

at the

University of Music and Performing Arts Graz

submitted by

Markus Guldenschuh

Institute of Electronic Music and Acoustics,
University of Music and Performing Arts Graz
8010 Graz, Austria

June 2014

Assessors:

Prof. Robert Höldrich (IEM/KUG, Graz, Austria)
Prof. Walter Kellermann (LNT/FAU, Erlangen, Germany)

<u>Markus Guldenschuh</u>   <u>0230947</u>
(Name in Blockbuchstaben)   (Matrikelnummer)

Erklärung

Hiermit bestätige ich, dass mir der *Leitfaden für schriftliche Arbeiten an der KUG*

bekannt ist und ich diese Richtlinien eingehalten habe.

Graz, den ……………………………………….

……………………………………………………..
Unterschrift der Verfasserin / des Verfassers

Kurzfassung

Kopfhörer mit aktiver Geräuschunterdrückung reduzieren den Umgebungslärm, indem sie diesen Umgebungslärm aufnehmen und gegenphasig wieder abspielen. Dadurch löschen sich der Umgebungslärm und der abgespielte 'Antilärm' im Kopfhörer (und damit direkt am Ohreingang der Benutzerin) aus. Diese Dissertation beschreibt sowohl für die vorwärts- als auch für die rückwärtsgekoppelte Geräuschunterdrückung neue Ansätze und geht dabei auch besonders auf Variabilitäten in der Kopfhörerakusik ein, die sich aus der Dichtheit der Tragesituation ergeben.

Für die Vorwärtskopplung wird ein adaptiver Ansatz vorgestellt, der auf einer Linearkombination von parallelen IIR-Filtern basiert. Der große Vorteil dieses Ansatzes ist, dass er relativ leicht mit analogen Filtern umgesetzt werden kann. Zudem werden Möglichkeiten aufgezeigt, wie die Dichtheit der Tragesituation auf effiziente und robuste Weise überprüft werden kann. Dabei werden die Vorteile der Digitaltechnik ausgenutzt, die eine Analyse der aufgenommen Mikrofonsignale in Echtzeit erlauben.

Neben den erhöhten Kosten ergeben sich aus der Digitaltechnik allerdings auch Latenzprobleme. Für Kopfhörer mit rückgekoppelter Geräuschunterdrückung wird deshalb ein Prädiktionsfilter vorgeschlagen, um die erhöhte Latenz zu kompensieren. Auch für die Rückkopplung wird eine Methode entwickelt, wie Änderungen der Tragesituation detektiert werden können.

Zum Schluss wird noch ein Verfahren zur Beurteilung der Geräuschunterdrückung vorgestellt, das sich an subjektiven Eindrücken von Benutzerinnen orientiert.

Abstract

Headphones with active noise control (ANC) cancel ambient noise by playing back destructively interfering 'anti'-noise. This thesis describes new approaches for feedforward as well as feedback ANC and treats the variability in the acoustics of headphones which results from differently tight wearing situations.

For feedforward ANC headphones, a new adaptive approach is presented that is based on a linear combination of parallel IIR filters. This approach is powerful because it allows for an analogue implementation. Besides, it is shown how the tightness of the headphones can be detected online by means of digital signal analysis.

Digital technology always suffers from the increased latency due to signal conversions. This is especially problematic in feedback ANC because the increased latency can lead to instability. A prediction filter is thus suggested to compensate for this latency. Additionally, a method for detecting changes in the wearing situation is developed for feedback ANC, too.

Finally, a procedure for the assessment of ANC headphones is derived over a regression analysis with subjective user ratings.

Preamble

In the last five years, I have been working at the Institute of Electronic Music and Acoustics for the K-projects Advanced Audio Processing (AAP) and Acoustic Sensing and Design (ASD). Both projects have been funded in context of the Competence Centres for Excellent Technologies (COMET) and are conducted by the Austrian Research Promotion Agency (FFG). Within the projects, I have been working on Active Noise Cancelling (ANC) headphones and I have developed an analogue prototype as well as digital adaptive prototypes on floating- and fix point processors. Moreover, my approach of adaptively combining analogue filters has been patented by our industrial partner AKG Acoustics GmbH. Beyond my work for the mentioned projects, I dealt with remaining unsolved topics in the area of ANC headphones.

The novel outcomes of my work have already been published in proceedings and journals and as a patent. All articles will be accessible via the IEEE Xplore Digital Library. This thesis is a compilation of my main publications. The main publications are:

- M. Guldenschuh, A. Sontacchi, R. Höldrich, 'Headphones for active noise suppression', Patent EP 2 677 765 A1, US 20 130 343 557, 12 25, 2013

  The idea for a basis transform of direction dependent ANC filters was originally formulated in the proposal of the AAP research project by A. Sontacchi and R. Höldrich. The idea to solve the direction dependent adaptation via parallel IIR filters and the Least Mean Square algorithm is my contribution and led to the referenced patent which I wrote together with a lawyer of the Patent attorneys Barger Piso & Partner. My contribution for the patent is 70%.

The content of the patent is also going to be published in the following scientific paper:

M. Guldenschuh, 'Least-Mean-Square Weighted Parallel IIR Filters in Active-Noise-Control Headphones', accepted for publication in the proceedings of EUSIPCO 2014, Sept. 2014, Lisbon

The article is entirely written by me and all related scientific work was done by me.

- M. Guldenschuh, 'Secondary-Path Models in Adaptive-Noise-Control Headphones', 3rd International Conference on Systems and Control, pp. 653-658, 29.-31- Oct. 2013, Algiers

All scientific work was entirely done by me.

- M. Guldenschuh, R. Höldrich, 'Prediction Filter Design for Active Noise Cancellation Headphones', IET Signal Processing, Vol. 7/6, pp. 497-504, Aug. 2013

Prof. Höldrich supported me with various ideas and suggestions on signal prediction and proofread the article. I did the literature review and the measurements, and I implemented and evaluated the prediction filter and wrote the article. My contribution to this work is 90%.

- M. Guldenschuh, R. de Callafon, 'Detection of Secondary-Path Irregularities in Active Noise Control Headphones' IEEE Transactions on Audio, Speech, and Language Processing, Vol. 22, Nr. 7, pp. 1148-1157, July 2014

Prof. de Callafon was my supervisor during my research stay at the University of California, San Diego. He supported me with discussions and ideas. Literature review, measurements, implementation and evaluation was done by me, and the article was entirely written by me. My contribution to this work is 90%.

- M. Guldenschuh, A. Sontacchi, M. Perkmann, M. Opitz, 'Assessment of Active Noise Cancelling Headphones', 2012 International Conference on Consumer Electronics - Berlin, pp. 299-303, 3.-5. Sept. 2012

This work was part of the AAP research project. M. Perkmann and M. Opitz were with the research and development group of AKG Acoustics GmbH. AKG did the laboratory measurements. I reviewed the

literature, conducted the listening test and did the statistical analysis. A. Sontacchi supported me with ideas and discussions. The article was entirely written by me and my contribution to this work is 70%. The article was awarded as best paper of the conference.

Danke

Danke an meinen Betreuer Prof. Robert Höldrich für die vielen Ideen und das Engagement, mit denen er diese Ideen eingebracht hat. (In diesem Sinne danke für die Matlab Codes und die händisch hergeleiteten Optimierungsalgorithmen.) Danke auch an Prof. Walter Kellermann für die Graz Reisen und die Zeit, die er sich für mich genommen hat.

Danke an unseren Institutsvorstand Dr. Alois Sontacchi. Vielen Dank für die Mitbetreuung und für den Aufbau dieses tollen Teams; und noch größeren Dank, dass ich Teil dieses Teams sein durfte.

An dieser Stelle folgt konsequenter Weise der Dank dem tollen Team und vor allem meinen Kollegen in der Petersgassenexklave für die vielen Diskussionen an den Whiteboards, in der Kaffee- oder Schokoschneckenpause und über die Schreibtische hinweg und natürlich besonders für den Spaß bei der Arbeit. (Danke meine DJs.)

Danke auch den vielen FreundInnen außerhalb des Instituts. Alle, die schon mal über einen meiner Witze gelacht haben, dürfen sich angesprochen fühlen!

Mit fortschreitendem Zeilenverbrauch werden Danksagung allem Anschein nach immer persönlicher. Es folgt mein Dank der Familie. Danke an meinen Vater für die Unterstützung, die er mir immer zuteil hat werden lassen, und danke meiner Schwester (und Leidensgenossin, die neben mir am Schreibtisch ihre Masterarbeit geschrieben hat), für ihren Rückhalt. Danke auch meinem Familienzuwachs, dem quasi-Schwager genauso wie dem Flock Clan. Ich fühle mich sehr aufgehoben bei euch.

Das bringt mich endlich zum wichtigsten Dank, dem meiner süßen Braut. Ich liebe dich! Danke, dass du bei mir bist.

<div align="right">Markus Guldenschuh, Graz, Juni 2014</div>

Contents

CONTENTS

Chapter 1

Introduction

Sound can be positive or negative in the double sense. Firstly, it is a fast change of pressure that propagates through a medium, whereas the pressure alternates between positive and negative values - meaning pushing and pulling the particles of the medium, respectively. Secondly, sound can be perceived as very useful (e.g. for communication) or very pleasant (e.g. as a nice piece of music) - thus as positive, but it can also be perceived as annoying, disturbing and negative when we talk of sound as noise.

This work uses the firstly mentioned (physical) dualism of sound to reduce the negative aspect of the latter (and psychological) dualism. It deals with headphones with Active-Noise-Cancellation (ANC).

Portable communication devices and music players allow users to communicate and enjoy music wherever they want. However in many environments, ambient noise might be so loud that it masks the music or communication signal and/or forces the user to harmfully loud playback volumes. Headphones with closed ear-cups already show a good passive attenuation of high-frequency ambient noise, but there is hardly any attenuation of the low-frequency noise as can be seen in Fig. 1.1.

Amplitude distribution
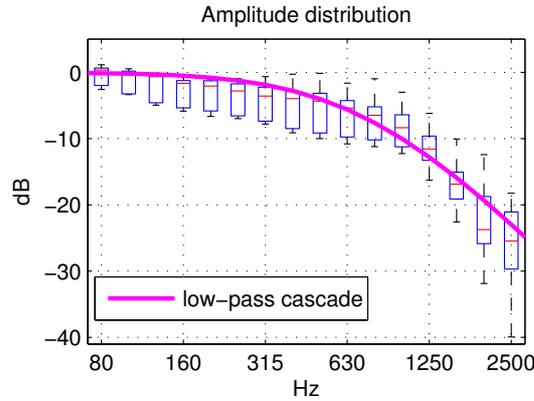
Figure 1.1: Passive attenuation of a closed ear cup. The blue boxes with their black whiskers show the distribution of passive attenuation for different directions of incident noise in 3rd octave bands. 50 % of all measured attenuations are within the blue boxes and the median passive attenuation is marked as red bar. The passive attenuation of the measured ear cup can be approximated by a cascade of two first order low pass filters with cut off frequencies at 500 Hz and 1000 Hz, respectively.

Headphones with Active-Noise-Cancellation fill this gap and reduce the low-frequency noise by playing back 'anti-noise' that destructively interferes with the ambient noise. The 'anti-noise' has the same magnitude as the penetrated noise but an opposite sign. The headphones thus produce positive directed sound pressure when the penetrated noise causes a negative sound pressure and vice versa. The superposition of both sound pressures eliminate each other. The principle of destructive superposition is shown in Fig. 1.2.

Since the 1950s, ANC-headphones have been produced for pilots because, in cockpits, clear communication is very important, but noise levels are very high. In the last decade, ANC-headphones also spread on the consumer market. Most of the currently available products use analogue circuitry because analogue circuits are less energy-consuming and less expensive than digital hardware. However, digital signal processing allows for more freedom
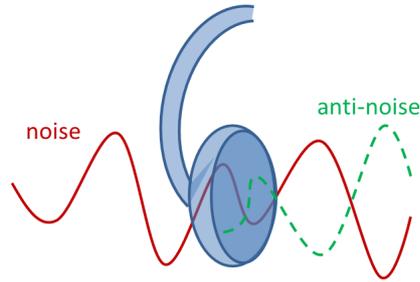
Figure 1.2: Principle of Active-Noise-Cancellation in a headphone. The headphone reproduces the penetrated sound wave, but with an opposite sign. As a consequence, both waves eliminate each other at the entrance of the ear. For illustration purpose, the sound waves are displayed as transversal waves, although in reality they are longitudinal waves, which means that the direction of pressure-change is the same as the propagation direction.

in designing and controlling ANC systems.

ANC technology can not only be divided into analogue and digital but also into feedforward and feedback control structures. Feedforward means that a reference noise from outside the headphones is fed forward to the loudspeakers as anti-noise. The structure is depicted in Fig. 1.3.
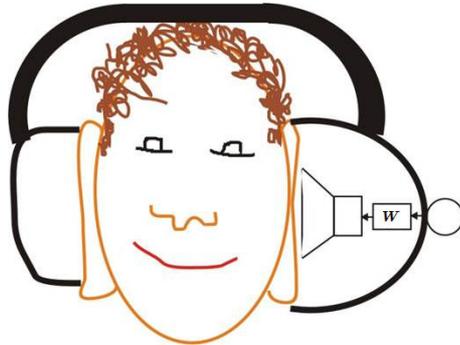
Figure 1.3: Feedforward ANC-headphones. A reference microphone outside the ear cup senses the ambient noise and feeds it forward through the controller W to the loudspeaker. The controller W passes through the inverted low frequencies of the sensed noise.

Feedback means that the residual noise (i.e. the superposition between the penetrated noise and the anti-noise) inside the headphones is fed back

to the loudspeakers as in Fig. 1.4.
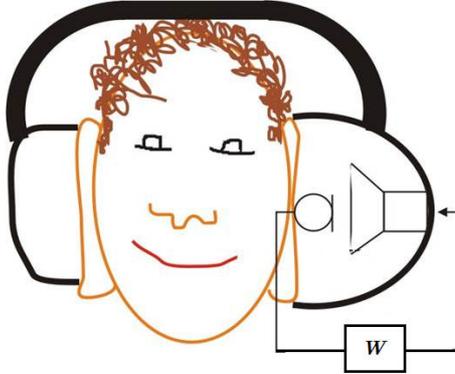
Figure 1.4: Feedback ANC-headphones. An error microphone inside the ear cup senses the residual error and feeds it back over the controller W to the loudspeaker.

While the term Active-Noise-Cancellation is mostly used for commercial products, the abbreviation ANC is mainly read as Active-Noise-Control in the scientific community. Both, feedforward and feedback controllers can be realized with analogue or digital circuits, forming the combinatory matrix of Fig. 1.5 that shortly lists the adavantages and disadvantages of each method.

Feedforward and feedback controllers will be reviewed in more detail in the following. A novel feedforward controller is presented in chapter 2 and a novel approach to design a feedback controller is presented in chapter 4. It is furthermore shown that adaptive controllers are very sensitive to changes in the acoustics of the headphones. Unfortunately, such changes unavoidably occur when the user lifts or presses against the headphones. Therefore several methods to detect and react on such changes are suggested for feedforward controllers in chapter 3 and for feedback controllers in chapter 5, respectively.

Although various different ANC headphones are commercially available
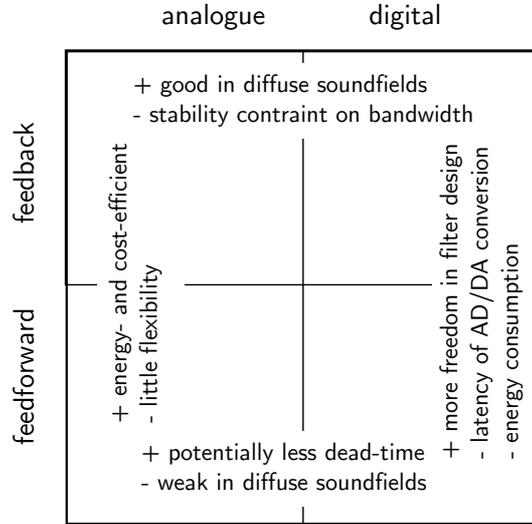
Figure 1.5: ANC can either be divided into analogue and digital technology or feedforward and feedback control.

by now, their performance assessment is still a difficult task. In appendix A, the difficulties of assessment are explained in more detail and a procedure for a reliable assessment is suggested.

1.1 Feedforward ANC

The block diagram of a feedforward ANC headphone is shown in Fig. 1.6. On the one hand, the ambient noise x penetrates the headphone and reaches the ear of the user as primary noise d via the primary path $P_1(j\omega)$. On the other hand, a reference of the noise is sensed outside the headphone (via transfer path $P_2(j\omega)$). The sensed reference noise is fed forward through the controller $W(j\omega)$. The controller output is played back as secondary anti-noise y via the loudspeaker of the headphone which - in this context - is called the secondary path $S(j\omega)$. Inside the ear cup the penetrated primary noise d and the secondary anti-noise y superpose and result in the residual noise or residual ANC error e.
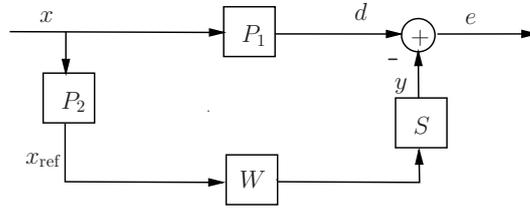
Figure 1.6: Feedforward ANC: The noise x penetrates the headphone via transfer path P_1 and reaches the (external) reference microphone via P_2. The microphone signal is fed forward through the controller W and is played back over the loudspeaker of the headphone (via the transfer path S). Inside the headphone the penetrated noise d and the played back anti-noise superpose and result in the residual noise (or residual error) e.

The primary path is often subsumed as $P(j\omega) = \frac{P_1(j\omega)}{P_2(j\omega)}$ which simplifies the block diagram to Fig. 1.7.
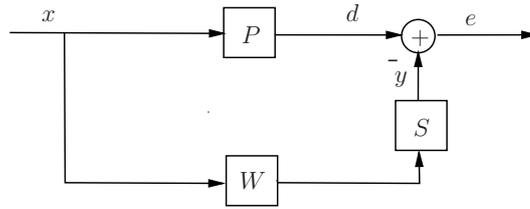
Figure 1.7: Feedforward ANC with $P = \frac{P_1}{P_2}$

The performance of an ANC system can be expressed by the sensitivity function

$$T(j\omega) = \frac{E(j\omega)}{D(j\omega)} \tag{1.1}$$

that relates the residual error to the primary noise in the frequency domain. For feedforward ANC the sensitivity function (with omitted dependency on $j\omega$) can be written as

$$T = \frac{XP - XWS}{XP} = 1 - W\frac{S}{P} \tag{1.2}$$

If the secondary noise y closely approximates the primary noise d, the

6

destructive superposition yields a small residual error e. This is the case if $W(j\omega)$ approximates the primary path $P(j\omega)$ and compensates the secondary path $S(j\omega)$. The ideal controller would thus be $W_{\text{opt}}(j\omega) = \frac{P(j\omega)}{S(j\omega)}$ as can also be deduced from eq. (1.2).

Both, the primary path $P(j\omega)$ and the secondary path $S(j\omega)$ are transfer functions to a yet unspecified target point inside the ear cup. This target point will be the position of ideal ANC. On the one hand, ANC is more effective the closer the target point is to the ear drum. On the other hand, the group delay of $S(j\omega)$ increases the further the target point is from the loudspeaker; and the relation $W_{\text{opt}}(j\omega) = \frac{P(j\omega)}{S(j\omega)}$ only remains causal, if the group delay of $P(j\omega)$ is larger than the group delay of $S(j\omega)$ for all ω.

Form Fig. 1.8 it is clear that $P(j\omega)$ itself can also become acausal e.g. if the noise incides from opposite of the external microphone as indicated by the red lines. Thus, a causal relation $W_{\text{opt}}(j\omega) = \frac{P(j\omega)}{S(j\omega)}$ is not always possible and it is advisable to choose the target point close to the loudspeaker to prevent further acausality of $W_{\text{opt}}(j\omega)$.
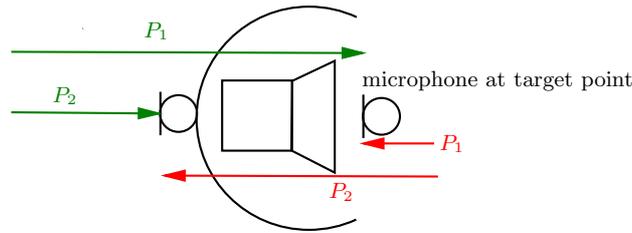
Figure 1.8: If the noise comes from the side of the external microphone, the group delay of P_1 is larger than the group delay of P_2 and the primary path $P = \frac{P_1}{P_2}$ is causal. If the noise comes from the opposite side, the group delay of P_1 is shorter than the one of P_2 and P is acausal.

Besides, ANC is desired in the low frequency band below 1000 Hz, where wavelengths and consequently also the zones of destructive superposition are large. Destructive superposition is possible as long as the phase error

between noise and anti-noise stays below 60° which corresponds to one sixth of the wavelength. At 1000 Hz, a sixth of the wavelength still has 6 cm which is much larger than the length of the ear canal which has approximately 2.5 cm. The proximity of the target point to the ear drum is thus not relevant. More important is the invertibility of $S(j\omega)$. In [Gan et al., 2005], the influence of different target points on $S(j\omega)$ is investigated. The authors show that a target point close to the intertragic notch of the pinna results in a flat frequency response of $S(j\omega)$ which is easy to invert.

In this work, on-ear prototype headphones have been used which are equipped with internal microphones close to the intertragic notch and with external microphones in the centre of the disc-shaped ear-cups. Fig. 1.9 shows measurements of the groupdelay of $P(j\omega)$ and $S(j\omega)$ of these prototype headphones. The group delay of $P(j\omega)$ is firstly determined for noise
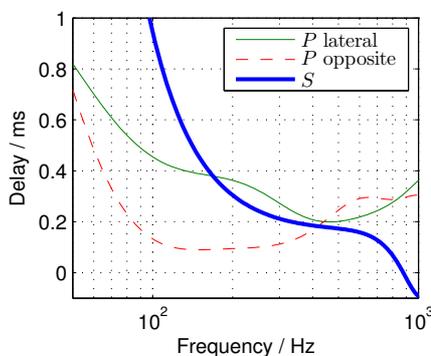
Figure 1.9: Group delay of P and S dependent on the direction of incident noise. If the noise comes from the direction of the external microphone (lateral), the group delay of P is larger than the group delay of S above 200 Hz. In this frequency range $W_{\text{opt}}(j\omega) = \frac{P(j\omega)}{S(j\omega)}$ is causal. If the noise comes from the opposite direction, the group delay of P is shorter and $W_{\text{opt}}(j\omega)$ would be acausal in a broader frequency range.

coming from the side of the external microphone. As expected, the group delay is positive in this case. Secondly, the group delay of $P(j\omega)$ is deter-

mined for noise from the opposite side. It can be seen that the group delay is still positive, although the internal microphone is about 1 cm closer to the noise source than the external microphone. The group delay is still positive because the noise experiences a phase shift when it enters the barrier of the ear cup; an effect that is also observed in [Rafaely and Jones, 2002]. Nevertheless, $W_{\mathrm{opt}}(j\omega)$ is acausal up to 400 Hz when the noise comes from the opposite direction, because the group delay of $S(j\omega)$ is larger than the one of $P(j\omega)$. For the low frequencies below 140 Hz, $W_{\mathrm{opt}}(j\omega)$ is even acausal if the noise comes from the side of the external microphone because of the large group delay of $S(j\omega)$.

The acausality of $W_{\mathrm{opt}}(j\omega)$ does not necessarily make noise cancellation impossible. Some noise cancellation is still possible for stationary noises as long as the phase lag that results from the negative group delay is smaller than 60°. In Fig. 1.9 e.g., the group delay difference Δ_g between $S(j\omega)$ and the lateral $P(j\omega)$ at $f = 100$ Hz is around 0.55 ms, which corresponds to a phase lag of only $360°\frac{\Delta_g}{\tau} = 20°$, where τ is the period length at 100 Hz.

Fig. 1.9 illustrates that $W_{\mathrm{opt}}(j\omega)$ depends on the direction of incident noise. The fitting of the headphones however also influences $P(j\omega)$ as well as $S(j\omega)$. If the headphones sit tight, they show good passive attenuation and cause a phase lag when the noise enters the headphones. For the sound reproduction, tight headphones are like a pressure chamber that allows for a strong reproduction of the low frequencies (via $S(j\omega)$) and a short group delay. If the headphones sit loose, the passive attenuation is worse and there is less phase lag when the noise enters the headphone as can be seen in Fig. 1.10 and Fig. 1.11.

For the sound reproduction, loose headphones mean that low frequency sound can radiate outside and consequently less sound pressure level (SPL)
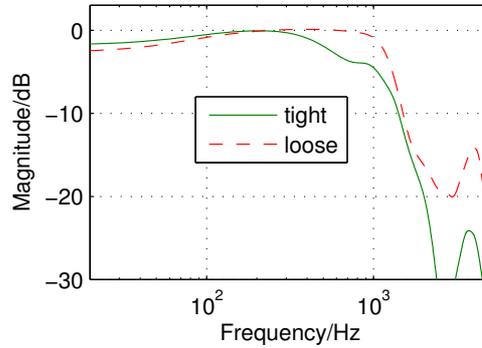
Figure 1.10: Magnitude response of P dependent on the fitting of the head-phones. Loose sitting headphones deteriorate the passive attenuation.
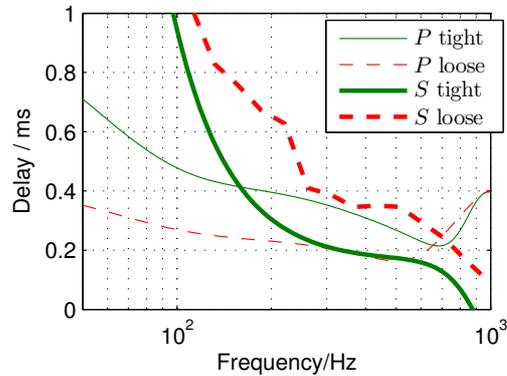
Figure 1.11: Group delay of P and S dependent on the fitting of the head-phones. The group delay of P decreases and the one of S increases if the headphones sit loosely. The consequences are that $W_{\mathrm{opt}}(j\omega)$ becomes acausal in a broader frequency range.

remains at the target point inside the ear cup. Thus an amplitude drop-off at low frequencies can be observed in the secondary path $S(j\omega)$, cf. Fig. 1.12.

Loose headphones thus have negative consequences on ANC in several aspects:

- The passive attenuation decreases.

- The temporal anticipation of the external microphone in relation to

10
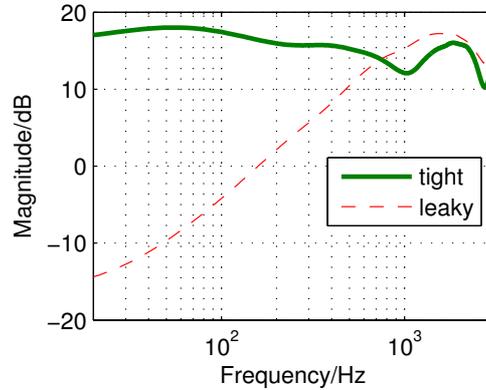
Figure 1.12: Magnitude response of S dependent on the fitting of the headphones. Due to the headphone amplifier, the magnitude of the secondary path is 20 dB above the reference voltage. For leaky sitting headphones, the magnitude drops off at low frequencies.

the interior of the ear cups decreases.

- The group delay of the secondary path increases which makes its temporal behaviour harder to invert.

- The low frequencies of $S(j\omega)$ drop off which makes its magnitude response harder to invert.

For the measurements, headphones with an external and internal microphone were worn by a test subject, while a broad band measurement signal was played back from the front of the subject in a distance of approximately 1 m. The subject was asked to once put on the headphones properly tight and an other time loosely.

Figs. 1.9 to 1.12 are worst case examples that illustrate the dependency of $W_{\mathrm{opt}}(j\omega)$ on the direction of incident noise and on the fitting of the headphones. The influences of the sound direction and the fitting of the headphones are evaluated more thoroughly in chapter 2. However, it is already clear that a static filter $W(j\omega)$ as controller can only yield limited

ANC because

- a real-life filter cannot be acausal

- a static control filter cannot be optimal for all directions of incident sound and all fittings of the headphones.

Adaptive feedforward ANC

As shown in the previous section, the optimal feedforward controller $W_{\mathrm{opt}}(j\omega)$ turns out to be acausal for some $P(j\omega)$ and $S(j\omega)$. A broadband realization of $W_{\mathrm{opt}}(j\omega)$ is not possible in these cases. However, it can still be possible to realize $W_{\mathrm{opt}}(j\omega)$ in a small bandwidth. E.g. for a static periodic primary noise x, the secondary noise only needs to have the opposite phase but it can be delayed by any integer multiple of the period.

From the above considerations, we can conclude that an adaptive controller has two main advantages over the static one:

1. It can adapt to a (causal approximation) of $W_{\mathrm{opt}}(j\omega)$ for any direction of incident sound and any fitting of the headphones.

2. It can consider the bandwidth of the excitation signal and approximate $W_{\mathrm{opt}}(j\omega)$ only in the bandwidth where it is currently needed.

In addition to the reference signal (i.e. the noise outside the headphones), an adaptive controller requires a second signal from behind the plant $P(j\omega)$ (i.e. inside the headphones). Thus a second microphone inside the ear cups is necessary as shown in Fig. 1.13.

An adaptive controller consists of two stages: Firstly an adaptive filter and secondly an adaptation algorithm that changes the transfer function of the filter. The typical realization of an adaptive filter is a digital transversal filter because of two important aspects:
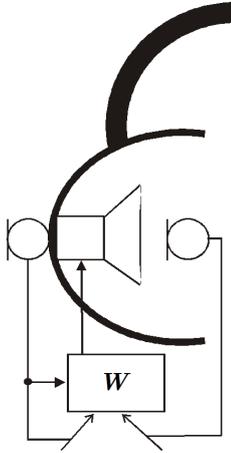
Figure 1.13: The adaptive controller requires a reference signal of the noise and an error signal from inside the headphones.

1. A transversal filter of sufficient length can approximate the impulse response of any $W(j\omega) = \frac{P(j\omega)}{S(j\omega)}$ as long as the energy of the impulse response is fading out in finite time and as long as $W(j\omega)$ is sufficiently bandlimited within $\frac{f_s}{2}$, where f_s is the sampling frequency.

2. Stability: The output of such a finite impulse response (FIR) filter is bounded as long as the filter coefficients are bounded.

Fig. 1.14 shows the block diagram of a digital adaptive ANC system.

The most popular algorithm to adapt the filter coefficients is the least mean square (LMS) because it is easy to implement and requires little computational cost [Haykin, 2001]. The LMS algorithm requires an error signal e and the reference x that caused the error. The details of the LMS will be reviewed in the beginning of chapter 2, but it can already be seen that the secondary path delays the anti-noise. Consequently, the error $e[n]$ at time index n corresponds to a reference noise sample $x[n+m]$ from m time steps ago (where m is the total delay in the secondary branch in discrete time

13
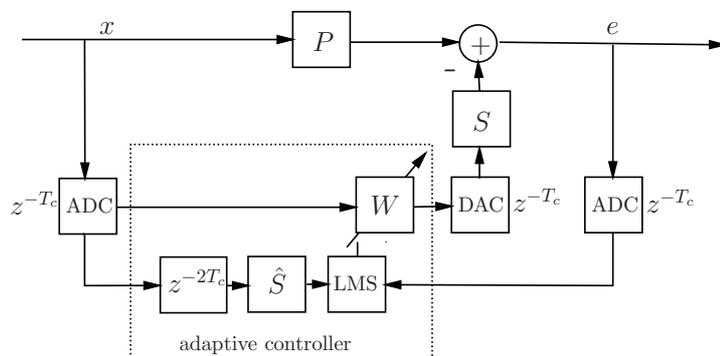
Figure 1.14: The common adaptive controller consists of an FIR filter W whose coefficients are adapted with the fxLMS algorithm.

steps). Hence, x is delayed and filtered with a model $\hat{S}(j\omega)$ of the secondary path in order to realign the reference with the error signal. The adaptation is therefore called filtered x LMS (fxLMS).

The digital to analogue converters (DACs) and analogue to digital converters (ADCs) introduce a further delay in the secondary path and the error branch. Under the assumption that DAC and ADC cause the same latency T_c, the anti-noise is additionally delayed by $2T_c$. Usually the latency of the converters is subsumed in $S(j\omega)$ and $\hat{S}(j\omega)$, respectively The additional delay in the fxLMS branch (the one that can be subsumed in $\hat{S}(j\omega)$) limits the convergence time of the adaptive filter, but it does not necessarily degrade the noise cancelling performance. However, the additional delay in the secondary branch (which can be subsumed in $S(j\omega)$) reduces the bandwidth for ANC or requires noise prediction if the noise is bandlimitted enough.

The latency is a severe problem in common audio CODECs (integrated ADC and DAC). They introduce a large latency because of their linear phase anti-aliasing and reconstruction filters. These CODECS have a talkthrough latency of at least 720 µs which impedes broadband ANC. For this reason, we developed a digital ANC prototype with high speed data converters and

minimum phase anti-aliasing filters, that reduce the talkthrough latency to 20 μs [Teschinegg, 2012]. This digital hardware allows broadband ANC up to 2000 Hz as can be seen in Fig. 1.15.
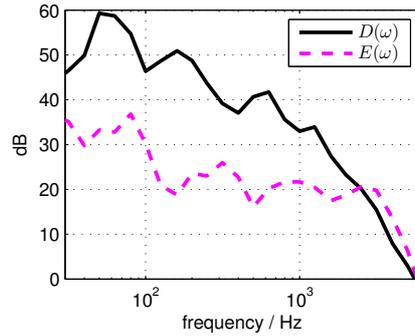
Figure 1.15: Measured primary noise spectrum $D(\omega)$ of a diffuse aeroplane noise and the error spectrum $E(\omega)$ after ANC with the digital prototype system.

Another strategy of adaptive ANC that completely avoids the latency of digital hardware is presented in chapter 2.

1.2   Feedback ANC

The block diagram of feedback ANC is shown in Fig. 1.16.
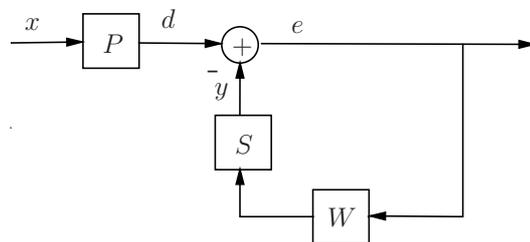
Figure 1.16: Feedback ANC

The residual error $E(j\omega)$ in this structure can be formulated (with omit-

15

ted dependency on $j\omega$) as

$$E = D - EWS = \frac{D}{1 + WS}.$$ (1.3)

Thus the sensitivity function follow directly to

$$T(j\omega) = \frac{1}{1 + W(j\omega)S(j\omega)}.$$ (1.4)

From the above two equations, it is clear that the error becomes small if the denominator of eq. (1.3) or (1.4) is large. The controller $W(j\omega)$ thus needs to amplify the sensed error signal. However, this amplification must not be unconstrained. Fig. 1.16 shows the block diagram of a negative feedback. Due to the negative feedback, the anti-noise y interferes destructively with the primary noise d. However, if the phase of the open loop $W(j\omega)S(j\omega)$ turns 180°, the feedback causes a constructive superposition, instead. The feedback would then run unstable if the error signal is not sufficiently damped at this frequency.

A phase shift of 180° occurs at the frequency where half the wavelength equals the group delay of the open loop. For a distance $d_S = 1$ cm between the loudspeaker membrane and the error microphone and a speed of sound of $c = 344$ ms, the group delay $\tau_g = \frac{d_S}{c}$ is at least 30 µs. Thus 180° phase shift would occure at 17.2 kHz latest. In Fig. 1.9 and 1.11, we have seen that the actual group delay of a real electroacoustic system is actually much larger. Our prototype headphones e.g. have the 180° phase shift at 4000 Hz as can be seen in Fig. 1.17.

Thus the controller needs to be a filter that amplifies the low frequencies with little phase shift and attenuates the high frequencies where the phase has turned 180°. Several methods have been proposed to design such a
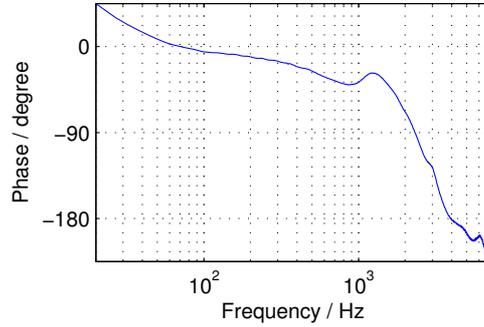
Figure 1.17: Phase response of a headphone's secondary path under regular tight wearing conditions.

controller [Laroche, 2007, Yu and Hu, 2001, Zhang et al., 2013, Francis and Zames, 1984, Doyle et al., 1989], but the resulting frequency responses tend to be very complex such that they cannot be realized by analogue filters with reasonable order. During my thesis, an analogue feedback controller has been developed, too, and a shelving filter turned out to be the most practical solution. It amplifies the low and attenuates the high frequencies with a phase lag of maximally 90°. It therefore yields good ANC results and it is comparatively easy to tune. The measured result can be seen in Fig. 1.18.

(a) Primary noise spectrum $D(\omega)$ and error spectrum $E(\omega)$ after ANC
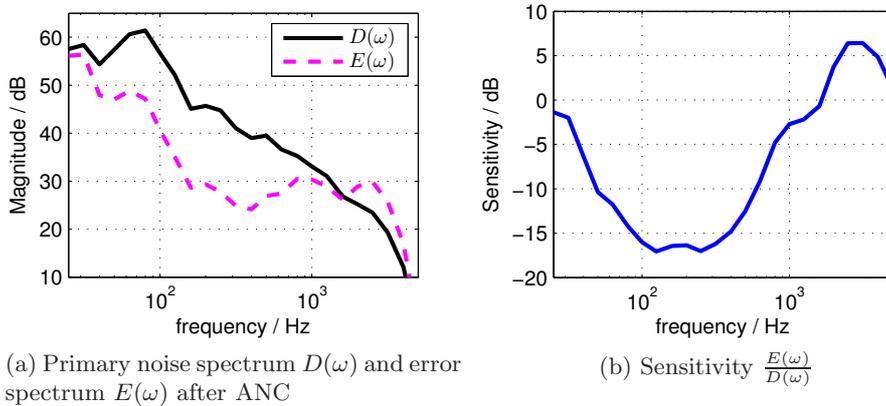
(b) Sensitivity $\frac{E(\omega)}{D(\omega)}$

Figure 1.18: Measured performance of the analogue feedback prototype for aeroplane noise: Noise spectra (a) and the corresponding sensitivity (b)

17

Digital filters are better suited to realize complex frequency responses, but the additional group delay of digital systems counteracts these benefits. Chapter 4 deals with digital feedback ANC and proposes a prediction filter design to compensate the latency.

1.3   Outline of the Thesis

Chapter 2 reviews the LMS algorithm and its convergence properties and suggests a parallel adaptive linear combiner (ALC) instead of the common adaptive transversal filter. The ALC consists of three parallel static filters whose outputs are scaled with an adaptive gain before they are added up. The combination of the three filters has to approximate all the variations in $P(j\omega)$ and $S(j\omega)$ which are dependent on the direction of sound and the fitting of the headphones.

The main variations in $P(j\omega)$ and $S(j\omega)$ are assessed via preliminary measurements and a principal component analysis of the measured transfer functions is used to derive the three prototype filters. These prototype filters can be modelled as IIR filter which saves computational power compared to FIR filters.

It is shown that the parallel IIR filters yield the same results as the adaptive transversal filter with fewer update weights. Chapter 2 is based on the article: M. Guldenschuh, 'Least-Mean-Square Weighted Parallel IIR Filters for Active Noise Control Headphones', accepted for publication in the proceedings of EUSIPCO 2014, Sept. 2014, Lisbon; and the method which is presented in this chapter is based on the patent: M. Guldenschuh et al., 'Headphones for active noise suppression', Patent EP 2 677 765 A1, US 20 130 343 557, 12 25, 2013

Chapter 3 treats the dependency of adaptive feedforward ANC on the accuracy of the secondary-path model $\hat{S}(j\omega)$. The secondary-path model can be derived over an initial measurement, but it is shown that changes in the wearing situation of the headphones severely change the secondary-path, too. Therefore it has been suggested in literature to do a online secondary-path estimation by constantly playing back broadband noise. Since this is very counterproductive for ANC headphones, chapter 3 suggests two robust methods for online secondary-path estimation that work without disturbing measurement noise.

In the first method, the noise cancelling performance is observed in real-time by comparing the signal powers of x and e. Good noise cancellation is only possible if the headphones sit tight. Poor noise cancellation indicates a leaky wearing situation and consequently the secondary-path model has to be changed accordingly.

The second method uses an infrasonic measurement signal. The infrasonic signal is not disturbing for the user but very suitable for detecting the changes in $S(j\omega)$ that occur when the headphones are lifted. A comparison with approaches from literature prove the advantages and robustness of the presented methods. Chapter 3 is based on the following article: M. Guldenschuh, 'Secondary-Path Models in Adaptive-Noise-Control Headphones', 3$^{\text{rd}}$ International Conference on Systems and Control, pp. 653-658, 29.-31- Oct. 2013, Algiers

Chapter 4 deals with the latency problem in feedback ANC. Adaptive systems try to compensate the latency by predicting the noise signal. Prediction is possible if the noise signal is bandlimited, and in feedback ANC the noise is naturally bandlimited due to the passive attenuation of the ear

cups. Since this bandlimitation is consistent over time, it is suggested to implement a static predictor instead of an adaptive filter in order to reduce the computational complexity.

It is shown that this static predictor yields the same performance as adaptive methods and is more robust against changes in the secondary path. Chapter 4 is based on the following article: M. Guldenschuh, R. Höldrich, 'Prediction Filter Design for Active Noise Cancellation Headphones', IET Signal Processing, Vol. 7/6, pp. 497-504, 2013

Chapter 5 treats the dependency of adaptive feedback ANC on the accuracy of the secondary path model. In adaptive feedback ANC, the model $\hat{S}(j\omega)$ is required two times. Firstly, to generate an internal model of the reference and secondly for the fxLMS adaptation. This makes feedback ANC especially sensitive to changes in $S(j\omega)$. It is shown that common changes in the wearing situation, like lifting the headphones, can drive the feedback unstable if no measures are taken. Chapter 5 proposes two measures to tackle this problem.

Firstly, a leaky version of the fxLMS is used to avoid divergence of the adaptive filter. Secondly, it is shown that the low frequenciy region of $S(j\omega)$ drops off if the headphones are lifted. The adaptive filter $W(j\omega)$ tries to compensate this drop-off by amplifying the low frequencies. Thus, changes in $S(j\omega)$ can be detected by the low frequency behaviour of $W(j\omega)$. It is further shown that it is sufficient to observe the DC-gain of $W(j\omega)$ which is beneficial because the DC-gain can easily be obtained by summing up the filter coefficients. The constraint on the DC-gain is compared with other stability constraints from literature and it is shown that it yields equal if not better results with less computational complexity. Chapter 5 is based on the

following article: M. Guldenschuh, R. de Callafon, 'Detection of Secondary-Path Irregularities in Active Noise Control Headphones' IEEE Transactions on Audio, Speech, and Language Processing, Vol. 22, No. 7, pp 1148-1157, 2014

The chapter 'General Conclusion' summarizes the essentials of each chapter and forms links between the individual contributions; and in the end, appendix A presents a procedure to assess ANC headphones. Usually, ANC headphones are put on some kind of artificial ear or mannequin that contains a measurement microphone in place of the ear drum. Measurement noise is played back from outside and the noise reduction is evaluated by comparing the measured noise with and without ANC. It is revealed that different artificial ears lead to different ANC results and that even the ranking of different headphones is not consistent over different artificial ears. Therefore, subjective user ratings are collected and these ratings are regressed with the measurement results. Finally, a objective measurement procedures is developed that reflects the user ratings. Appendix A is based on the following article: M. Guldenschuh et al., 'Assessment of Active Noise Cancelling Headphones', 2012 International Conference on Consumer Electronics - Berlin, pp. 299-303, 3.-5. Sept. 2012

Chapter 2

Adaptive Linear Combiner for feedforward ANC Headphones

based on:

Markus Guldenschuh 'Least-Mean-Square Weighted Parallel IIR Filters in Active-Noise-Control Headphones', accepted for the Proceedings of EUSIPCO 2014, Lisbon;

and

M. Guldenschuh et al., 'Headphones for active noise suppression', Patent EP 2 677 765 A1, US 20 130 343 557, 12 25, 2013

As mentioned in the introduction, a static feedforward controller can only yield optimal ANC for limited directions of sound and for one defined fitting of the headphones. If the controller is designed as an adaptive filter, these limitations are overcome, but the implementation of an adaptive filter is more complex and costly. In this chapter, it is shown that the adaptive

combination of three static filters yields equally good results as a common fully adaptive filter, and it requires less adaptation effort.

The common adaptive filter is an FIR filter. FIR filters have the advantage of producing a bounded output if the input is bounded (i.e. they are BIBO stable), but they need considerably more filter coefficients to yield similar transfer functions as IIR filters. Also, FIR filters can only approximate pole-zero systems with additional zeros, while IIR filters are able to exactly reproduce these systems. However besides stability issues, IIR adaption via the LMS algorithm might not converge to the Minimum Mean Square Error (MMSE) or it might result in biased coefficients [Shynk, 1989]. Therefore it has been proposed to keep the poles fixed and adapt the zeros only [Bank, 2008, Carusone and Johns, 2003, Williamson and Zimmermann, 1996, Kaelin et al., 1993, Zeng and de Callafon, 2003, Karjalainen and Paatero, 2006]. The pole locations are suggested to be chosen according to preliminary offline measurements of the system.

In the case of ANC headphones, the primary path $P(j\omega)$ and the secondary-path $S(j\omega)$ have to be measured for different directions of incident sound and different fittings of the headphones. Since not only the pole but also the zero locations are readily known from such measurements, this chapter proposes to determine the relevant zero-pole combinations via a principal component analysis (PCA) and to implement the resulting fixed IIR filters in parallel as adaptive linear combiner (ALC) like in Fig. 2.1.

The ALC weights the filter output with the scalar coefficients $w_{1\ldots3}$ and the superposition of the weighted outputs are played back as anti-noise. The coefficients can be updated via the fxLMS algorithm, equivalently to the common adaptive filter. The performance of the parallel IIR filters is compared with a common adaptive transversal-filter and an adaptive-
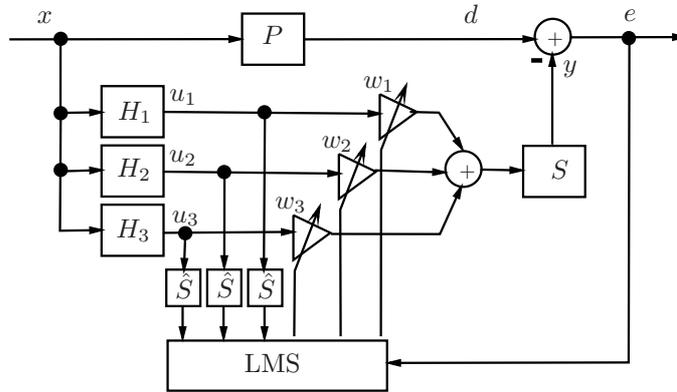
Figure 2.1: Adaptive noise control with three parallel filters, an adaptive linear combiner and fxLMS update.

zeros/fixed-pole system. Although the fixed pole approaches have hardly gained any consideration, they show equally good results as the popular transversal filter, but have better convergence properties and require fewer adaptation weights. This is especially beneficial if an analogue implementation of the IIR filters and the LMS algorithm is desired. Analogue implementations are faster and less energy consuming, but they suffer from wrong weight updates due to DC offsets in analogue active circuits [Johns et al., 1991, Shoval et al., 1995]. Fewer adaptive weights therefore mean less effort to handle the DC offset. Our idea of implementing an adaptive linear combiner with parallel analogue filters has been patented in [Guldenschuh et al., 2013b]. Although this chapter deals with discrete time IIR filters, it implies the methodology of the patent and can be read as instruction to derive appropriate analogue filters, too.

2.1 Review of the Least Mean Squares (LMS) adaptation

2.1.1 Method of Steepest Descent

The LMS algorithm as practical case of the method of steepest descent is well described in [Widrow et al., 1976, Kuo et al., 1996, Snyder and Hansen, 1990, Haykin, 2001, Widrow et al., 1975, Burgess, 1981, Nelson et al., 1993, Kuo and Morgan, 1999]. In the following the most important properties of the algorithm are reviewed.

The method of steepest descent and the LMS respectively are used to find the minimum of a multidimensional convex cost function. In the case of ANC, a convex cost function J is given by the expected squared adaptation-error $J(\mathbf{w}) = E\{e^2\}$ which depends on the weight vector $\mathbf{w}$. The weight vector can either contain the coefficients of a common adaptive FIR filter or the output weights of the parallel filters as in Fig. 2.1. With a neglected secondary-path, the cost function reads as

$$J(\mathbf{w}) = E\{|d - \mathbf{w}^{\mathrm{T}}\mathbf{u}|^2\} = E\{d^2\} - 2\mathbf{w}^{\mathrm{T}}E\{\mathbf{u}d\} + \mathbf{w}^{\mathrm{T}}E\{\mathbf{u}\mathbf{u}^{\mathrm{T}}\}\mathbf{w}, \quad (2.1)$$

where $\mathbf{u}$ is the vector of inputs to the adaptive weights. In case of the parallel structure of Fig. 2.1, $\mathbf{u}$ contains the outputs of the L parallel filters at time instance n: $\mathbf{u} = [u_1[n], u_2[n], \dots, u_L[n]]^{\mathrm{T}}$. In the case of the common adaptive transversal filter, $\mathbf{u}$ contains the last L reference noise samples: $\mathbf{u} = [x[1], x[2], \dots, x[L]]^{\mathrm{T}}$. The influence of the secondary path will be treated in the following in section 2.1.2.

With the cross-correlation vector $\mathbf{p} = E\{\mathbf{u}d\}$ and the auto-correlation

matrix $\mathbf{R} = E\{\mathbf{u}\mathbf{u}^{\mathrm{T}}\}$ and with $\sigma_d^2 = E\{d^2\}$, the cost function reads as

$$J(\mathbf{w}) = \sigma_d^2 - 2\mathbf{w}^{\mathrm{T}}\mathbf{p} + \mathbf{w}^{\mathrm{T}}\mathbf{R}\mathbf{w}. \tag{2.2}$$

The method of steepest descent finds the minimum of $J(\mathbf{w})$ by iteratively changing the weight vector $\mathbf{w}$ according to the negative gradient ∇ of the cost function.

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \mu\nabla, \tag{2.3}$$

where μ is the step size that controls the speed of convergence.

The gradient is the deviation of eq. (2.2) and is given as

$$\nabla = \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^{\mathrm{T}}} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}. \tag{2.4}$$

The auto-correlation matrix is symmetric and can be decomposed into $\mathbf{R} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$, where $\mathbf{Q}$ is a unitary matrix of eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues λ_l.

To asses the convergence properties, it is useful to define a misalignment vector $\mathbf{v}[n] = \mathbf{w}[n] - \mathbf{w}_{\mathrm{opt}}$ with the optimal solution of the weight vector $\mathbf{w}_{\mathrm{opt}}$. By rotating the misalignment vector to the principal axis of the multidimensional cost function $\mathbf{v}' = \mathbf{Q}^{-1}\mathbf{v}$, a decoupled system of equations is yielded

$$v_l'[n+1] = (1 - 2\mu\lambda_l)v_l'[n], \tag{2.5}$$

for $l = 1 \ldots L$. With the initial value $v_l'[0]$, eq. (2.5) can be written as a geometric sequence

$$v_l'[n] = (1 - 2\mu\lambda_l)^n v_l'[0]. \tag{2.6}$$

This allows deducing the convergence behaviour. The method of steepest

27

descent converges if $0 < \mu < \frac{1}{\lambda_{\max}}$ and it converges fastest if $\mu = \frac{1}{2\lambda_{\max}}$, where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{R}$ [Kuo et al., 1996, Snyder and Hansen, 1990, Haykin, 2001].

The geometric sequence can be described by the exponential function

$$(1 - 2\mu\lambda_l) = e^{(\frac{-1}{\tau_l})}, \tag{2.7}$$

where τ_l is the convergence time along the lth principal axis of the cost function. With a linear estimation of the exponential function, the over all convergence time τ lies between [Haykin, 2001]

$$\frac{1}{2\mu\lambda_{\max}} \leq \tau \leq \frac{1}{2\mu\lambda_{\min}}. \tag{2.8}$$

Since the step size is ought to be set to $\mu = \frac{1}{2\lambda_{\max}}$, the largest possible convergence time is determined by the ratio between the largest and the smallest eigenvalue, i.e. the condition number of $\mathbf{R}$.

2.1.2 Least Mean Square Algorithm (LMS)

The LMS uses the method of steepest descent, but since $\mathbf{p}$ and $\mathbf{R}$ are difficult to obtain, the LMS utilises the estimated gradient

$$\nabla = -2e[n]\mathbf{u}[n],$$

which leads to the following update

$$\mathbf{w}[n + 1] = \mathbf{w}[n] + 2\mu e[n]\mathbf{u}[n]. \tag{2.9}$$

In many text books (e.g. [Haykin, 2001]), the scalar factor 2 is subsumed in the step size μ.

In acoustical systems, the anti-noise y is delayed by the secondary-path $S(j\omega)$ (i.e. the transfer function from the loudspeaker to the error microphone, including ADC and DAC in digital systems). Thus, the noise x has to be filtered with a model $\hat{S}(j\omega)$ of the secondary-path in order to apply the same delay to the reference input of the LMS, cf. [Kuo and Morgan, 1999]. This algorithm is therefore called filtered-x-LMS (fxLMS) and its update reads as

$$\mathbf{w}[n+1] = \mathbf{w}[n] + 2\mu e[n]\mathbf{x}'[n], \tag{2.10}$$

where $\mathbf{x}'$ is the filtered input vector.

In the case of the common transversal adaptive filter, $\mathbf{x}'^{\mathrm{T}} = \hat{\mathbf{s}}^{\mathrm{T}}\mathbf{X}$, where $\hat{\mathbf{s}}$ is the N taps long impulse response of the secondary-path estimate and $\mathbf{X}$ is a matrix of reference noise samples, i.e.:

$$\mathbf{x}'^{\mathrm{T}} = \begin{bmatrix} \hat{s}_{[1]} & \hat{s}_{[2]} & \cdots & \hat{s}_{[N]} \end{bmatrix} \begin{bmatrix} x_{[n]} & x_{[n-1]} & \cdots & x_{[n-L+1]} \\ x_{[n+1]} & x_{[n]} & \ddots & x_{[n-L+2]} \\ \vdots & \ddots & \ddots & \vdots \\ x_{[n+N-1]} & x_{[n+N-2]} & \cdots & x_{[n+N-L]} \end{bmatrix}. \tag{2.11}$$

If we take the convolution matrix $\hat{\mathbf{S}}$ of the secondary-path estimate instead of the impulse vector $\hat{\mathbf{s}}$, we yield a convolution matrix of the filtered

input $\mathbf{X}' = \hat{\mathbf{S}}\mathbf{X}$, i.e.:

$$
\mathbf{X}' = \begin{bmatrix}
0 & \cdots & 0 & \hat{s}_{[1]} \\
\vdots & \cdot^{\cdot^{\cdot}} & \hat{s}_{[1]} & \hat{s}_{[2]} \\
\vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\
0 & \hat{s}_{[1]} & \cdot^{\cdot^{\cdot}} & \hat{s}_{[N-1]} \\
\hat{s}_{[1]} & \hat{s}_{[2]} & \cdot^{\cdot^{\cdot}} & \hat{s}_{[N]} \\
\hat{s}_{[2]} & \hat{s}[3] & \cdot^{\cdot^{\cdot}} & 0 \\
\vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\
\hat{s}_{[N-1]} & \hat{s}_{[N]} & \cdot^{\cdot^{\cdot}} & \vdots \\
\hat{s}_{[N]} & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
x_{[n]} & x_{[n-1]} & \cdots & x_{[n-L+1]} \\
x_{[n+1]} & x_{[n]} & \ddots & x_{[n-L+2]} \\
\vdots & \ddots & \ddots & \vdots \\
x_{[n+N-1]} & x_{[n+N-2]} & \cdots & x_{[n+N-L]}
\end{bmatrix}.
$$

(2.12)

This convolution matrix is required to compute the input auto-correlation matrix $\mathbf{R} = \mathbf{X}'^{\mathrm{T}}\mathbf{X}'$.

In the case of the parallel adaptive linear combiner, the reference noise is additionally filtered with the filters $H_l(j\omega)$ yielding $\mathbf{X}' = \mathbf{H}^{\mathrm{T}}\hat{\mathbf{S}}\mathbf{X}$, where the columns of $\mathbf{H}$ are the impulse responses of the filters $H_l(j\omega)$,

$$
\mathbf{X}' = \begin{bmatrix}
h_{1[1]} & h_{1[2]} & \cdots & h_{1[2N-1]} \\
h_{2[1]} & h_{2[2]} & \cdots & h_{2[2N-1]} \\
\vdots & \cdots & \cdots & \vdots \\
h_{L[1]} & h_{L[2]} & \cdots & h_{L[2N-1]}
\end{bmatrix} \hat{\mathbf{S}}\mathbf{X}.
$$

(2.13)

The secondary path model $\hat{S}(j\omega)$ and the filers $H_l(j\omega)$ respectively influence the convergence behaviour because they influence the auto-correlation matrix $\mathbf{R} = \mathbf{X}'^{\mathrm{T}}\mathbf{X}'$ and its eigenvalues.

Furthermore, the group delay of $S(j\omega)$ impedes an immediate update and decreases the maximum allowable step-size $\mu_{\max}$ by the factor $\sin[(\pi/(2(2\Delta+$

1))] where Δ is the group delay in samples [Snyder and Hansen, 1990]. Additionally, $\mu_{\max}$ is reduced by modelling errors of $\hat{S}(j\omega)$, and phase modelling errors larger than 90° could even lead to divergence of the adaptation process [Snyder and Hansen, 1990, Lopes and Piedade, 2004, Snyder and Hansen, 1994].

In cases of large phase errors, the divergence can be prevented by penalizing the norm of the adaptive filter. The cost function then reads as $J(\mathbf{w}) = E\{e^2 + \gamma(\mathbf{w}^T\mathbf{w})\}$ yielding the update algorithm

$$\mathbf{w}[n+1] = (1 - \mu\gamma)\mathbf{w}[n] + 2\mu e\mathbf{x}', \tag{2.14}$$

where γ is a trade-off factor between the constraint on $\mathbf{w}$ (and thus stability) and the minimum mean square error (MMSE) which will be increased by the suboptimal $\mathbf{w}$.

2.2 Primary and Secondary Path Variations

The optimum feedforward controller has to approximate the primary path $P(j\omega)$ and compensate for the secondary-path $S(j\omega)$. In the frequency domain, its transfer function therefore reads as $K(j\omega) = \frac{P(j\omega)}{S(j\omega)}$. It is advantageous to have an adaptive filter because both paths, $P(j\omega)$ and $S(j\omega)$, may change during the usage of the ANC headphones.

- The primary path $P(j\omega)$ is the sound pressure relation between the internal microphone and the external microphone whereas both sensors are excited by incident noise. Firstly, this pressure relation depends on the passive attenuation of the earcups and therefore on the tightness of the wearing situation. Secondly, the phase relation the outer and inner microphone signal depends on the direction of incident sound.

- The secondary-path $S(j\omega)$ is the transfer function from the loud-speaker to the error microphone inside the earcup and includes the group delay of AD and DA conversion in digital applications as in Fig. 1.14. This transfer function again depends on the tightness of the headphones. In the tight case, the earcup is like a pressure chamber that allows reproducing enough low frequency sound pressure to get a flat magnitude response of $S(j\omega)$. In the leaky case, the low frequency sound power radiates outside and consequently it is diminishes inside the headphones. The consequence is a magnitude drop-off at low frequencies of $S(j\omega)$.

It has to be assumed that the direction of incident sound changes permanently while wearing ANC headphones. Be it because the noise source moves (e.g. traffic noise), or be it because the user turns the head. And also changes in the tightness of the headphones may occur if the headphones are shifted or slightly lifted - especially if non-circumaural headphones are used.

2.2.1   Plant Measurements

$P(j\omega)$ and $S(j\omega)$ are measured for all these main potential variations. The headphones were put on a mannequin which was placed on a turntable in the centre of a circular vertical loudspeaker array as in Fig. 2.2. The turntable turns in 30° steps and the 14 loudspeakers are spaced in 12.5°. Hence, $P(j\omega)$ was measured for 168 different directions.

The 168 measurements were performed once for tight sitting headphones and once for leaky sitting headphones. For the latter, the headphones were shifted back such that a leak of approximately 2 mm in diameter was provoked between the earcups and the intertragic notch of the mannequin's ear. The secondary path $S(j\omega)$ was measured by playing back the measurement
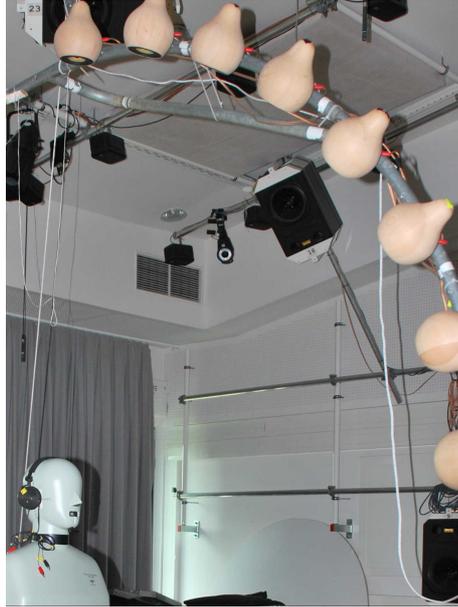
Figure 2.2: Setup to measure $P(j\omega)$ and $S(j\omega)$ for different directions of incident sound and differently tight sitting headphones. The loudspeakers are arranged on a vertical arc with radius 1.5 m and spaced every 12.5° from -75° to 87.5°. The headphones are in the centre of the arrangement. They are put on a mannequin which stands on a turntable. The turntable turns 12 times 30°.

signal via the headphones and the results are shown in Fig. 2.3. As mentioned earlier, a magnitude drop-off can be observed at low frequencies if the headphones are worn in a leaky manner.

The maximum phase difference between the outer and inner microphones occurs when the sound comes from the side as in Fig. 1.8. Disregarding the phase shift due to the physical barrier, the maximum phase difference can be approximated as $\Delta_\phi = 2\pi\frac{\mathrm{d}}{\lambda}$, where d is the distance between the microphones and λ is the wave length. Consequently, the maximum phase difference between two opposite lateral sounds yields $4\pi\frac{\mathrm{d}}{\lambda}$.

Phase mismatches between the anti-noise y and the penetrated noise d larger than +/-60° lead to constructive instead of destructive interference. Direction independent ANC is therefore only possible as long as $\frac{2\pi}{3} > 4\pi\frac{\mathrm{d}}{\lambda}$;
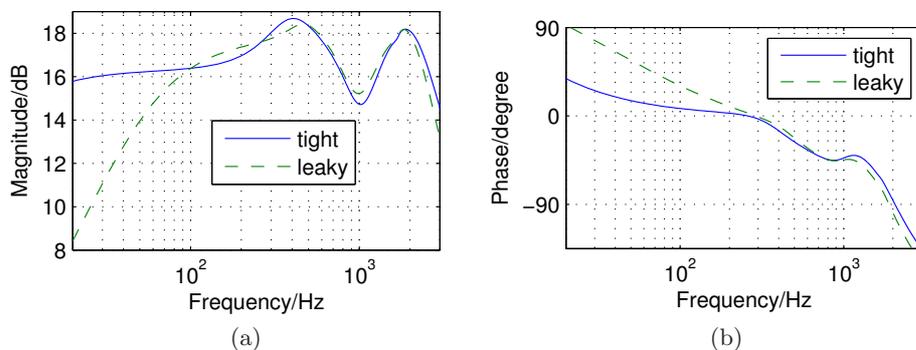
Figure 2.3: (a) Magnitude and (b) phase response of $S(j\omega)$ for tight and leaky positioned headphones: In the leaky case, a magnitude drop-off at low frequencies can be observed.

or in terms of the frequency

$$f < \frac{c}{6\mathrm{d}}, \tag{2.15}$$

where c is the speed of sound of $344\,\mathrm{m/s}$. For a distance d of approximately 2.5 cm, 2300 Hz can thus be considered as cut-off frequency for broadband ANC.

Fig. 2.4 shows the measured magnitude and phase variation of $P(j\omega)$ in the horizontal plane. It can be seen that 60° phase deviation from the mean are actually reached at approx. 2600 Hz.

In the median plane, the phase difference between the inner and outer microphone is always the same. Thus hardly any variation of $P(j\omega)$ is expected. Fig. 2.5 shows these variation for tight and leaky sitting headphones. As expected, there is hardly any variation if the headphones sit tight. However, it can be seen that leaks between the headphones and the ears have a considerable influence on $P(j\omega)$.

In total, we have 336 measurements of $K_{\eta,i}(j\omega) = \frac{P_\eta(j\omega)}{S_i(j\omega)}$ for $\eta = 1\ldots168$ and $i = 1, 2$. For the sake of simplicity, the indices η, i will be subsumed as $m = 1\ldots336$.

(a) (b)
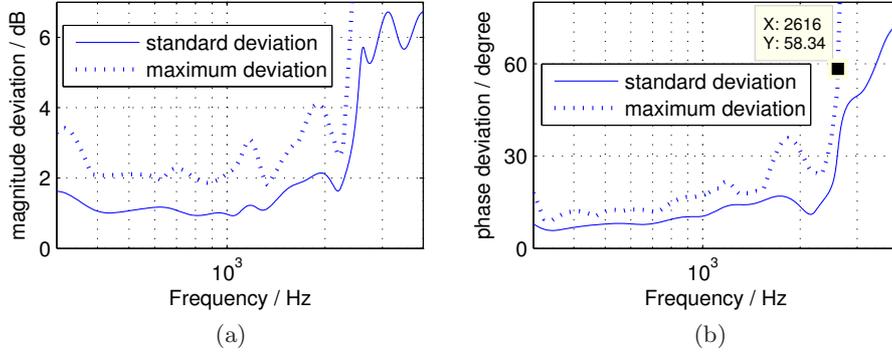
Figure 2.4: Deviation of (a) the magnitude and (b) the phase from the average primary path for tight sitting headphones over 12 azimuth angles (from 0° to 330°) evaluated at 0° elevation. At approx. 2600 Hz the maximum phase deviation reaches 60° which marks the maximum bandwidth of direction independent ANC.

(a) (b)
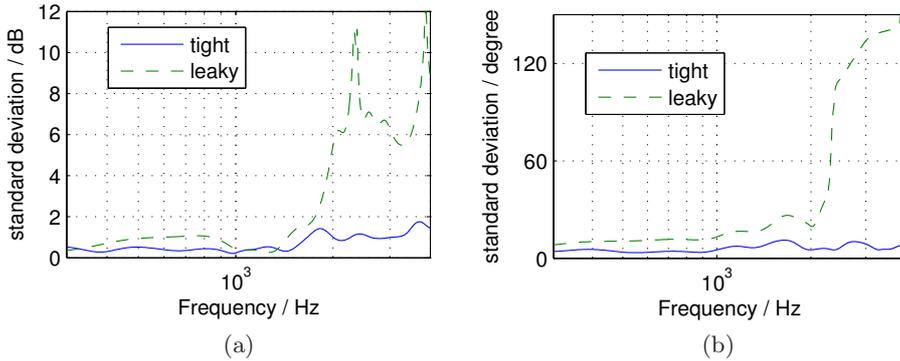
Figure 2.5: Standard deviation of the magnitude (a) and phase (b) of $P(j\omega)$ over 14 different elevation angles and 0° azimuth.

In the next step, the standard deviation of all $K_m(j\omega)$ is examined. Fig. 2.6 shows the spatial measurement points that result form the measurement setup described in Fig. 2.2. It can be seen that the poles of the spherical measurement grid are sampled much denser than the equator. Therefore, each measurement point is weighted by its surrounding spherical surface in order to yield a comparable set of data. This is equivalent to weighting the sampling points according to their elevation angle ϑ with $\cos(\vartheta)$.

35
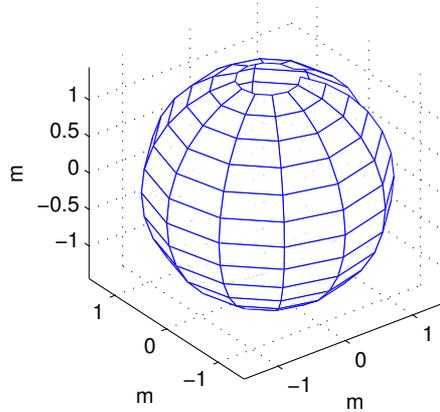
Figure 2.6: The junctions of the spherical grid indicate the different measurement positions which result form the setup that is described in Fig. 2.2.

Fig. 2.7 shows the standard deviation of the weighted $K_m(j\omega)$. Between
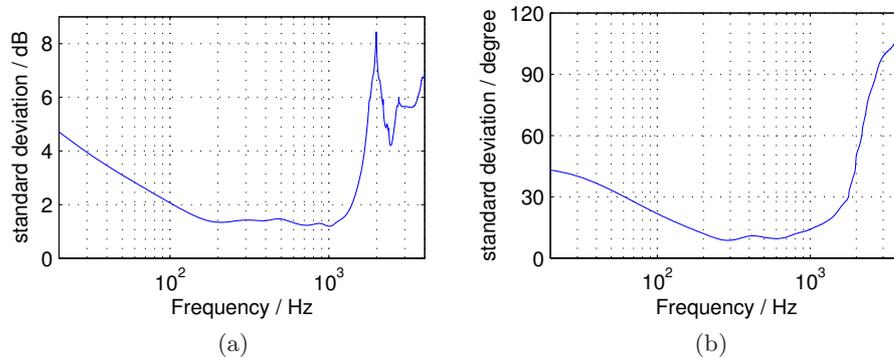
Figure 2.7: Weighted standard deviation of (a) magnitude and (b) phase of $K(j\omega)$ over all measurements.

200 Hz and 1000 Hz neither the magnitude nor the phase varies a lot. In this frequency band, robust ANC would be possible with one static filter only. However, for the frequencies below 200 Hz and above 1000 Hz, an adaptive filter is required to yield robust ANC performance.

Since we already know (the main) required filter variations, we do not

need a fully adaptive filter. The problem is much more: Is there a set of filters $H_l(j\omega)$ that can model all 168×2 transfer-functions $K_m(j\omega)$ in the form of the ALC in Fig. 2.1. A solution can be obtained by means of a Principal Component Analysis (PCA).

2.3 Extraction of Prototype Filters for the Adaptive Linear Combiner

2.3.1 Principal Component Analysis of the Required ANC Filters

The PCA changes the basis of a coordinate system with the goal to reduce the redundancy of dimensions for a given set of data [Jackson, 1991]. Its first principal component is the new coordinate in which the data has the largest variance. The second principal component is orthogonal to the first and covers the largest remaining variance of the data, and so forth.

Our set of data consists of the transfer functions $K_m(j\omega) = \frac{P_\eta(j\omega)}{S_i(j\omega)}$ which can be gathered column-wise into the matrix $\mathbf{K}$. The secondary-path in the denominator is the transfer-function from the loudspeaker to the error microphone inside the earcup, and it is not minimum phase. Thus its inverse is acausal. The same applies to $P(j\omega)$ for those directions where the sound hits the error microphone earlier than the reference microphone. Since the acausal parts cannot be cancelled by the ALC, only the causal parts of $K_m(j\omega)$ are considered in $\mathbf{K}$.

In addition, it is the goal to obtain real-valued gain factors w_l. Therefore, the used data set has to be real valued, too. With $K_m(k)$ being the discrete Fourier transform of the transfer functions for the frequency bins $k = 1 \ldots N$ There are two ways to achieve a real valued matrix $\mathbf{K}$.

1. $K_m(k)$ can be transformed into the time domain. Then, matrix $\mathbf{K}$ contains the impulse responses of $K_m(k)$ in its columns.

2. $K_m(k)$ has to be split into real and imaginary parts. The columns of $\mathbf{K}$ are then concatenated as

$$[\Re\left\{[K_m(0), K_m(1), K_m(2), \ldots, K_m(N/2)]\right\},$$
$$\Im\left\{[K_m(0), K_m(1), K_m(2), \ldots, K_m(N/2)]\right\}]^{\mathrm{T}}$$

The PCA transforms $\mathbf{K}$ into a new matrix

$$\mathbf{H} = \mathbf{Q}_{\mathrm{k}}^{\mathrm{T}}\mathbf{K} \tag{2.16}$$

where the columns of $\mathbf{Q}_{\mathrm{k}}$ are the eigenvectors of $\mathbf{K}^{\mathrm{T}}\mathbf{K}$ which again have been weighted with the corresponding spherical surface segment. The weighted eigenvalues λ_l are proportional to the variance that is explained by the principal component.

Fig. 2.8 shows the amount of explained variations for the time domain and the frequency domain PCA. There is hardly any difference between both approaches but it can be seen that the amount of explained variation increases with the number of principal components.

The impulse response $\mathbf{h}_1$ (i.e. the vector of $\mathbf{H}$ corresponding to the largest eigenvalue) already models 89% of all required impulse responses with a scalable gain only, and four impulse responses would already model over 95%. The required transfer-functions $K_m(j\omega)$ are thus all very similar.

The resulting impulse responses $\mathbf{h}_l$ could directly be implemented as FIR filters in the ALC. However, an IIR representation of the impulse responses is beneficial in order to reduce computational burden of the prototype filters.
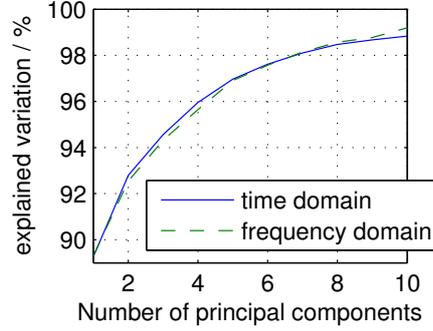
Figure 2.8: Explained variation of the PCA functions.

2.3.2 IIR Filter Design

The outputs of the parallel filters are superposed with the real valued weights w_l to yield an approximation $\tilde{K}(j\omega)$ of the actual plant $K(j\omega) = \frac{P(j\omega)}{S(j\omega)}$

$$\tilde{K}_m(j\omega) = w_1(m)H_1(j\omega) + w_2(m)H_2(j\omega)\ldots + w_L(m)H_L(j\omega). \qquad (2.17)$$

If the filters $H_l(j\omega)$ shall be modelled as IIR filters, the plant approximation $\tilde{K}(j\omega)$ reads as

$$\tilde{K}_m(j\omega) = w_1(m)\frac{B_1(j\omega)}{A_1(j\omega)} + w_2(m)\frac{B_2(j\omega)}{A_2(j\omega)}\ldots + w_L(m)\frac{B_L(j\omega)}{A_L(j\omega)}. \qquad (2.18)$$

where the numerator and denominator frequency responses can be written as polynomials of order $\mathcal{O}$ in dependency of the z-transform variable

$$\frac{B_l(z)}{A(z)} = \frac{b_{0,l} + b_{1,l}z^{-1}\ldots + b_{\mathcal{O},l}z^{-\mathcal{O}}}{1 + a_{1,l}z^{-1}\ldots + a_{\mathcal{O},l}z^{-\mathcal{O}}} \qquad (2.19)$$

The weights $w_l(m)$ as well as the coefficients $a_{\mathcal{O},l}$ and $b_{\mathcal{O},l}$ (with $\mathcal{O} = 0\ldots\mathcal{O}$) have to be chosen in order to minimize the squared residual ANC error $e^2(k,m)$ (i.e. the modelling error of $\tilde{K}$) over all $k = 1\ldots N$ frequency

39

bins and all $m = 1 \ldots M$ measurements:

$$\min \sum_k \sum_m e^2 = \min \sum_k \sum_m \left| K(k)_m - \tilde{K}_m(k, b_{\mathcal{O},l}, a_{\mathcal{O},l}, w_{m,l}) \right|^2. \qquad (2.20)$$

Due to the denominator coefficients $a_{\mathcal{O},l}$, this minimization can only be solved by nonlinear optimization, e.g. by the Gauss-Newton algorithm [Dennis and Schnabel, 1996].

Gauss-Newton Algorithm

The basic Newton optimization algorithm finds the minimum of a function $f(\theta)$ in an iterative procedure and has the following update rule for multidimensional problems:

$$\boldsymbol{\theta}(n+1) = \boldsymbol{\theta} - \boldsymbol{\mathcal{H}}^{-1}\boldsymbol{\nabla}f(\boldsymbol{\theta}), \qquad (2.21)$$

where $\boldsymbol{\nabla}$ is the gradient and $\boldsymbol{\mathcal{H}}$ is the Hessian matrix of second order derivatives.

The Gauss-Newton algorithm neglects the second order derivatives and can only be applied to squared and summed function values $f^2(\boldsymbol{\theta})$. The Gauss-Newton algorithm approximates the Hessian matrix over the Jacobian matrix $\mathbf{J}$ as $2\mathbf{J}^{\mathrm{H}}\mathbf{J}$, where the elements of $\mathbf{J}$ are $J_{i,j} = \frac{\partial f_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j}$, and the superscript $^{\mathrm{H}}$ denotes complex conjugate transposition. The gradient follows to $\boldsymbol{\nabla} = \mathbf{J}^{\mathrm{H}}f(\boldsymbol{\theta})$.

We have $f = e(k, m)$ and the error is considered for M different measurements and for N frequency bins. The error can be written into a vector $\boldsymbol{e} = [e(1,1), \ldots, e(k,m), \ldots, e(N,M)]^{\mathrm{T}}$ which yields the squared and summed cost function $f^2(\boldsymbol{\theta}) = \boldsymbol{e}^{\mathrm{H}}\boldsymbol{e}$. Our vector $\boldsymbol{\theta}$ consists of the variables $a_{\mathcal{O},l}, b_{\mathcal{O},l}$ and $w_l(m)$ and since we only allow real valued weights and coeffi-

cients, our update of the Gauss-Newton algorithm with step-size parameter ν reads as

$$\boldsymbol{\theta}(n+1) = \boldsymbol{\theta}(n) - \nu \Re\{(\mathbf{J}^{\mathrm{H}}\mathbf{J})^{-1}\}\Re\{\mathbf{J}^{\mathrm{H}}\mathbf{e}\}. \tag{2.22}$$

In order to account for the frequency dependent sensitivity of the human ear, the error is weighted using an A-weighting curve and with a logarithmic frequency spacing. The latter weights the frequency bins k such that every octave band yields equal power. This can be achieved by weighting the signal power with $1/k$. For the unsquared error, the weight follows to $1/\sqrt{k}$. Fig. 2.9 shows the combined A and $1/\sqrt{k}$ weighting.
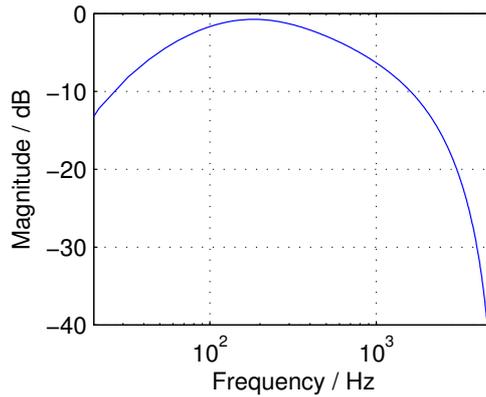
Figure 2.9: A-weighting curve together with an $1/\sqrt{k}$ weighting.

It can be seen that these weights put a lot of emphasis on frequency band around 200 Hz. Since the plant $K(j\omega)$ has low-pass character, the absolute error puts a lot of emphasis on this frequency band, too. Therefore, it is suggested to take the relative error $e_{\mathrm{rel}} = \frac{K(k)_m - \tilde{K}_m(k)}{|K(k)_m|}$ instead. The frequency dependent weights and the direction dependent weights $\cos(\vartheta)$ are subsumed in the weight variable $\psi(k,m)$. Thus, we are looking for the minimum of

$$e^2(k,m) = \left|\psi(k,m)\left(\frac{K_m(k) - \tilde{K}_m(k)}{|K_m(k)|}\right)\right|^2 \tag{2.23}$$

The entries of the Jacobian matrix in the z domain follow to

$$\frac{\partial e(k,m)}{\partial w_l(m)} = \psi(k,m)\frac{-H_l(z)}{K_m(z)} \tag{2.24}$$

$$\frac{\partial e(k,m)}{\partial b_{\mathcal{O},l}} = \psi(k,m)\frac{-w_l(m)z^{-\mathcal{O}}}{A_l(z)K_m(z)} \tag{2.25}$$

$$\frac{\partial e(k,m)}{\partial a_{\mathcal{O},l}} = \psi(k,m)\frac{w_l(m)z^{-\mathcal{O}}H_l(z)}{A_l(z)K_m(z)} \tag{2.26}$$

Our cost function $e^2(k,m)$ is not convex, therefore it is not guaranteed that the Gauss-Newton algorithm finds the global minimum. Thus, it is required to start the minimization with an initial $\theta(0)$ which is close to the global minimum. The PCA outputs are the ideal transfer functions. An IIR model of these transfer functions and the resulting weights $w_l(m)$ are thus a good initialization vector for the Gauss-Newton iterations. The numerator and denominator polynomials of the PCA frequency responses can be determined via the Steiglitz-McBride algorithm [Steiglitz and Mcbride, 1965] by iteratively solving

$$\min \sum_k \left| H_l(k) - \frac{B_l(k)}{A_l(k)} \right|^2. \tag{2.27}$$

The Gauss-Newton algorithm requires a predefined number L of filters and a predefined filter order $\mathcal{O}$. It is apparent that a larger number of filters and a larger filter order will yield better results, but for the sake of low complexity, few filters with low order are desired. Still, it is unclear if e.g. two fourth order filters or four second order filters yield better results. Therefore the Gauss-Newton algorithm was run for various filter numbers and various filter orders. The variation of $K_m(j\omega)$ which is explained by the different combinations is shown in Fig. 2.10a
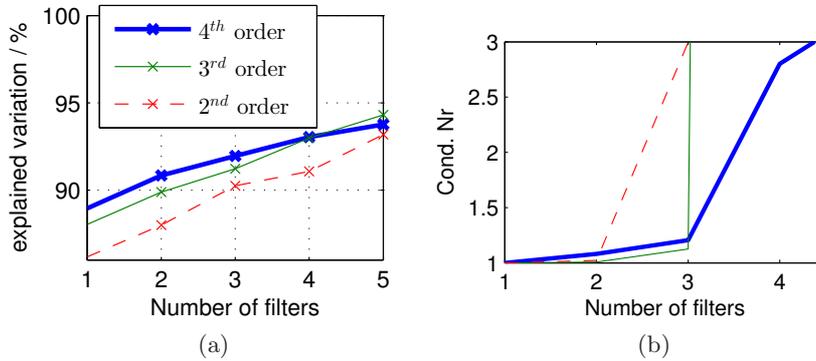
Figure 2.10: (a) Explained variation of the IIR filters in percent. (b) Condition number of $\mathbf{H}^{\mathrm{T}}\mathbf{H}$: Three 3rd order filters yield a condition number of 1.125. Four such filter would already lead to an increased condition number of 78.

The three adaptively combinable 3rd order filters yield better results than two 4th order filters or five 2nd order filters, respectively.

In addition, the resulting three 3rd order filters are almost orthogonal which increases the convergence speed of the LMS adaption [Johns et al., 1989, Erdol and Basbug, 1996, Widrow and Walach, 1984, Narayan et al., 1983]. Fig. 2.10b shows the condition number of $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ as indication of their orthogonality for different orders and different numbers of used filters. Three 3rd order filters yield a low condition, while the 2nd order filters already lead to an increased condition number if more than two filters are used.

The pole-zero plot of the three chosen filters is shown in Fig. 2.11, and the frequency responses are shown in Fig. 2.12.

2.4   Analysis of the ANC Performance

2.4.1   Systems for Comparison

The ALC of Fig. 2.1 with the fixed parallel filters is implemented with a sampling frequency of 11025 Hz and will be compared with the common
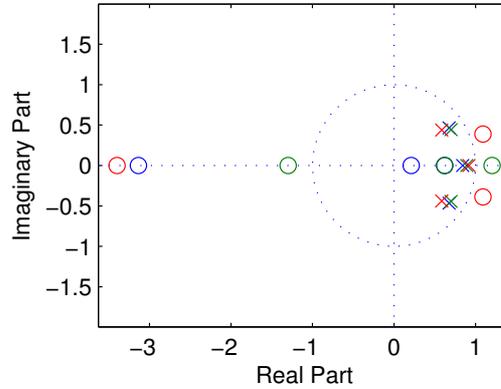
Figure 2.11: Zero and pole locations of the three IIR filters.

| (a) | (b) |
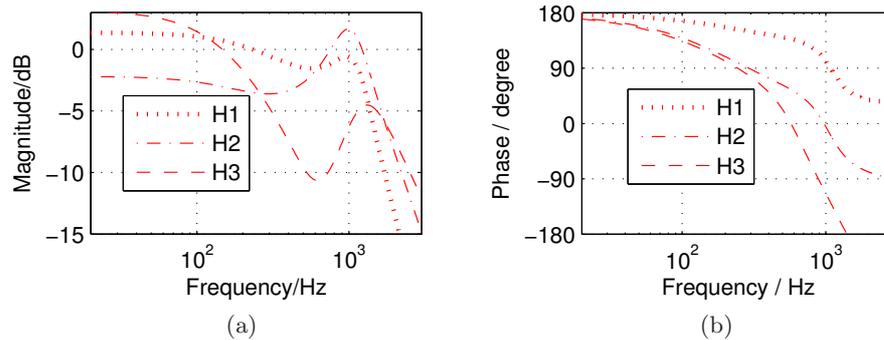
Figure 2.12: (a) Magnitude and (b) phase response of the three chosen IIR filters.

transversal adaptive filter and with a system that has fixed poles but adaptive zeros as in [Bank, 2008, Williamson and Zimmermann, 1996].

The fixed parallel third order filters require three multiply and accumulate operations (MACs) for the numerator polynomial, four MACs for the denominator (including the output weighting by w_l) and three MACs for the weight updates. In total, the three parallel IIR filters thus require 24 MACs. This system is compared with a fixed-pole/adaptive-zero system and with a conventional adaptive filter of similar complexity as compared in Table 2.1.

44

	3 IIRs	8 fixed poles	adapt. FIR
denom.	9	8	-
num.	12	9	12
weight update	3	9	12
total MACs	24	26	24

Table 2.1: Required MAC operations for the compared systems.

The fixed poles for the adaptive-zeros system should be the dominant poles of $\mathbf{K}$. The Steiglitz-McBride algorithm derives a pole- (and/or zero-) model of a single impulse response by iteratively minimizing the output error [Steiglitz and Mcbride, 1965]. The same approach is used to derive eight dominant poles of the data set, but the cost function is defined as the sum over all modelling errors for each variation in $\mathbf{K}$ as in [Williamson and Zimmermann, 1996]. This results in four conjugate complex pole pairs as shown in Fig. 2.13. The frequency responses of the four pole pairs are shown in Fig. 2.14.
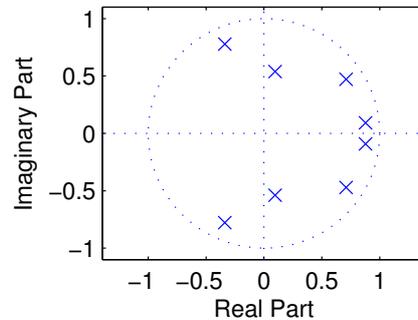
Figure 2.13: Pole locations of the fixed-pole system.

The fxLMS requires a model $\hat{S}$ of the secondary-path to filter the reference input x. $\hat{S}$ should be close to the tight secondary-path S_{tight} since this is the regular use case. S_{tight} has a flat magnitude response and can be modelled with a scaled and delayed impulse.
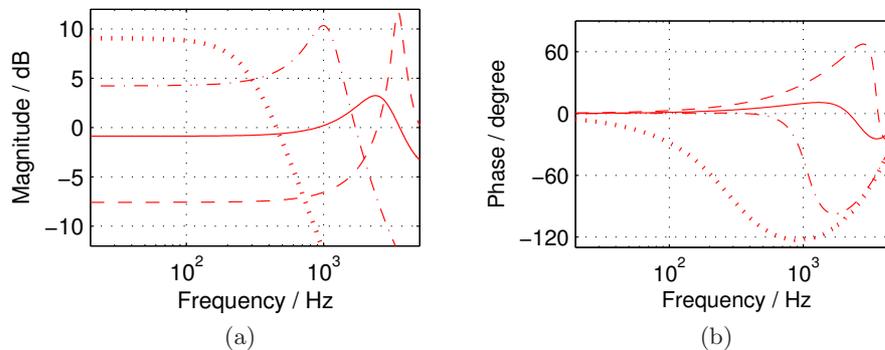
Figure 2.14: (a) Magnitude and (b) phase response of the fixed poles.

Fig. 2.15 shows the impulse response s_{tight} of the tight secondary-path and the scaled impulse $\hat{s}$, and Fig. 2.16 compares the frequency response of $\hat{S}$ with the frequency response of the tight and leaky S, respectively.
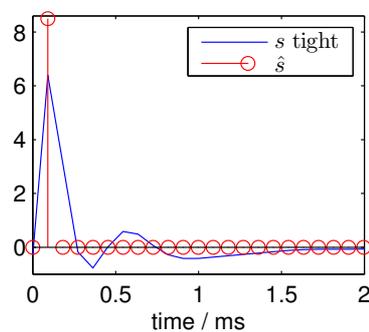
Figure 2.15: Impulse response $s_{\text{tight}}[n]$ and $\hat{s}[n]$ at the sampling frequency of 11025 Hz.

The chosen model $\hat{S}$ can be implemented on a DSP with the cost of only one additional buffer entry and Fig. 2.17 shows that the phase error stays below 90° in almost the complete bandwidth. Only for narrowband excitation below 20 Hz, stability issues would arise that would need to be treated with the leaky fxLMS.

(a)

(b)
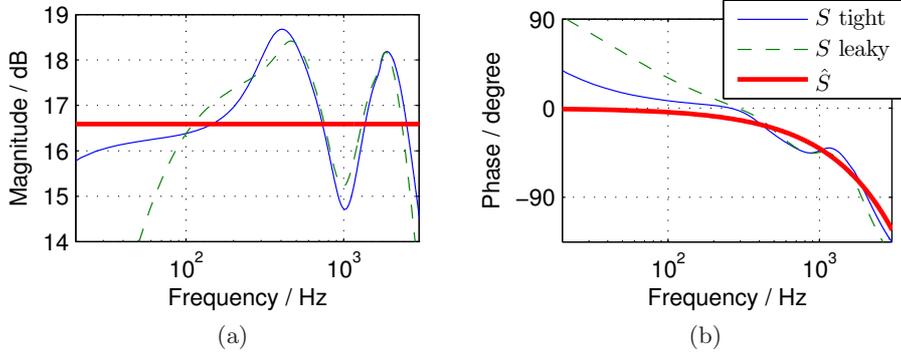
Figure 2.16: (a) Magnitude and (b) phase response of $S(j\omega)$ for tight and leaky positioned headphones together with the model $\hat{S}(j\omega)$
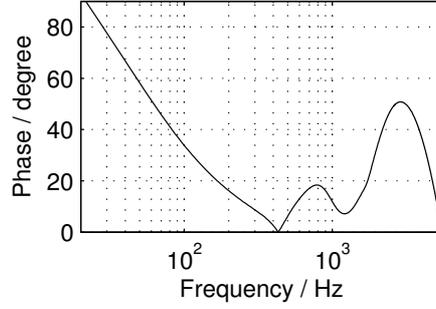
Figure 2.17: Maximum phase error between $S(j\omega)$ and $\hat{S}(j\omega)$

2.4.2 Theoretical Results

To derive the theoretically maximum performance, we take $\mathbf{w}_{\text{opt}}$ for every possible realization of K_l because it allows us to examine the effects of the convergence in the next step.

The noise-cancelling error of our parallel IIR filter can be estimated as the difference between the measured variations $\mathbf{K}$ and the adaptively combined IIRs.

$$\mathbf{E} = \mathbf{K} - \mathbf{H}\mathbf{W}_{\text{opt}}, \qquad (2.28)$$

where $\mathbf{H}$ holds the three IIRs and $\mathbf{W}_{\text{opt}}$ is the matrix of the three optimum adaptive weights times 336 measured variations of $K(j\omega)$. The optimal

47

weights can be derived by minimizing

$$|\mathbf{K} - \mathbf{H}\mathbf{W}_{\mathrm{opt}}|^2 \qquad (2.29)$$

with the least-squares solution

$$\mathbf{W}_{\mathrm{opt}} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{K}. \qquad (2.30)$$

Equally, impulse responses of the fixed poles can be gathered to a matrix $\mathbf{H}_{\mathrm{fxp}}$ as in [Bank, 2008] and the respective zeros can be estimated like the weight matrix $\mathbf{W}_{\mathrm{opt}}$ in eq. (2.30).

The optimal solution of the transversal FIR are the first 12 samples of the impulse responses in $\mathbf{K}$ that can be seen in Fig. 2.18a. At a sampling frequency of 11025 Hz, the 12 samples correspond to 1.1 ms. Fig. 2.18b shows that more than 96% of the impulse energy has passed after 1.1 ms or 12 samples, respectively.
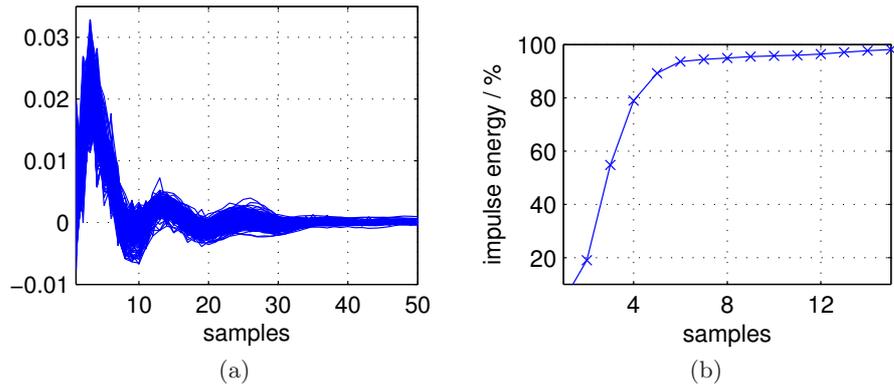
Figure 2.18: Impulses responses $\mathbf{w}_l$ (a) and percentage of the mean passed impulse-energy over time (b) with a sampling frequency of 11025 Hz.

With the above considerations, the noise-reduction of the three systems can readily be estimated. The residual error is set in relation to the pene-

trated noise d and it is express it in dB as

$$e_{\mathrm{dB}}(m) = 10 \log \frac{\sum_n e_m[n]^2}{\sum_n d_m[n]^2}, \qquad (2.31)$$

Fig. 2.19 shows the distribution of e_{dB} over all $m = 336$ variations of $K(j\omega)$. It can be seen that there is hardly any difference in the optimum performance of the three systems.
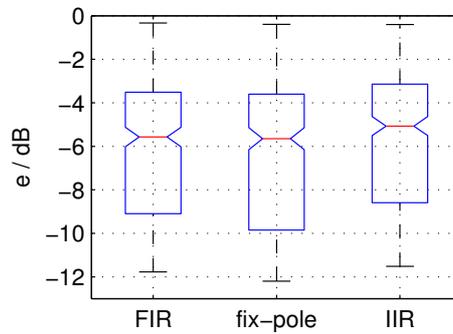
Figure 2.19: Estimated amount of active-noise-cancellation in box plots. The boxes include for 50% of the variations in e_{dB} and the whiskers show the distribution of all 336 variations. The bars in the middle of the boxes denote the median residual error.

The second performance criteria for adaptive noise control is the convergence speed that is bounded by the condition number of $\mathbf{R}$. Table 2.2 compares the condition number of the three systems for white and pink noise.

	3 IIRs	8 fixed poles	adapt. FIR
cond. white noise	4	421	1
cond. pink noise	18	7 900	54
$\lambda_{\max}$ pink noise	189	5 000	723

Table 2.2: Condition number and $\lambda_{\max}$ of the input correlation-matrix $\mathbf{R}$.

For a fully white input, the transversal FIR yields a condition number of 1, while the three 3rd order IIR filters yield a condition number of 4. In the

case of pink noise, which is a more realistic approximation of environmental noises, the three parallel IIRs already yield a smaller condition number than the adaptive FIR, which shows that they are more robust to colourations in the ambient noise.

The fixed-pole/adaptive-zero system yields the largest condition number for both noises. Thus, it is possible that this system only converges very slowly. The actual performance and convergence speed however are evaluated through experiments in the following section.

2.4.3   Experimental Results

In the experiments described in this section, noise was played back from all directions mentioned in section 2.2.1 with tight and leaky sitting headphones. The primary path $P(j\omega)$ and the corresponding $S(j\omega)$ varied every 0.5 seconds and the weights $\mathbf{w}$ were updated by the fxLMS. In the theoretical considerations, we assumed the convergence to $\mathbf{w}_{\text{opt}}$. This time, the imperfect estimate $\hat{S}$ and the fast changes of the plant $K(j\omega)$ are expected to deteriorate the results.

In real-life applications, a predefined μ would be normalized by the input noise energy because it is infeasible to constantly calculate the maximum eigenvalue of $\mathbf{R}$. However, for the experiments, we precalculate $\lambda_{\max}$ in order to compare the experimental with the theoretical results. The experiments were done for pink filtered noise, and the corresponding $\lambda_{\max}$ are given in Table 2.2.

In the first experiment, pink noise was used to evaluate the speed of convergence. Therefore, the weights $\mathbf{w}$ are reset to zero before every change of $P(j\omega)$ and $S(j\omega)$. Fig. 2.20 shows the median e_{dB} over the time. The convergence speed is similar for all three systems. Above all, it can be seen

that the fixed-pole system does not converge as slowly as it could have been expected due to the large condition number. This shows that the condition number of $\mathbf{R}$ only indicates an upper bound for the convergence time, but that the actual convergence time τ lies below this boundary as stated in eq. (2.8).
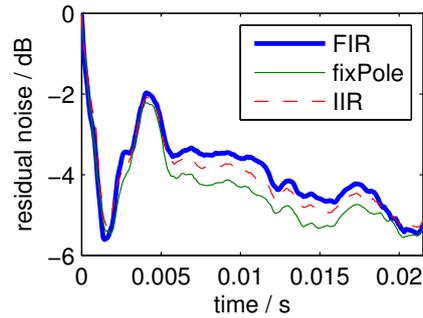
Figure 2.20: Learning curves of the fxLMS adaptation for pink noise: The median residual error converges to the same value for all three compared systems.

In the second experiment, $P(j\omega)$ and $S(j\omega)$ still vary every 0.5 s, but this time the weights $\mathbf{w}$ are not reset to zero. The first few seconds of the simulations are shown in Fig. 2.21. The results over all 336 changes in
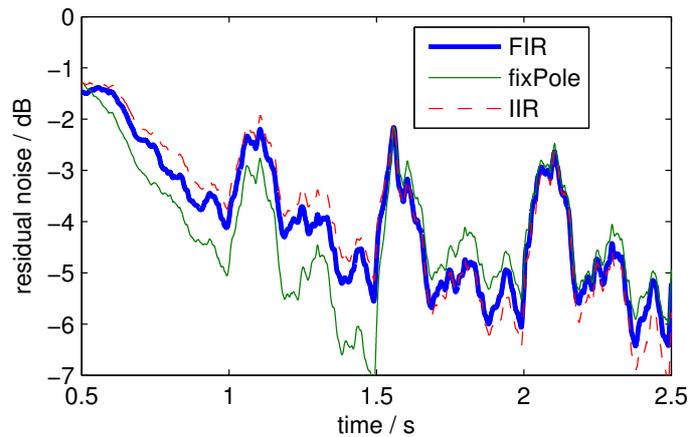
Figure 2.21: Residual error over time.

$P(j\omega)$ and $S(j\omega)$ are shown in the box plot of Fig. 2.22. As expected, the noise reduction is deteriorated compared to the theoretical considerations. However, the three parallel IIR filters show the best performance in the experimental evaluation. The median noise reduction is 1 dB better than for the fixed-pole system and 2 dB better than for the transversal FIR filter.
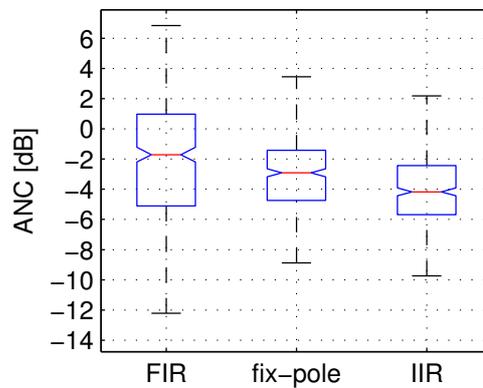
Figure 2.22: Boxplot of the residual errors for pink input noise.

The transversal FIR filter shows good performance in the low frequency band but only with the cost of amplifying the high frequency band. The advantage of the three IIR filters is that they all show low-pass behaviour and consequently they cannot boost the high frequency band as strongly as the FIR filter does. Fig. 2.23 shows that the three IIR filters yield ANC up to 2000 Hz in median, where the other two systems already start to amplify the noise.

At 2500 Hz, the three IIR filters only show a marginal median amplification of 2 dB, while the FIR filter amplifies the noise by 10 dB. The amplification results from the large phase error of $\hat{S}(j\omega)$ in this frequency bin, cf. Fig. 2.17. This phase error can result in wrong fxLMS updates and requires additional measures like reducing the step size or introducing
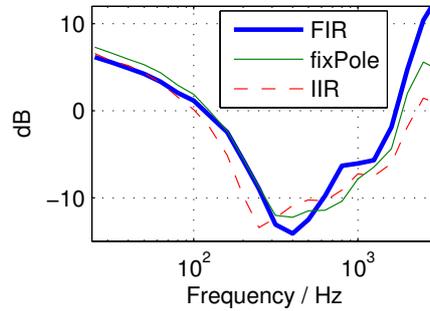
Figure 2.23: Median residual error over frequency for pink noise.

a leakage factor γ as in eq. (2.14). The influence of the leaky fxLMS will be treated in more detail in chapter 5.

The poor performance in the very low frequency band that can be seen in the sensitivity diagrams of Fig. 2.23 results from the large group delay of $S(j\omega)$ and is similar for all three systems.

2.5 Conclusion

In this chapter, different adaptive systems for ANC headphones are investigated. Typically, adaptive transversal FIR filters are used in ANC because they have little stability issues, but it is shown that fixed-pole filters with adapted zeros and full parallel IIR filters with adaptive output gains are equally if not more suitable because the plant variations of the headphones can be accessed via preliminary measurements. The contributions of this chapter are:

- The investigation of the plant variation with consideration of different noise directions and different acoustic properties of the secondary-path.

- The deduction of parallel IIR prototype based on the aforementioned

measurements.

- The comparison of the common adaptive transversal FIR filter with a fixed-pole/adaptive-zero system and the adaptive linear combiner of parallel IIR filters.

Restricted to a similar computation power, the three systems perform equally well. However, it is shown that the parallel IIR filters are more robust in real life conditions although only three adaptive weights are used. The method is applied to ANC headphones where the plant can be described by rather short impulse responses. The benefit of our approach is even more valuable for applications where longer impulse responses have to be equalized. In ANC headphones, the advantage of little adaptation effort becomes especially important if an analogue realization of adaptive ANC headphones is planned because the weights are easily biased by the DC offset of analogue active circuits [Johns et al., 1991, Shoval et al., 1995].

Chapter 3

Secondary-Path Models in Adaptive Feedforward Noise Control Headphones

based on: M. Guldenschuh, 'Secondary-Path Models in
Adaptive-Noise-Control Headphones', 3rd International Conference on
Systems and Control, pp. 653-658, 29.-31- Oct. 2013, Algiers

In the previous chapters, it has been shown that the secondary path of
ANC headphones has a non-negligible phase delay. As a consequence, the
reference signal and the error signal are out of phase which impedes the
correct filter update of the LMS. The solution is filtering the reference signal
with a model of the secondary path, which is then called the filtered-x-LMS
(FxLMS) [Widrow et al., 1981, Burgess, 1981]. Analyses of the FxLMS
[Snyder and Hansen, 1994, Lopes and Piedade, 2004] show that the algorithm
is robust against errors in the secondary path model as long as the phase
error stays below 90°. However, the secondary-path phase of tight sitting

headphones differs from lifted headphones by more than 90° as will be shown in section 3.2. Thus, the FxLMS could diverge, if the headphones are lifted.

There are four main approaches to tackle the problem of changing secondary paths in FxLMS applications, but they all have severe drawbacks and/or limitations.

- An additional analogue feedback controller does not only change the magnitude but also the phase of the sensitivity function. With a preliminary measurement of a worst case $S(j\omega)$ (whose phase differs by more than 90° to the nominal $S(j\omega)$) the feedback controller can be designed to reduce the phase change in the sensitivity function with constraint optimization techniques. The sensitivity function of the additional feedback loop can then be regarded as stabilized secondary path for the fxLMS algorithm [Song et al., 2005]. However, this approach increases the hardware complexity and the controller design is non-trivial.

- Online secondary-path estimation as in [Zhang et al., 2001, Jin and Zhang, 2009, Gan et al., 2005, Akhtar et al., 2006, Liu et al., 2010] aims to track changes in the secondary-path. However these methods fail in large and sudden changes, and/or they inject white noise into the headphones which is counter-productive for a noise-cancelling application.

- Constraints or penalties on the norm of the adaptive-filter prevent the filter from diverging completely [Zhang et al., 2003, Kinney and Callafon, 2009, Cartes et al., 2002, Kamenetsky and Widrow, Nov]. However, they do not avoid wrong filter updates.

- Keeping the norm of the secondary-path model smaller than the ac-

tual secondary path increases the robustness of the FxLMS stability [Ardekani and Abdulla, 2012]. However, manipulating the norm only does not prevent divergence if the phase error exceeds 90°.

We present three solutions that are simple to implement, do not require white noise injection and yield noise-control with robust stability and a correct filter-update by the FxLMS algorithm. All approaches are based on secondary-path measurements under different conditions reaching from very tight to completely loose headphones.

In the first method (solution A), a secondary-path model is designed whose phase does not differ more than 70° from all measured secondary-paths. As an alternative approach to improve the tracking behaviour of the adaptive filter, two models of the secondary-path are implemented; one for the regular tight headphones and one for the lifted headphones. The contribution of the two models to the FxLMS is weighted according to the actual secondary-path, which is either detected via the level-ratio between the error and the reference input (in solution B) or via an infrasound test signal (in solution C). The results are compared with the system of [Zhang et al., 2003], because it combines two of the previously mentioned approaches from literature. The results proof that our solutions are superior to the existing approaches with respect to ANC-headphones.

3.1 Review of the Filtered x Least Mean Square

Fig. 3.1 shows the block diagram of feedforward ANC with the FxLMS. The normalized fxLMS reads as

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu \frac{e(n)\boldsymbol{x}'(n)}{\boldsymbol{x}'^{\mathrm{T}}(n)\boldsymbol{x}'(n)}, \tag{3.1}$$

where $\boldsymbol{x}'(n)$ is a vector of the latest N filtered input samples $[x'(n), x'(n-1), ..., x'(n-N+1)]$ as in Fig. 3.1.
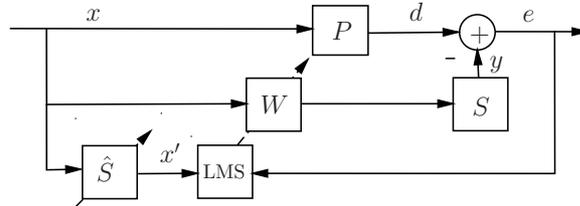
Figure 3.1: Block diagram of feedforward ANC with the FxLMS filter update.

The LMS-input x' is filtered with $\hat{S}$ which introduces the same phase delay to the reference signal as the real secondary-path S to the anti-noise signal. This shall ensure that the error sample $e(n)$ is well-correlated with the reference sample that actually caused this error. Only this time-aligned correlation is a correct gradient for the LMS update.

If $\hat{S}$ deviates from S, the correlation between the filtered reference $\boldsymbol{x}'$ and the error e deviates from the real gradient. For small deviations, the gradient is slightly wrong, but the LMS will still converge. For large deviation, (i.e. for phase differences $> 90°$), the gradient points into a completely wrong direction and the LMS diverges.

3.2 Secondary-Path Measurements

Under laboratory conditions, the secondary-path of adaptive ANC headphones can easily be determined since those headphones already have an internal error-microphone. Thus, S is determined by playing back an appropriate broadband signal (e.g. a swept cosine) via the headphones and measuring the response at the internal microphone. The measurements can be done on test subjects, but also on mannequin heads or on other suitable

artificial ears.

As stated above, small mismatches of the secondary-path model $\hat{S}$ do not matter much in the FxLMS [Kuo and Morgan, 1999]. However, fast and violent changes, e.g. when the headphones are suddenly lifted, can drive the adaptive filter unstable. We therefore extended the secondary-path measurements of our prototype-headphones to four very different conditions:

1. The headphones sit tightly on the ears.

2. A tube with approx. 1 mm radius is inserted between the artificial ear and the cushion of the headphones.

3. Two such leaks are introduced between the artificial ear and the head-phones.

4. The headphones are completely loose.
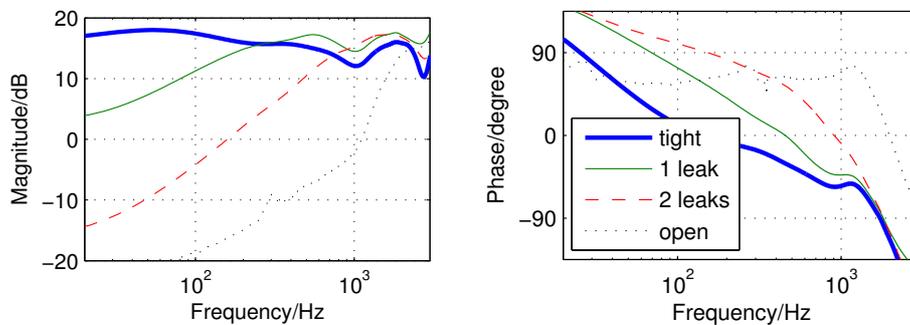
Figure 3.2: Frequency response of the secondary-path for a tight, leaky and completely loose headphone. The increased leakage leads to a magnitude drop-off at low frequencies.

Fig. 3.2 shows the magnitude and phase responses of the measurements. Due to the headphone amplifier the magnitude level lies at 20 dB above the reference voltage, but it can be seen that an increased leakage leads to a magnitude drop-off at low frequencies and to a positive phase shift.

59

The secondary path under tight conditions is assumed to be the regular use case. The phase error from the tight secondary path to the leaky and loose secondary-paths is shown in Fig. 3.3.
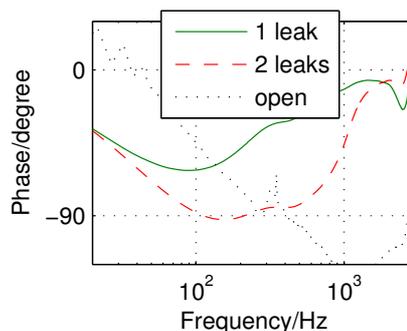
Figure 3.3: Phase error between the tight and the leaky/loose secondary paths.

The error slightly exceeds 90° around 150 Hz for the secondary path with two inserted leaks. For the completely loose secondary path, this phase-error tolerance is violated above 400 Hz. As a consequence, the FxLMS would diverge when the headphones are suddenly lifted.

Approaches with online secondary-path estimation try to track changes in S, but are too slow if S changes abruptly. Therefore, a wrong filter update cannot be avoided, and only a restriction on the growth of W prevents instabilities. In the following, we present three solutions to prevent divergence of the FxLMS without restricting the filter-update.

3.3 Robust Secondary-Path Models

3.3.1 Secondary-Path Model Design from Phase Information

The easiest way to prevent divergence of the FxLMS is to omit the online estimation and to implement a secondary-path model whose phase differs less than 90° from all measured scenarios. This is possible, since the maxi-

mum difference between the phase of the measured secondary-paths is always smaller than 180° within the relevant frequency range as can be seen from Figs. 3.2 and 3.3.

However, it is still favourable if $\hat{S}$ resembles the tight secondary-path for this is the regular use case of ANC-headphones. Therefore, we choose a phase response $\phi_{\hat{S}}(\omega)$ that minimizes the phase deviation to the tight secondary-path $\phi_{S_{\text{tight}}}(\omega)$ in a least square sense and keeps a phase difference of less than 70° to all measured secondary paths. This can be derived by constrained convex-optimization as

$$\min \left| \phi_{\hat{S}} - \phi_{S_{\text{tight}}} \right|_2$$

subject to

$$\phi_{\hat{S}}(k) - \phi_{S_{\max}}(k) < 70°$$

$$\phi_{\hat{S}}(k) - \phi_{S_{\min}}(k) < 70°,$$

where $\phi_{S_{\max}}(k)$ and $\phi_{S_{\min}}(k)$ are the maximum and minimum value of all measured phase responses per frequency bin k.

From the phase response $\phi_{\hat{S}}(k)$, a causal FIR sequence $\hat{s}(n)$ is derived over a fixed-point iteration as in [Hayes et al., 1980]. In this approach, an initial $\hat{s}(n)$ is computed over an inverse Discrete Fourier Transform (DFT) from $\left| \hat{S}_0(k) \right| e^{j\phi_{\hat{S}}(k)}$ where $\left| \hat{S}_0(k) \right|$ is an initial guess of the magnitude of $\hat{S}$, e.g. the magnitude of the tight secondary-path measure. In an iterative process, the second half of the time domain sequence $\hat{s}(n)$ is forced to zeros to assure its causality. Then, the magnitude of the DFT of $\hat{s}(n)$ is taken as new initial guess for $\left| \hat{S}_0(k) \right|$.

The magnitude and phase of the converged $\hat{S}$ is shown in Fig. 3.4, and

the phase error from $\phi_{\hat{S}}$ to the four measured secondary-paths is depicted in Fig. 3.5. As forced in the constrained convex optimization, the phase error stays below 70° in the whole bandwidth.
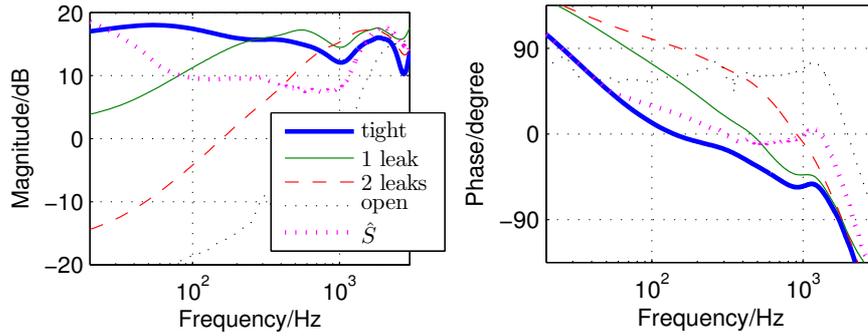
Figure 3.4: Frequency response of the measured secondary-paths and the chosen secondary-path model $\hat{S}$.
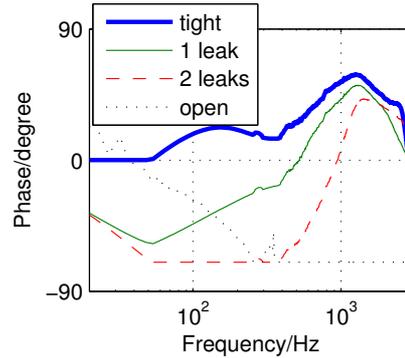
Figure 3.5: Phase error between $\hat{S}$ and the measured secondary-paths. The phase error to all measured scenarios stays below 70°.

$\hat{S}$ is close to the tight secondary-path, but it deviates by 45° at 1000 Hz and by approximately 7 dB between 30 Hz and 1000 Hz. Thus in the beginning of the FxLMS adaption (when the headphones sit regularly tight), the correlation between $\boldsymbol{x}'$ and e will not be as strong as when the secondary path is perfectly modelled. As a consequence, there is less gain on the coefficient update and the FxLMS will be slower in tracking changes in P. A comparison of the temporal behaviour will follow in section 3.4.

3.3.2   Two Secondary-Path Models with Noise-Cancelling Analysis

To improve the tracking performance of the FxLMS, two secondary-path models can be implemented in parallel: $\hat{S}_1$ for the regular tight use-case and $\hat{S}_2$ for the lifted headphones. The latter is derived over a fixed-point iteration (as in section 3.3.1) of the mean phase between the secondary-path measure with 2 inserted leaks and the one with the completely loose headphones. Its frequency response is shown in Fig. 3.6.
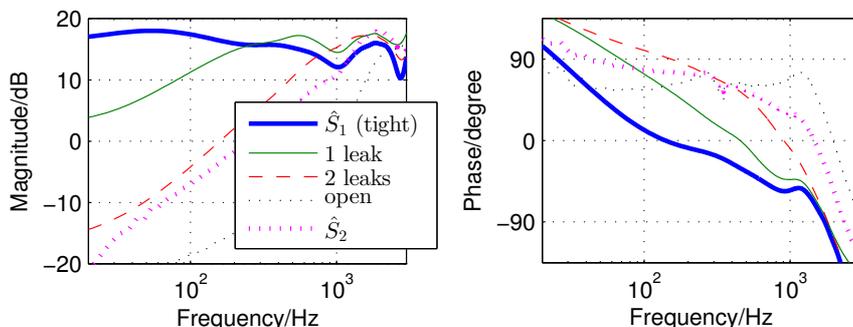
Figure 3.6: Frequency response of the measured secondary-paths and the chosen secondary-path models $\hat{S}_1$ and $\hat{S}_2$.

Depending on the use-case, the contribution of each secondary-path model $\hat{S}_i$ to the FxLMS is weighted with an adaptive gain g_i as in Fig. 3.7. The use-case itself can be identified by comparing the residual-error power P_e with the input power P_x.

Fig. 3.8 shows the relation $\frac{P_e}{P_x}$ for frequencies below 500 Hz. The relation is obtained from a FxLMS-simulation with white noise input, a 2nd order passive-attenuation P (with a cut-off frequency at 500 Hz), a step size of $\mu = 0.002$ and a sampling frequency of $f_s = 44.1$ kHz. The simulation is done for the two extreme cases of the tight headphones and the completely loose headphones, each with the models $\hat{S}_1$ and $\hat{S}_2$ respectively.
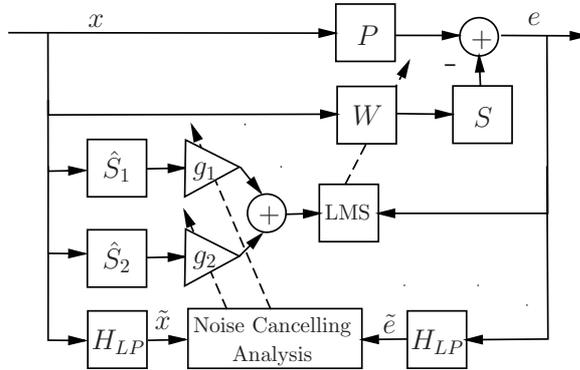
Figure 3.7: Parallel implementation of two secondary-path models. $\hat{S}_1$ is for the regular tight use case and $\hat{S}_2$ is for the lifted headphone.
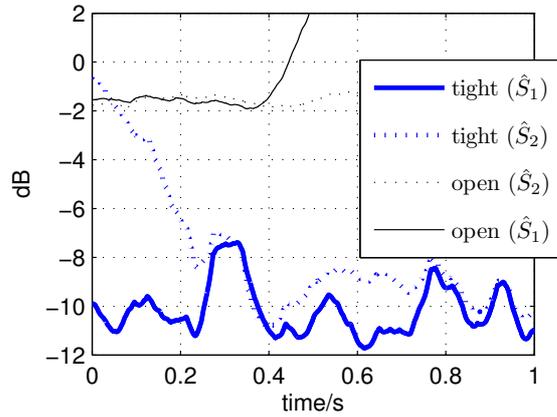
Figure 3.8: Relation between the residual-error power P_e and the input power P_x below 500 Hz. The relation results from FxLMS simulations with different secondary paths and one of the two suggested models which are indicated in the brackets.

In the leaky case, hardly any noise-cancellation is possible, even with the good fitting model $\hat{S}_2$ because the headphones cannot reproduce enough low-frequency power, as shown in Fig. 3.2. With the wrong model $\hat{S}_1$, the FxLMS diverges after 0.4 seconds. This is the worst case; for any other measured S, the FxLMS would diverge slower or not at all. In the tight case, the FxLMS also converges with the wrong model and yields noise-cancellation of 10 dB after 0.3 s. Thus, the current secondary-path can be

detected at latest after 0.3 s.

If the relation $\frac{P_e}{P_x}$ is below -10 dB (which corresponds to the factor 0.1), noise cancellation is very pronounced which means that the headphones sit very tight. Consequently, weight g_1 should be close to 1 and g_2 close to 0. On the other hand, if $\frac{P_e}{P_x} > -2$ dB (factor 0.6), the headphones sit leaky, and we want g_2 to be close to 1 and g_1 close to 0. Hence, a heuristic model for the calculation of the weights g_1 and g_2 yields

$$g_1(n) = 1 - g_2(n)$$
$$g_2(n) = \frac{\frac{P_e}{P_x}(n) - 0.1}{0.6 - 0.1}$$

subject to the restriction of the weights to the interval [0,1]. If the weights should lie outside this interval they are rounded to 0 or 1 respectively.

The relation between P_e and P_x can also be used to detect divergence of the FxLMS when $P_e > P_x$. The power P_x (and accordingly P_e) itself are calculated over a leaky integral as

$$P_x(n) = \alpha P_x(n-1) + (1-\alpha)\tilde{x}^2(n), \tag{3.2}$$

where $\tilde{x}$ is the low-pass filtered input and α is the leakage factor.

The weighted combination of both secondary-path models results in a time-varying filter

$$\hat{S}(n) = g_1(n)\hat{S}_1 + g_2(n)\hat{S}_2 \tag{3.3}$$

Since the phase difference between $\hat{S}_1$ and $\hat{S}_2$ is below 90°, the superposition of the two models yields a magnitude- and phase-interpolation between $\hat{S}_1$ and $\hat{S}_2$.

The filter W is initialized with a low-pass FIR. Therefore, a reasonable

first estimate of the anti-noise is generated from the beginning of the usage. Consequently, also the relation between P_e and P_x is valid from the beginning of the usage. However, the noise cancelling analysis is only valid if there is a low-frequency noise excitation below 300 Hz. This is mostly the case when ANC-headphones are used, e.g. in aeroplanes, in traffic-noise and also in speech-noise, but there might be situations where a user is exposed to high frequency disturbances only. In the following, we present an identification of the secondary-path for ambient noise with little low-frequency content.

3.3.3 Two Secondary-Path Models and Infrasound Secondary-Path Identification

In literature, a complete identification of the secondary-path is done with the injection of white noise [Zhang et al., 2001, Jin and Zhang, 2009, Akhtar et al., 2006, Liu et al., 2010]. However, from Fig. 3.2 it is clear that it is only necessary to identify the low-frequency response to be able to distinguish between the regular tight and the lifted use-case. In fact, it suffices to identify the magnitude at a single low frequency only. We propose to measure the secondary-path with an infra-sound that cannot be heard by humans e.g. with a 18 Hz tone.

The infrasound test-tone is injected into the headphones as in Fig. 3.9. Then the error signal e (which consists of the residual noise and the injected test-tone) is correlated with a delayed version of the original test-tone. The delay corresponds to the group delay of the tight secondary-path at 18 Hz such that the two signals are in phase if $S = S_{\text{tight}}$. Thus, the correlation gain reaches its maximum when the headphones are worn regularly tight.
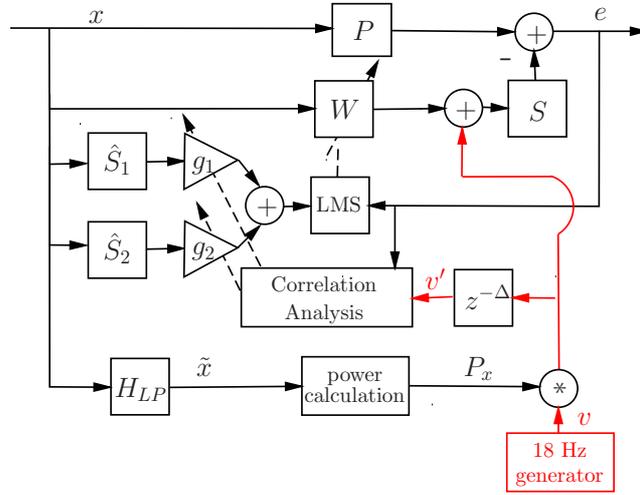
Figure 3.9: A 18 Hz tone detects whether the headphones are regularly tight or lifted and loose.

The correlation gain ξ is calculated over a leaky integral as

$$\tilde{\xi}(n) = \beta\tilde{\xi}(n-1) + (1-\beta)e(n)v'(n)\frac{\gamma}{P_x}$$

where v' is the delayed version of the test-tone, β is the leakage factor, and γ is a constant gain.

The signal-to-noise-ratio (SNR) of the correlation suffers from the residual ambient noise. The SNR can either be increased by the integration time or by the level of v within the error e. We have to set the level of v to 15 dB above the level of the input noise because the test-tone already has a long periodicity and the application requires a fast detection of changes in the secondary path. Therefore the input noise x is filtered with a second order low pass at 50 Hz and the power of the filtered input $\tilde{x}$ (calculated as in eq. (3.2)) schedules the level of v to approximately +15 dB.

This solution is only applied if the low-frequency excitation is below 40 dB$_\text{SPL}$[1]. Thus, the playback-level of the test-tone reaches maximally

---

[1]dB$_\text{SPL}$ has a reference Sound Pressure Level (SPL) of 20μPa.

75 dB$_{\text{SPL}}$ which is not harmful for human ears. Since it is an infrasound frequency, it cannot be heard by the users and it does not deteriorate the FxLMS because the test-tone periodicity is about 38 times longer than our adaptive-filter length.

The correlation-gains for the four measured secondary-paths are shown in Fig. 3.10. They are normalized by the maximum possible correlation-gain which occurs for the tight sitting headphone. In the leaky cases, ξ varies around 0. The filter weights are therefore defined as

$$g_1(n) = \xi_{[0,1]}(n)$$
$$g_2(n) = 1 - g_1(n)$$

where $\xi_{[0,1]}$ is the normalized correlation gain with the negative values set to 0.
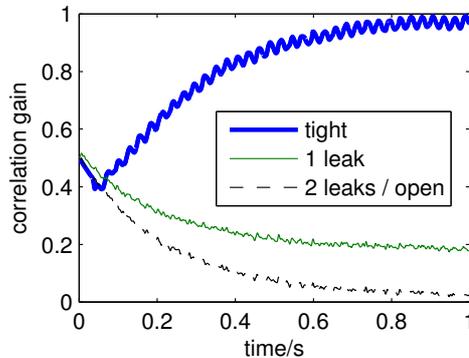
Figure 3.10: Correlation gain over time for the 4 measured secondary-paths. If the headphones sit tight, the 18 Hz tone is played back loud enough and yields a large correlation gain. In the other cases, especially for the very leaky and the loose headphones, the 18 Hz tone is hardly played back. This is why the correlation gain approaches 0. The tight and the leaky cases can easily be distinguished after about 0.2 seconds.

3.4 Results

To evaluate the proposed approaches, simulations were performed. We are interested in the FxLMS behaviour when the phase of the secondary path changes more than 90°. Therefore, the secondary-path is changed every second according to the following sequence: $S_{\text{tight}} \rightarrow S_{\text{2leaks}} \rightarrow S_{\text{tight}} \rightarrow S_{\text{loose}} \rightarrow S_{\text{tight}}$. The passive attenuation P is modelled again by a second-order low-pass at 500 Hz.

First, the noise-cancelling performance is evaluated for white-noise input. Since this input has enough low-frequency components, the noise-cancelling-analysis is used to detect whether the headphones sit regularly tight or not.

The ANC performance is compared for 4 systems:

- An ANC system with a static secondary-path model

- Zhang's method

- Our approach A (whose secondary-path model S_ϕ provokes maximally 70° phase-error)

- Our approach B (with the noise-cancelling analysis)

The results are shown in Fig. 3.11. The static system operates with the tight secondary-path model $\hat{S}_1$. As a consequence it increases the noise when the secondary-path changes to S_{2leaks} and it diverges completely when it changes to S_{loose}. Zhang's method keeps the FxLMS stable, but it shows no noise-cancellation at all for the leaky and loose headphones. Our approach A (with S_ϕ) yields a better result when the headphones are leaky or loose because it still appropriately updates the filter coefficients. However, it performs worse under tight wearing conditions because it does not track changes in the input as fast as Zhang's method due to the secondary-path
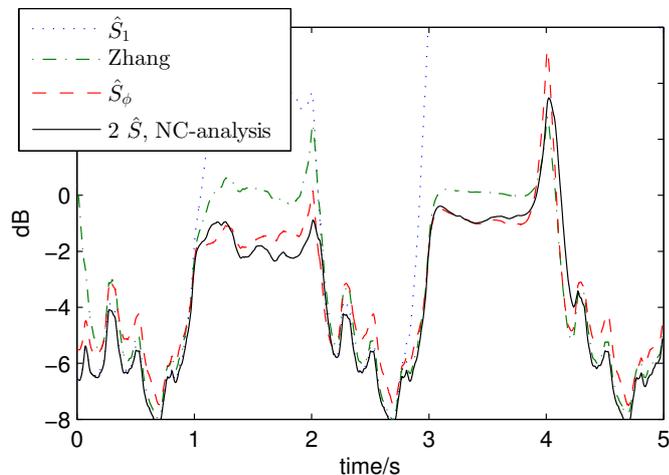
Figure 3.11: ANC-performance for white noise input. Negative dB values denote active noise cancellation whereas positive values mean that the noise inside the headphones is being amplified.

mismatch. Our approach B finally, combines the robust stability for the loose headphones with the good performance under tight wearing conditions.

For the second comparison, the primary noise input consists of white noise that is filtered with a second-order high-pass at 600 Hz cut-off frequency. This input signal lacks low-frequency components and therefore the infrasound test-tone is used to detect changes in S instead of the noise-cancelling analysis.

Fig. 3.12 shows the ANC results. Again, the FxLMS without secondary-path considerations diverges and our approach A keeps the FxLMS stable. For this input signal, Zhang's constraint on the coefficient-update impedes optimal noise-cancellation even when the headphones sit regularly tight. Our approach C, with the infrasound-test tone however, yields optimal ANC again.

The third comparison is done for a narrow-band input. White noise has been filtered with a second-order peak filter at a centre-frequency at 800
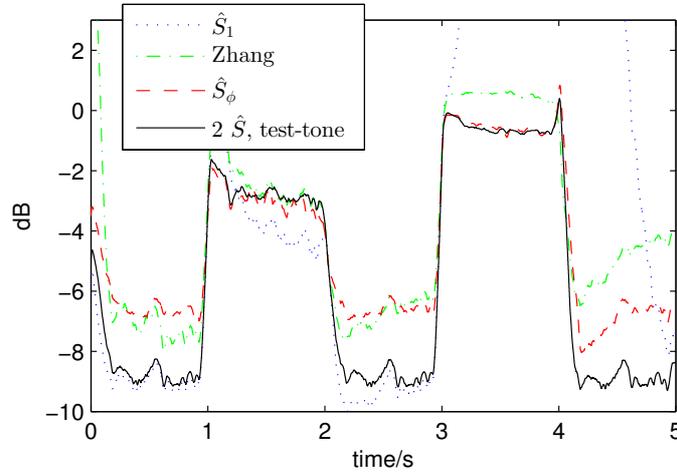
Figure 3.12: ANC-performance for high-pass filtered white noise. This time, the approach with the infrasound-test tone yields the best results. Again, the FxLMS can be kept stable with $\hat{S}_\phi$.

Hz. Again the infrasound test-tone is used since there is no low-frequency excitation in the input signal.
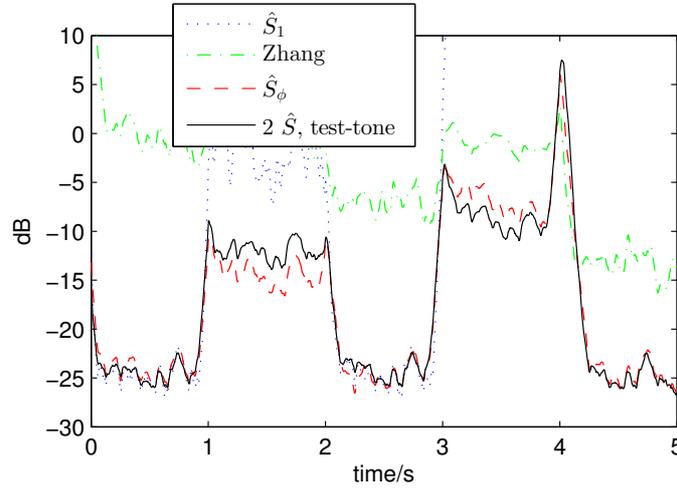
Figure 3.13: ANC-performance for a narrowband input. Here, the (non-adaptive) model $\hat{S}_\phi$ already yields optimal ANC. The approach with the 2 models and the infrasound secondary-path detection does not bring further improvements. Zhang's constraint on the coefficient-update impedes optimal ANC for this narrowband excitation.

Fig. 3.13 shows that approach C does not bring further improvement over approach A this time. For this more deterministic signal, the slower tracking-behaviour due to the imperfect secondary-path model $\hat{S}_\phi$ does not have a negative influence as long as P does not change drastically. However, Zhang's constraint on the coefficient-update is even more problematic for this narrow-band input. It prevents the FxLMS from diverging, but our approaches yield up to 10 dB more ANC.

3.5   Conclusion

The FxLMS is a very efficient algorithm for ANC-headphones that relies on the knowledge of the secondary-path. It has been shown that the secondary-path differs considerably between loose and tight headphones. As a consequence, the FxLMS drives unstable if the headphones are suddenly lifted. Three methods are proposed to overcome this problem. First, a stabilizing secondary-path model can be derived from the measurements of the tight and loose headphones. This secondary-path model does not only prevent the FxLMS from diverging; it also shows optimal performance for narrow-band excitation and almost optimal performance for white noise input. The performance for broadband and high pass excitations is further improved by implementing two secondary-path models in parallel. One for the regular tight use-case and another for the lifted headphones. The models are weighted according to the current use-case. Two different methods demonstrate how the current use-case can be detected.

If the headphones are loose, they cannot produce enough low-frequency power for the anti-noise signal and noise cancellation fails. Thus one approach compares the input power with the residual-error power to detect whether noise has been cancelled or not and deduces whether the head-

phones sit tight or not. If there is not enough low-frequency excitation for a reliable comparison, an infrasound test-tone is played-back to determine the low-frequency response of the secondary-path. The infrasound test-tone cannot be heard by the user, but it suffices to reliably detect whether the headphones sit tight or not. The presented methods are easy to implement and very cost efficient. Additionally, simulations show that these methods are superior to existing approaches in literature.

Chapter 4

Prediction Filter Design for Feedback ANC Headphones

based on: M. Guldenschuh, R. Höldrich, 'Prediction Filter Design for Active Noise Cancellation Headphones', IET Signal Processing, Vol. 7/6, pp. 497-504, 2013

As already mentioned, digital signal processing allows for more flexibility in filter design than analogue technology, but the latency of conventional audio converters severely limits the ANC performance [Snyder and Hansen, 1990]. Chapter 2 showed a way how to maintain the power of adaptive feedforward ANC with a reduced adaptation effort that allows for an analogue implementation. This chapter presents an alternative for a digital feedack ANC systems: Due to the latency of the converters, the adaptive filter has to predict the noise to compensate for the delay. In the feedback ANC approach, the prediction is based on noise that actually entered the headphone. This has two advantages. Firstly, ANC is independent from the direction of incident noise and also works in diffuse sound-fields [Zangi, 1993, Rafaely and

Jones, 2002]. And secondly, the higher frequencies of the entered noise are damped by the ear cup. This low-pass characteristic is advantageous when it comes to signal prediction [Jain and Ranganath, 1981].

In literature, prediction is mostly done by different kinds of the least mean squares (LMS) algorithm [Kuo et al., 2006, Gan et al., 2005, Kannan et al., 2011] or by iterated one-step-ahead predictions [Oppenheim et al., 1994]. Both algorithms are based on sequential updates of the prediction filter and account for changes in the noise-signal characteristics. However since most environmental noises, such as traffic noises, are broadband, the signal characteristic of the penetrating noise mainly depends on the assumed constant damping of the ear cups.

Therefore in this chapter, a prediction filter is designed that is only based on this passive damping characteristic. The proposed filter does not require run-time coefficient updates which makes its application simple and economical, and still it yields better and more robust ANC results than the adaptive methods.

The structure of digital feedback ANC is reviewed in the following section and the prediction filter design is outlined in section 4.2 and 4.3. Section 4.4 shows simulation and measurement results and compares the proposed prediction with LMS and iterated one-step-ahead predictions before the conclusion and an outlook are given in the last section.

4.1 Feedback ANC with Internal Model Controller

Headphones with feedback ANC, as depicted in Fig.4.1, only require one microphone inside each ear cup. This microphone measures the residual error signal $e(t)$ which is the superposition of the entered noise $d(t)$ and the played-back anti noise $y(t)$.
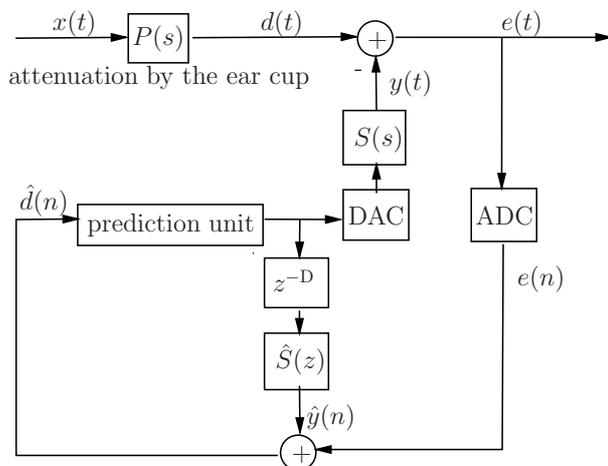
Figure 4.1: Digital feedback ANC: An estimate of the noise inside the ear cup $\hat{d}(n)$ is used as input for the prediction unit. The inverted output is played back to cancel the entered noise $d(t)$.

Due to the analogue/digital converters (ADC and DAC) and the group delay of the secondary path $S(s)$, there is a delay in the feedback loop; a delay which considerably limits the bandwidth of the active noise control. To increase the bandwidth, this delay has to be compensated by predicting future samples of d.

Since d is not directly available, an estimate $\hat{d}$ is obtained by subtracting an estimate $\hat{y}$ of the played-back anti-noise from the sensed residual error, $\hat{d} = e - \hat{y}$. The estimate $\hat{y}$ itself is obtained by feeding the output of the prediction unit through a model $\hat{S}(z)$ of the secondary path and delaying the signal according to the latency of the converters. Because of the internally generated disturbance signal $\hat{d}$, this feedback controller is called an internal model controller (IMC). For a detailed description of IMCs, we refer to [Morari and Zafiriou, 1989].

4.2 Noise Signal Prediction

The prediction unit builds a weighted sum of the available past samples of $\hat{d}$ to predict a future noise sample d',

$$d'(n + \mathrm{D_{DAC}}) = \sum_{i=0}^{L-1} w_i \hat{d}(n - i), \tag{4.1}$$

where L is the prediction order, w_i are the coefficients of the linear prediction filter, and $\mathrm{D_{DAC}}$ is the delay of the DAC in samples that has to be predicted.

Assuming that a correct estimate of d is given, $\hat{d}(n)$ equals $d(n - \mathrm{D_{ADC}})$, where $\mathrm{D_{ADC}}$ is the delay of the ADC in samples. With this assumption, the prediction follows to

$$d'(n + \mathrm{D_{DAC}}) = \sum_{i=0}^{L-1} w_i d(n - \mathrm{D_{ADC}} - i). \tag{4.2}$$

Thus the current predicted sample $d'(n)$ reads as

$$d'(n) = \sum_{i=0}^{L-1} w_i d(n - \mathrm{D_{ADC}} - \mathrm{D_{DAC}} - i), \tag{4.3}$$

and with the total delay $\mathrm{D} = \mathrm{D_{ADC}} + \mathrm{D_{DAC}}$

$$d'(n) = \sum_{i=0}^{L-1} w_i d(n - \mathrm{D} - i). \tag{4.4}$$

This sum of products can be written as vector operation

$$d'(n) = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{d}(n - \mathrm{D}), \tag{4.5}$$

where $\boldsymbol{w}$ is the vector of filter coefficients, and $\boldsymbol{d}(n - \mathrm{D})$ is a signal vector starting from D samples in the past $\boldsymbol{d}(n - \mathrm{D}) = [d(n - \mathrm{D}), d(n - \mathrm{D} -$

$1), \dots, d(n - \mathrm{D} - L + 1)]^{\mathrm{T}}.$

The predicted noise samples are played back by the loudspeakers of the headphones which means that they are convolved with the impulse response of the secondary-path. The convolution operation can be written with the $L \times M$ large convolution matrix $\mathbf{S}$. By using the commutative properties of cascaded linear and time invariant filters, we apply $\mathbf{S}$ directly to the input samples of the predictor and yield the ANC output y

$$y = \boldsymbol{w}^{\mathrm{T}} \mathbf{S} \boldsymbol{d} = \boldsymbol{w}^{\mathrm{T}} \begin{bmatrix} s_1 & 0 & \dots & 0 \\ s_2 & s_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ s_M & s_{M-1} & \ddots & s_1 \\ \vdots & \ddots & \ddots & \vdots \\ s_N & s_{N-1} & \ddots & s_M \\ 0 & s_N & \ddots & s_{M+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & s_N \end{bmatrix} \begin{bmatrix} d(n - \mathrm{D}) \\ d(n - \mathrm{D} - 1) \\ \vdots \\ d(n - \mathrm{D} - M + 1) \end{bmatrix}, \quad (4.6)$$

where s_i are the coefficients of the N taps long impulse response of S. The ANC error follows to

$$e(n) = d(n) - \boldsymbol{w}^{\mathrm{T}} \mathbf{S} \boldsymbol{d}(n - \mathrm{D}), \qquad (4.7)$$

The minimum of the expected squared error leads to the well known Wiener filter [Haykin, 2001], the optimal prediction solution,

$$\boldsymbol{w}_{opt} = (\mathbf{S} \boldsymbol{R}_d \mathbf{S}^{\mathrm{T}})^{-1} \mathbf{S} \boldsymbol{r}, \qquad (4.8)$$

where $\boldsymbol{R}_d$ is the autocorrelation matrix of the latest available noise samples

$$\boldsymbol{R}_d = \begin{bmatrix} r_0 & r_1 & \dots & r_{L-1} \\ r_1 & r_0 & \dots & r_{L-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{L-1} & r_{L-2} & \dots & r_0 \end{bmatrix}$$

and $\boldsymbol{r}$ is a vector of autocorrelation elements starting from lag D

$$\boldsymbol{r} = \begin{bmatrix} r_{\mathrm{D}} \\ r_{\mathrm{D+1}} \\ \vdots \\ r_{\mathrm{D}+L-1} \end{bmatrix}.$$

The matrix inversion in eq. (4.8) can be avoided with the Levinson-Durbin algorithm [Vaidyanathan, 2007], but only if D = 1. The resulting one-step-ahead predictor can still be used for a delay of D samples when the linear prediction of eq. (4.2) is iterated D times. However, since the prediction filter is stable, the recursive one-step-ahead prediction converges to zero. Thus, the filter outputs have to be amplified to get a reasonable multi-sample prediction.

To avoid this problem, gradient search algorithms are often used for prediction problems with more than one sample delay. The normalized filtered x LMS (fxLMS) e.g. use the noise estimate and the error signal to recursively calculate the (multi-sample) prediction filter

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu \frac{e(n)\hat{\mathbf{S}}\hat{\boldsymbol{d}}}{\hat{\boldsymbol{d}}^{\mathrm{T}}\hat{\boldsymbol{d}}}. \tag{4.9}$$

Fig. 4.2 shows the block diagram of the feedback ANC system with an

adaptive filter and the fxLMS update.
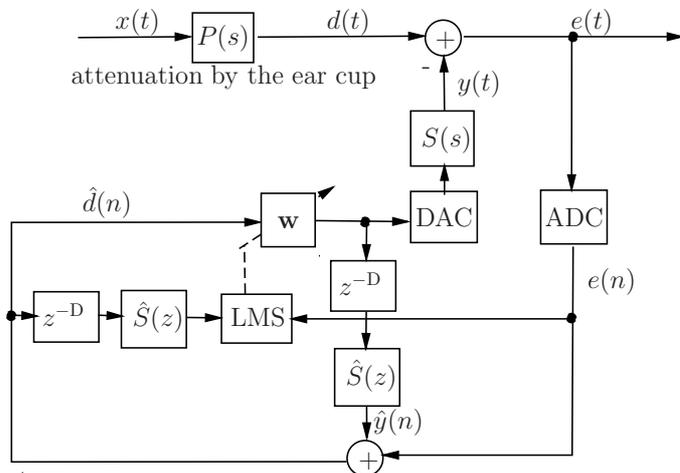
Figure 4.2: Digital feedback ANC with an adaptive filter $\mathbf{w}$ and fxLMS update.

The steady update of the filter is unnecessary if the main noise characteristic always stays the same. And in our ANC application, this is the case because the high frequencies of the outside noise x (as in Fig. 4.1 and 4.2) will always be damped by the physical barrier $P(s)$ of the ear cup. This passive attenuation can be written as a convolution operation

$$\boldsymbol{d} = \mathbf{P}\boldsymbol{x},$$

where $\mathbf{P}$ is a convolution matrix of a low-pass impulse response that simulates the passive attenuation. With this, the autocorrelation matrix $\boldsymbol{R}_d$ follows to

$$\boldsymbol{R}_d = \mathbf{P}^{\mathrm{T}}\boldsymbol{R}_x\mathbf{P}. \tag{4.10}$$

Assuming that x is white noise, its autocorrelation matrix $\boldsymbol{R}_x$ reduces to an identity matrix and $\boldsymbol{R}_d$ solely depends on $\mathbf{P}$. Thus for white noise x, the

calculation of the optimal prediction filter (eq.(4.8)) simplifies to

$$\mathbf{w}_{opt} = (\mathbf{SP^TPS^T})^{-1}\mathbf{Sp}, \tag{4.11}$$

where $\mathbf{p}$ is a vector of autocorrelation elements of the low-pass impulse response starting from time lag D.

In section 4.4, it will be shown that this simplified prediction filter also works for non-white noises, especially if the noises are broadband and their spectral characteristic is still shaped by the passive attenuation P.

The big advantage of eq. (4.11) is that it allows for an a priori filter design where no real-time calculation of the filter coefficients is needed. All required data ($S(\omega)$ and $P(\omega)$) can be measured and designed in advance. Furthermore, we do not need to rely on an autocorrelation matrix $\mathbf{R}_x$ which will never be obtained exactly in real life applications. The prediction filter that follows from this a priori calculation can be used as a fixed prediction filter in the ANC headphones.

4.3   Filter Design for Prototype Headphones

4.3.1   Passive Attenuation of the Headphones

The proposed non-adaptive prediction filter (eq.(4.11)) depends mainly on the passive attenuation of the headphones. The passive attenuation itself however slightly depends on the tightness of the wearing situation and the direction of incident noise. The measurement results of the passive attenuation are already presented in chapter 2. The median passive attenuation over all measurements increases approximately with 6 dB per octave above 500 Hz and 12 dB per octave above 1000 Hz. This corresponds to a simple cascade of two first order filters with cut off frequencies at 500 Hz and 1000

Hz, respectively as can be seen in Fig. 4.3. This cascade could be used for the design of the prediction filter. In order to achieve a more accurate prediction, we propose to design a minimum phase filter from the median values of the passive attenuation. The impulse response of this minimum phase filter is then used for the prediction filter design.
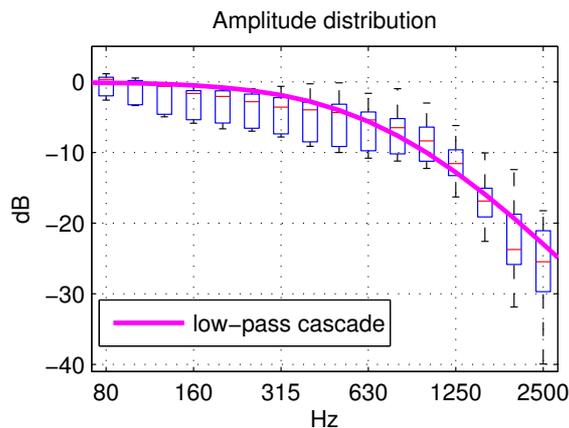
Figure 4.3: Variation of the passive attenuation in third octave bands for different users, different leakage situations, and different directions of incident sound. The boxes include 50% of the measured attenuations and the whiskers show the distribution of all measurements. The bars in the middle of the boxes denote the median values of the measured attenuations. The passive attenuation can be approximated by a cascade of two first order filters.

After these a-priori measurements in laboratory conditions, no further online measurements of P are necessary for the design of the prediction filter. The deviations of the real measured attenuation from the used median value and their influence onto the ANC performance are examined in section 4.4.

4.3.2 Secondary Path

The feedback ANC system of Fig. 4.1 predicts the penetrated noise such that it interferes destructively. However the anti-noise is played back and modified by the secondary path S (i.e. the loudspeaker) of the headphones.

For a perfect noise cancellation, this secondary path would be needed to be compensated. The compensation of S is also indicated in eq.(4.11) by the pseudoinverse of the convolution matrix $(\mathbf{SS^T})^{-1}\mathbf{S}$. However, since the loudspeaker response S involves a delay, its inverse is acausal. A perfect compensation of S is thus impossible. Two steps can nevertheless be undertaken to approximately compensate for the secondary path:

1. The group delay of S can be estimated and this group delay can be considered in the prediction filter design as additional time lag $\hat{D}$ for vector $\mathbf{p}$ in eq.(4.11).

2. The amplitude of S can be compensated with a single gain factor $\frac{1}{g}$.

The prediction filter is thus calculated as

$$\mathbf{w}_{opt} = \frac{1}{g}(\mathbf{P^T P})^{-1}\mathbf{p}, \qquad (4.12)$$

where g is a mean amplitude value of S, and $\mathbf{p}$ is a vector of the low-pass autocorrelation elements starting from time lag $D + \hat{D}$.

Chapter 3 presented measurements of the secondary-path for four provoked and very different leaky wearing situations. This time, we would like to know the average secondary-path under regular real-life usage conditions. Therefore, four persons were asked to put on the headphones normally for measuring the secondary path. We added a measurement on a mannequin which gives us a a total of ten measurements for the left and the right ear. Fig. 4.4 shows the group delay variation of S for these measurements.

The group delay is not constant over the frequencies, but the prediction filter design demands one fix time lag $\hat{D}$. We therefore take the delay of the main pulse in the impulse response of S as fix time lag. This delay is about 100 µs as can be seen in Fig. 4.5.
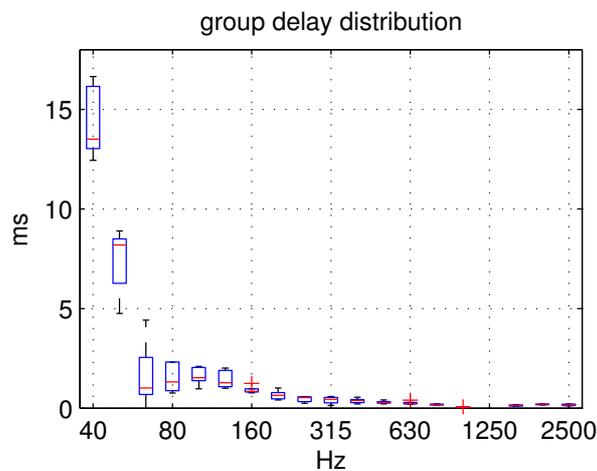
Figure 4.4: Variation of the group delay of S, for 4 users and a dummy head. Again, the boxes include 50% of the measured delays and the whiskers show the distribution over all 10 measurements. The group delay increases largely below 100 Hz.
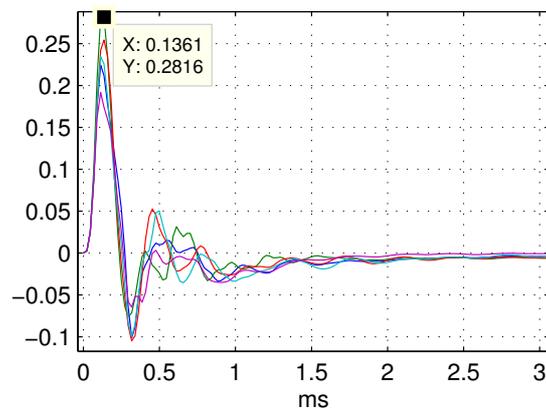
Figure 4.5: Impulse responses of S for 4 users and a dummy head in regular wearing situations. The main pulses appear after about 100 μs.

Fig. 4.6 shows the amplitude variation of the secondary path for the four test subjects and the dummy head. We propose to take the mean value of the amplitudes between 200 and 600 Hz (i.e. 15 dB) as compensation factor g because firstly this is the frequency band where we desire good ANC and secondly the amplitudes in this band are very consistent for all

85
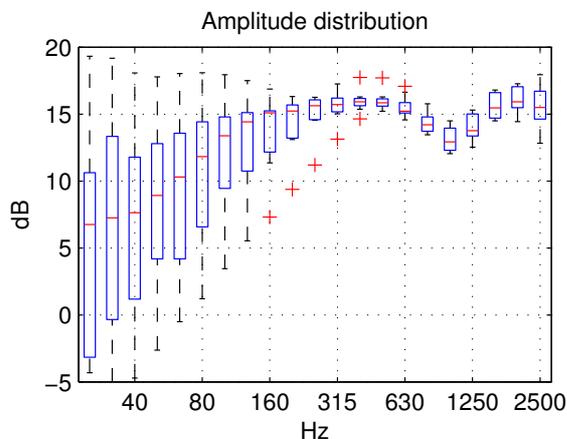
Figure 4.6: Variation of S in third octave bands, for 4 users and a dummy head. The amplitudes are roughly constant and do hardly vary above 100 Hz while there are large differences between the users and a general high-pass characteristic below 100 Hz.

test subjects. On the other hand, this means that frequencies below 100 Hz will not always be cancelled in the best possible way since their amplitude is lower than these 15 dB.

For the prediction filter design, we take the scalar value $g = 15$ dB that we derived from the laboratory measurements. However, an exact model of S is still required for the internal model controller in order to get a correct estimate of the penetrated noise $\hat{d}$. Therefore, it is crucial for the ANC system that a good model $\hat{S}$ is acquired. While the passive attenuation P was only measured once for all different use cases, the impulse response of $\hat{S}$ has to be measured for every user because it may vary from the laboratory measurement [Guldenschuh et al., 2012].

Since the headphones are used for ANC, the secondary path is usually measured in very noisy conditions. We need at least 10 dB signal to noise ratio (SNR) for the measurement because the secondary path S has a dynamic of about 10 dB as can be seen in Fig. 4.6. Therefore,

- we firstly measure a sample of the penetrated noise and apply an A

weighting filter to get an estimate of the dBA level.

- Secondly, we use a measurement signal with roughly the same spectral characteristic as the penetrated noise, i.e. the low-pass characteristic from Fig. 4.3.

- And thirdly, we play back the measurement signal with the dBA level of the noise and increase the SNR to 10 dB by extending or repeating the measurement signal ten times.

$\hat{S}$ is the discrete Fourier transform ($\mathcal{DFT}$) of the measurement response y_S divided by the $\mathcal{DFT}$ of the measurement signal x_S ($\hat{S} = \frac{\mathcal{DFT}(y_S)}{\mathcal{DFT}(x_S)}$).

We propose a sequence of pink noise (which is filtered with the aforementioned low-pass filter) as measurement signal because it is unobtrusive and might even not be noticed within the broadband noise from outside. However any broadband signal like sine sweeps or maximum length sequences (MLS) will suffice this purpose [Weinzierl et al., 2009, Farina, 2007, Rife and Vanderkooy, 1987, Xiang and Schroeder, 2003]. In any case, the unextended measurement signal needs to be at least 50 ms long in order to be able to measure frequencies down to 20 Hz. With the extension/repetition of the measurement signal the whole procedure lasts a little bit longer than one second; after this time, ANC can be started.

In literature, complex methods for online secondary path identification are proposed [Zhang et al., 2003, Akhtar et al., 2006, Gan et al., 2005, Kuo et al., 2006, Liu et al., 2010, Gholami-Boroujeny and Eshghi, 2010]. They track changes in the secondary path during the active noise cancellation, but the more advanced methods are only suitable for feedforward ANC. Also the method of online noise-cancelling analysis that is proposed in chapter 3 is only suitable for feedforward ANC. The method of the injected infra-

sound signal is also possible in feedback ANC headphones and it is presented in [Guldenschuh et al., 2013a], but it is limited to noises with little frequency content below 20 Hz. Another method to track the constitution of the secondary-path is presented in the following chapter. In this chapter, the focus stays on the prediction filter and we only consider the initially measured secondary-path. The consequences of changes in the secondary path onto our ANC system is neverthelessss examined in the following section.

4.4 Simulation and Measurement Results

The following simulations are based on measurement data from our prototype headphones. We have a total latency of approximately 190 µs, including the loudspeaker response of the headphone and the converter latency. This corresponds to $\hat{D}+D = 8$ samples at 48 kHz sampling frequency. If not indicated differently, we use a 0.05 seconds long impulse response for S and a 11 ms long version of the same impulse response for $\hat{S}$.

4.4.1 Investigation on the Spectral Characteristic of the Noise

The 8-samples fixed prediction filter is compared with adaptive methods like LMS and iterated one-step linear prediction. Although the fixed prediction filter is derived under the assumption of white input noise, this comparison is now done for noises with non-white spectral characteristics:

- Static and broadband aeroplane noise.

- Time varying narrow band noise of an accelerating engine.

The filters of the adaptive as well as of the fixed predictor all have a length of 32 samples and the step size of the normalized LMS algorithm is set to 0.003 if not indicated differently. This step size is a factor 6 below the maximum

stable value for the given inputs and it yields a convergence after less than 20 ms as can be seen in Fig. 4.7. For the following comparison, only the steady state error after convergence is considered.
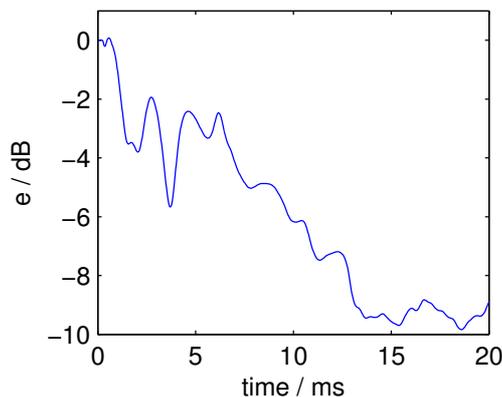
Figure 4.7: Learning curve of the LMS algorithm for the broadband aeroplane noise. The curve shows the power of the residual error related to the power of the noise signal d over time in dB.

Fig. 4.8 shows the spectrum of the aeroplane noise and the spectra of the residual noises after ANC. Our fixed prediction filter yields slightly better results than the iterated one-step prediction filter and - at the same time - consumes far less processing power. The LMS performs slightly better than the two prediction filters between 150 and 300 Hz, and it produces a smaller error above 2000 Hz. However, it therefore only cancels noise up to 600 Hz, which makes it less favourable compared to the fixed prediction filter.

Denote that the LMS filter converges to the H_2 optimal Wiener filter and that the fixed prediction filter can only reach a larger bandwidth because it allows more error in the upper frequency band. In [Rafaely and Elliott, 1999], a constraint on the noise enhancement in the upper frequencies is suggested for ANC in a headrest. In the case of headphones however, the error in the upper frequency bands has a negligible SPL because of the pronounced passive attenuation.
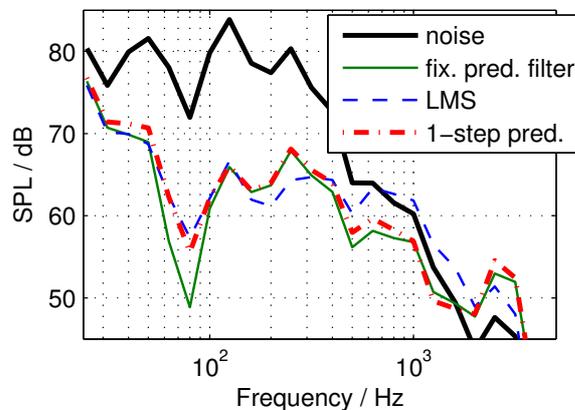
Figure 4.8: Spectra of the aeroplane noise d that entered the headphones and the residual noises after ANC with the proposed prediction filter, an LMS filter and after iterated 1-step ahead prediction. The 1-step and the fixed prediction filters yield ANC up to 1.5 kHz, while the LMS only cancels noise up to 600 Hz. The deteriorated ANC of all methods below 80 Hz results from the long group delay of the loudspeaker in the low frequency band.

The excitation noise of the previous simulation was broadband. Thus the penetrated noise d was above all coloured by the passive attenuation P. We could therefore expect a good result of the proposed fixed prediction filter. In the next simulation, we use the noise of an accelerating engine which is narrow band and time varying.

Fig. 4.9 shows that the LMS algorithm has advantages here, but only if the step size is increased to $\mu = 0.01$. However, this step size is only possible if a very accurate model $\hat{S}$ is implemented. It bears the risk of driving the LMS algorithm unstable for wrong estimates of S as can be seen in the following section. A solution could be an adaptable step size as suggested in [Akhtar et al., 2006]. In either case, the proposed fixed prediction filter yields a much larger ANC bandwidth up to 1.5 kHz again.
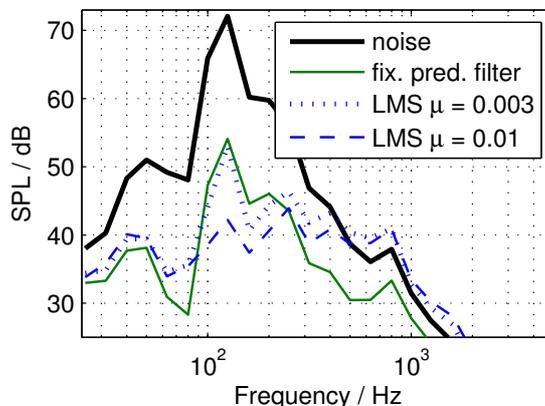
Figure 4.9: Engine noise d that entered the headphone and the residual noises after ANC. Again, the fixed prediction filter yields ANC in a larger band width than the LMS algorithm. However, if the step size μ is increased to 0.01, the LMS performs better at the spectral peak around 120 Hz.

4.4.2   Investigation on Variations in the Secondary Path and the Passive Attenuation

In the previous simulations, we showed that the proposed fixed prediction filter yields the largest ANC bandwidth independently of the spectral characteristic of the noise. Here, we investigate the influence of variations in S and P on the ANC performance. We investigate the influences of these transfer functions separately although in reality they might be coupled. For example, a leaky wearing situation might deteriorate the passive attenuation and, at the same time, weaken the reproduction of low frequencies in the secondary path.

First, we compare the ANC of the proposed prediction filter for two different passive attenuations. The frequency responses of these attenuations are shown in Fig. 4.10.

In the simulation, the secondary path stays unchanged and the excitation signal is pink noise. Although one measured passive attenuation is much
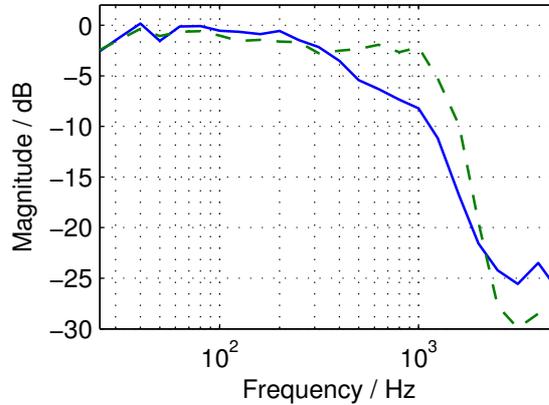
Figure 4.10: Two different frequency responses of the passive attenuation. The solid one was measured under tight wearing conditions. Therefore, the passive attenuation already starts at 400 Hz. The dashed frequency response does hardly attenuate up to 1 kHz. It was measured under a leaky condition.

poorer than the other, the difference of the ANC in third octave bands from 20 Hz to 4000 Hz stays below 0.5 dB. Thus the exact knowledge of P is not necessary for the prediction filter design.

Next, we investigate the influence of a wrong secondary path estimate on the noise prediction. Again, in reality a change in S will not only affect the prediction but the whole feedback ANC system, where a model of S is also needed to estimate the penetrated noise $\hat{d}$. However, we will investigate the influences separately to gain a deeper insight in the ANC system.

Fig. 4.11 shows the frequency response of the used S and $\hat{S}$. While $\hat{S}$ has a very flat frequency response, S has a strong drop off below 200 Hz. On the one hand, this means that the given gain correction $\frac{1}{g}$ will not hold for those frequencies, and we have to expect a poor performance at low frequencies. On the other hand, our prediction filter does not depend on the exact frequency response of S and is thus robust against changes. The fxLMS algorithm however needs a good estimate $\hat{S}$ for its coefficient update as indicated in eq.(4.9).
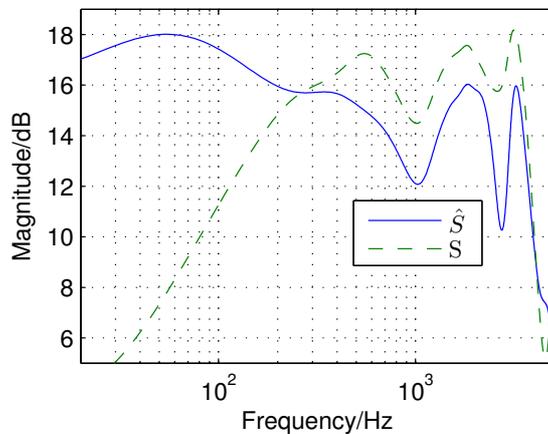
Figure 4.11: Real and estimated secondary path. They differ considerably, especially below 200 Hz.

Fig. 4.12 shows that the fxLMS with a step size of $\mu = 0.003$ diverges, while the fixed prediction filter still yields an ANC of more than 10 dB between 100 and 200 Hz.
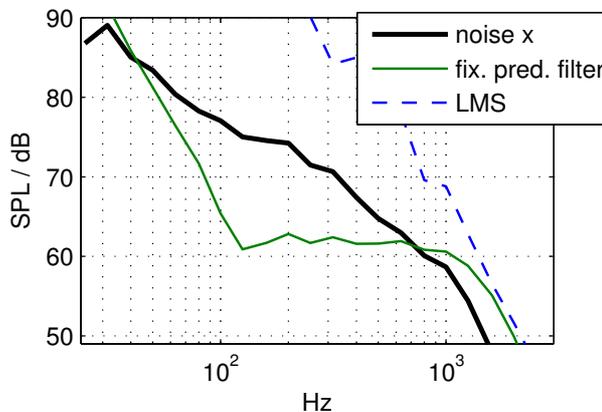
Figure 4.12: Pink noise d and the residual noises after ANC with a wrong estimate of S in the prediction unit. The fxLMS algorithm which depends on an accurate model of S diverges, while the proposed prediction filter still yields a good ANC performance.

Finally, we consider the full consequences when S and P change after the initial identification process. Thus, in this simulation also the estimation

of $\hat{d}$ via the feedback branch through $\hat{S}$ is affected. The feedback ANC system with the fxLMS algorithm runs unstable before one second. With
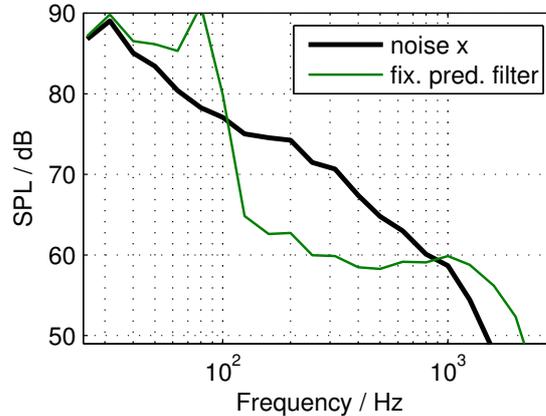
Figure 4.13: Pink noise d and the residual noises after ANC with a wrong estimate of S also in the feedback branch $\hat{S}$ which estimates $\hat{d}$. While the fxLMS runs unstable within 1 second. The proposed prediction filter still yields good results above 100 Hz, but increases the error considerably below 100 Hz.

the transfer functions from Fig. 4.11, the ANC system with the proposed prediction filter still yields a stable feedback, but it produces an error of more than 10 dB around 80 Hz as can be seen in Fig. 4.13. In general, it has to be said that - without further constraints - stability cannot be guaranteed for all deviations from the initially measured secondary path. However, a robust IMC design that accounts for uncertainties in the secondary path model can be found in [Rafaely and Elliott, 1999] and [Kinney et al., 2008].

4.4.3 Measurement Result

Finally, we implemented the feedback ANC system on a digital signal processor (DSP) in order to compare the proposed fixed prediction filter with the LMS algorithm under real conditions. In the implementation, we use a 6 ms impulse response of $\hat{S}$ instead of the 11 ms impulse response. The shorter

impulse response saves computational power, especially for the fxLMS which needs two convolutions with $\hat{S}$: One for the update of its coefficients and a second one to get the estimate $\hat{d}$ of the penetrated noise (which is generally needed in the feedback ANC with an internal model controller.)

As excitation signal, we used pink noise. Fig. 4.14 shows the results of the measurement.
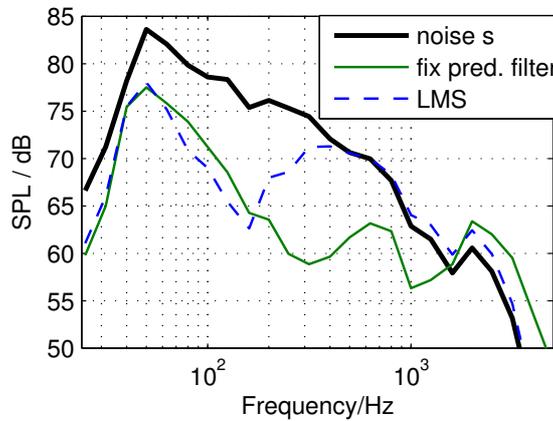
Figure 4.14: Pink noise d and the residual noises after ANC. The proposed prediction filter yields a much better ANC than the fxLMS algorithm between 200 and 1500 Hz with the cost of producing more error above 1500 Hz. Below 50 Hz, the loudspeakers do not produce enough SPL and additionally, the large group delay of S in these frequencies disables a better ANC.

The proposed prediction filter yields ANC up to over 1000 Hz while the fxLMS algorithm only cancels noise up to 400 Hz. As already shown in the simulations, the fxLMS algorithm is much more sensitive to uncertainties in the secondary-path estimate. Also, the adaptive algorithms are more sensitive to sensor noise, which is why the use of a Kalman filter is proposed in [Oppenheim et al., 1994]. Our prediction filter is very robust against sensor noise and inaccurate estimates of S and, at the same time, much more economical and efficient than the adaptive methods. It only produces a larger error in the high frequency band, which can be accepted since the

passive attenuation at these frequencies is already very good.

4.5   Conclusion and Outlook

This chapter presented a new predictive solution for one of the main prob-
lems in digital feedback ANC headphones: The unavoidable delay of the
secondary path and the AD/DA conversion. It has been shown that a main
part of the noise characteristic is determined by the passive attenuation due
to the ear cups. With this information, we design an a priori prediction
filter that does not need any real-time coefficient updates and is therefore
very economical. To the knowledge of the author, this elegant prediction
solution, elaborated by [Jain and Ranganath, 1981, Jain & Ranganath],
has never been suggested for feedback ANC applications. Simulations show
that this filter reaches equal - if not better - results than adaptive predic-
tion methods like the LMS or an iterated one-step-ahead linear prediction.
At the same time, it is more stable against changes in the acoustics of the
headphone. For a prototype system, we could show that the feedback ANC
system can be kept very simple without any adaptation, but for a com-
mercial product some kind of online secondary path estimation or stability
check will be necessary. The proposed prediction filter might also be used
in hybrid ANC systems as proposed in [Song et al., 2005, Rafaely and Jones,
2002, Schumacher et al., 2011], where it is expected to lead to improved ANC
results, too.

Chapter 5

Detection of Secondary-Path Irregularities in Feedback Active Noise Control Headphones

The last chapter showed a practical design procedure for digital feedback ANC. It has been shown that the proposed fixed prediction filter is more robust against changes in the secondary-path than other methods, but it does not contain a detailed treatment of stability. This chapter puts more emphasis on robust stability and proposes a method to detect changes in the secondary-path. Although it has been shown that the adaptive predictor

97

has hardly any advantages over the fixed prediction filter, this chapter uses an adaptive predictor because it allows monitoring the secondary path.

Fig. 5.1 shows a typical realization of a digital feedback ANC-system (compare [Kuo and Morgan, 1999]). It uses a model $\hat{S}(j\omega)$ of the secondary-path (including ADC and DAC) to estimate the primary noise $d(n)$. This internal model controller (IMC) allows designing the filter $W(j\omega)$ as feed-forward controller [Morari and Zafiriou, 1989] and therefore as Least-Mean-Square (LMS) adaptive filter. An adaptive filter is beneficial when facing noises with changing spectral characteristics, and the LMS is a cost efficient adaptive algorithm [Haykin, 2001].
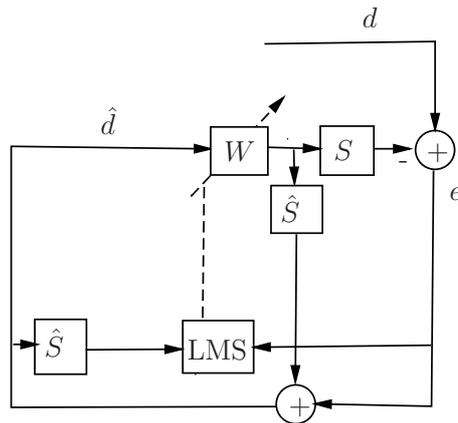
Figure 5.1: Feedback ANC with an internal model controller: The internal model $\hat{S}(j\omega)$ of the secondary-path is used to derive an estimate $\hat{d}(n)$ of the noise. Besides, it is required to filter the input of the LMS adaptation. The AD- and DA converters are subsumed in S and $\hat{S}$, respectively.

The performance and stability of an IMC feedback ANC-system depends on the accuracy of the secondary-path model $\hat{S}(j\omega)$. An initial nominal model can easily be determined off-line by injecting an appropriate broad-band signal (e.g. a swept cosine) into the headphones and measuring the system response with the microphones inside the ear-cups. The secondary-path $S(j\omega)$ however changes considerably once the headphones are lifted or

pulled away completely as is shown in chapter 3. The deviation of $S(j\omega)$ from the nominal model $\hat{S}(j\omega)$ can then drive the system unstable.

In [Song et al., 2005], it is shown that an additional analogue feedback-controller can reduce the influence of changing $S(j\omega)$ onto the error signal, but the analogue controller design is non-trivial and stability cannot be assured in general. Consequently, the controller has either to incorporate an uncertainty about the secondary-path model [Rafaely and Elliott, 1999, Rafaely and Elliott, 2000, Kinney et al., 2008] or an on-line secondary-path estimation has to be implemented that tracks changes in $S(j\omega)$ [Gan et al., 2005, Akhtar et al., 2006, Zhang et al., 2003]. The first method suffers from a loss of performance under optimal conditions and/or requires real-time Fourier transforms to check the uncertainty constraints; and the latter method fails when there are large and sudden changes in the secondary-path, and it requires the injection of a broadband noise into the headphones which is counter-productive for a noise-cancelling application. Chapter 3 used an infrasound signal to detect changes in S in feedforward ANC headphones. In [Guldenschuh et al., 2013a], it has been shown that the infrasound measurement can be used for feedback ANC, too, but only if there is little low frequency excitation in the ambient noise.

In this chapter, a simple and efficient method is introduced to identify changes in the secondary-path without the need of injecting additional noise into the headphones and without the need of real-time Fourier transforms. In particular, it is shown that lifting and pulling away the headphones mainly affects the low frequency region of $S(j\omega)$ and that the adaptive filter $W(j\omega)$ which tries to invert $S(j\omega)$ can be used to detect those low frequency changes. Once this irregularity in $S(j\omega)$ is detected, the LMS adaptation is interrupted and the filter temporarily changes to a stable default-setting.

This way, instabilities are avoided even during sudden and large changes in $S(j\omega)$.

In the given adaptive feedback-ANC-system, two stability issues arise: (i) Stability of adaptation and (ii) stability of the feedback loop. First, the theoretical conditions for both issues are discussed separately in section 5.1 and 5.2. Then experimental data from prototype headphones are presented in section 5.3. With the experimental data, the adaptation stability is reviewed in section 5.4 and the feedback stability in section 5.5. Finally, experimental results of our algorithm are compared with existing approaches from literature in section 5.6.

5.1   Theoretical Considerations for the Stability of Adaptation

The sensitivity function $T(j\omega)$ (i.e. the transfer function from the input noise $D(j\omega)$ to the residual error $E(j\omega)$) of the feedback system in Fig. 5.1 reads as (with omitted dependency on $j\omega$)

$$T = \frac{E}{D} = \frac{1}{1 + S\frac{W}{1 - \hat{S}W}} = \frac{1 - W\hat{S}}{1 - W(\hat{S} - S)}. \tag{5.1}$$

In the case of $S = \hat{S}$, the denominator of T is equal to unity and the filter W becomes $W = \hat{S}^{-1} = S^{-1}$ in order to minimize $|T|$. However, the inverse of S in general will not exist since S will not have minimum phase. Thus, the filter W can only try to compensate for the phase delay and the dynamics of S e.g. in an H_2 or H_∞ optimal sense. The accuracy of the compensation depends on the bandwidth in which S shall be compensated. It is easier to compensate for the phase delay and the magnitude at a single frequency than

over a broad bandwidth. Thus, the optimal filter W depends on the current spectral characteristic of the input noise d. It is therefore advantageous to implement an adaptive filter W that yields the compensation in the band where it is currently needed. Note that the adaptive filter tries to do a system identification of the inverse secondary-path. This fact will be used later to detect changes in the secondary-path.

Stability of LMS Adaptation

The convergence properties of the LMS have already been treated in chapter 2. For this chapter, it is only important to recall that phase errors between S and $\hat{S}$ larger than 90° can lead to the divergence of the adaptive filter; and that only a penalty on the filter gain can keep the FxLMS stable in these cases. The penalty is introduce in the cost function of the LMS as $J(\mathbf{w}) = E\{e^2 + \gamma(\mathbf{w}^T\mathbf{w})\}$ and it leads to the leaky FxLMS algorithm whose normalized version reads as

$$\boldsymbol{w}(n+1) = (1 - \mu\gamma)\boldsymbol{w}(n) + \mu\frac{e(n)\hat{\mathbf{S}}\hat{\boldsymbol{d}}(n)}{\hat{\boldsymbol{d}}^{\mathrm{T}}(n)\hat{\boldsymbol{d}}(n)}, \qquad (5.2)$$

where $\hat{\mathbf{S}}$ is the convolution matrix of the secondary-path model and $\hat{\boldsymbol{d}}(n)$ is a vector of the latest L estimated noise input-samples as in Fig. 5.1. The leaky LMS prevents divergence and avoids large filter gains in general which can be important in real life conditions [Kuo and Morgan, 1999] as will also be shown in section 5.4.2.

5.2 Theoretical Considerations for Feedback Stability

5.2.1 Constraint on the Real Part of the Open Loop

The stability of the feedback loop depends on the poles of the sensitivity function T in eq. (5.1) that all have to lie within the unit circle in the z-plane. Since we can assume stable plants, stability is ensured if and only if the open loop transfer function $L = \frac{WS}{1-W\hat{S}}$ does not encircle the point (-1,0) in the z-plane [Morari and Zafiriou, 1989]. Visually, it is easy to check if the Nyquist contour of L encircles the point (-1,0), but analytically and even numerically it is less so. One sufficient, yet not necessary, condition would be to force L to lie on the right hand side of (-1,0). This can be checked analytically, because it means that the real part of L is constrained to

$$\Re\left\{\frac{WS_i}{1 - W\hat{S}}\right\} > -1. \tag{5.3}$$

Without real-time information about the secondary-path, every possible variation $S_i(j\omega)$ has to be tested in condition (5.3). This makes the condition only suitable if it is known that the secondary-path can only vary between a few distinct realizations. Even for one realization of $S(j\omega)$, condition (5.3) is computationally quite demanding. The more variations in $S(j\omega)$ that are possible, the more processor intensive becomes checking the condition.

5.2.2 Constraint on the Norm of WU

A less processor intensive constraint results from regarding the deviation of the nominal model from the real secondary-path as additive uncertainty $U(\omega) = |\hat{S}(j\omega) - S_i(j\omega)|$. We are now interested in the maximum uncertainty

$U_{\mathrm{max}}(\omega)$ over all possible variations $S_i(j\omega)$. $U_{\mathrm{max}}(\omega)$ can be regarded as radius around every frequency point ω of the Nyquist contour of $\hat{S}(j\omega)$. Thus, all $S_i(j\omega)$ have to lie within this band of radii. The uncertainty radius around the frequency points of the open loop follows to

$$r(j\omega) = \left| \frac{W(j\omega)U_{\mathrm{max}}(j\omega)}{1 - W(j\omega)\hat{S}(j\omega)} \right|. \tag{5.4}$$
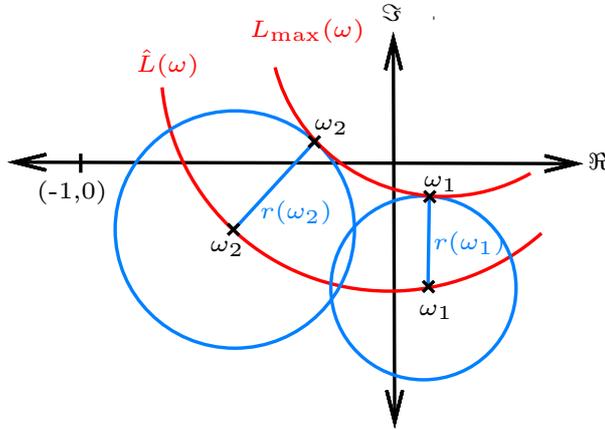
Figure 5.2: Graphical derivation of the stability condition: The nominal open-loop $\hat{L}$ is displayed together with L_{max} the open loop with the largest uncertainty. No open-loop with a smaller uncertainty can encircle the point (-1,0) if the distance of $\hat{L}(j\omega)$ from (-1,0) is larger than $r(j\omega)$.

From Fig. 5.2 it is clear that no open loop (for any S_i) can encircle the point (-1,0) if the distance of the nominal open loop $\hat{L}$ from (-1,0) is larger than r. Mathematically expressed, this means

$$\left| 1 + \frac{W\hat{S}}{1 - W\hat{S}} \right| > \left| \frac{WU_{\mathrm{max}}}{1 - W\hat{S}} \right|. \tag{5.5}$$

The condition can be simplified to

$$1 > |WU_{\mathrm{max}}| \tag{5.6}$$

which yields the practical constraint for IMC [Morari and Zafiriou, 1989]

$$|W(j\omega)| < \frac{1}{U_{\max}(\omega)}. \tag{5.7}$$

This would be a necessary and sufficient condition if the circles of uncertainty would be fully covered by different realizations of S. However, it is a rather conservative constraint if there are only a few distinct secondary-path realizations in the band of uncertainty.

In each case, it only requires the frequency domain representation of $W(j\omega)$ and is far less computationally demanding than constraint (5.3). Still, it is advisable to implement the frequency-domain LMS. The resulting filter $W(j\omega)$ can be transformed into the time-domain to avoid the latency of block-processing in the anti-noise path as in [Morgan and Thi, 1995, Park et al., 2001]. In the following, we show a constraint that allows implementing the FxLMS completely in the time-domain.

5.2.3 Constraint on Single Frequency Bins of $|WU|$

The poles of a system are its resonances and they can be recognized by peaks in the magnitude response. The closer a pole (or a complex conjugate pole-pair, respectively) approaches the unit circle, the sharper the resonance becomes until it can be heard as ringing. As long as the pole stays inside the unit circle, the ringing will eventually die out, but as soon as the pole crosses the unit circle, the system looses its stability.

It is thus very likely that the instability is caused by a distinct pole or pole-pair at a distinct frequency. If the resonance frequencies that most likely turn into unstable poles are known, it can be sufficient to check constraint (5.7) only at the corresponding frequency bins.

For a single frequency bin k, the squared constraint (5.7) reads as

$$\left| \frac{1}{L} \sum_{l=1}^{L} w_l e^{2\pi j \frac{k}{L} l} \right|^2 < \frac{1}{U_{\max}^2(k)} \tag{5.8}$$

where w_l are the coefficients of the L taps long filter $\mathbf{w}$. The exponential Fourier Kernel can be separated into cosine and sine operations and the squared constraint avoids computing the root. This decomposition yields a real numbered expression

$$\left(\frac{1}{L} \sum_{l=1}^{L} w_l \cos\left(2\pi \frac{k}{L} l\right) \right)^2 + \left(\frac{1}{L} \sum_{l=1}^{L} w_l \sin\left(2\pi \frac{k}{L} l\right) \right)^2 \tag{5.9}$$

on the left hand side of inequality (5.8) that has a complexity of only $\mathcal{O}(N)$. Constraint (5.7), for comparison, has a complexity of at least $\mathcal{O}(N \log_2 N)$ because it requires a complete Fourier transform of $w(l)$.

5.2.4 Measures to Preserve Feedback Stability

We know that feedback stability cannot be guaranteed if one of the above conditions is violated. In [Rafaely and Elliott, 1999] and [Rafaely and Elliott, 2000] it is shown, that constraint (5.7) can be directly incorporated in the controller design, and in [Kinney et al., 2008], the same constraint is checked before each filter update. We choose the latter approach, since it can be applied to any of the above introduced constraints.

If one of the above constraints is violated, W must not be updated. Instead, it makes sense to change W to a stable default filter because if S deviates from $\hat{S}$, the FxLMS update is not reliable anymore. From equation (5.7), we know that the default filter $\tilde{W}$ yields stability if its absolute value is smaller than the inverse of the maximum uncertainty. The easiest implementation of such a time-domain default filter is a scaled Kronecker impulse

$\delta(l)$:

$$\tilde{w}_l = \frac{\delta(l)}{\max(U_{\max}(\omega)) + \epsilon}, \tag{5.10}$$

where l is the index of filter coefficients and ϵ is a small quantity that ensures the inequality of (5.7).

In order to prevent time variances during the filter process, $\mathbf{w}$ should not change abruptly, but converge smoothly to the scaled impulse. Therefore the difference $\boldsymbol{\Delta}_w = \tilde{\boldsymbol{w}} - \boldsymbol{w}(n)$ between the coefficients of the default filter $\tilde{\boldsymbol{w}}$ and the current filter $\boldsymbol{w}(n)$ is used as update. If the difference-vector is normalized and scaled by the step-size parameter μ, the update yields a similar effective step-size as the normalized FxLMS

$$\boldsymbol{w}(m+1) = \boldsymbol{w}(m) + \mu \frac{\boldsymbol{\Delta}_w}{\boldsymbol{\Delta}_w^T \boldsymbol{\Delta}_w}, \tag{5.11}$$

where $m = n \ldots n + M$ with $M = \text{round}\left(\frac{\boldsymbol{\Delta}_w^T \boldsymbol{\Delta}_w}{\mu}\right)$. The time index m of the interim update thus starts with the recent sample n and the update is repeated M times until $\boldsymbol{w}(m) = \tilde{\boldsymbol{w}}$. The difference vector $\boldsymbol{\Delta}_w$ stays the same during the whole interim update which results in a linear fade to $\tilde{\boldsymbol{w}}$.

After the given number of repetitions, the NFxLMS is applied again. Apart from possible short breaks, the NFxLMS is thus constantly running also if S constantly deviates form $\hat{S}$. In that case, the filter starts growing again and it eventually violates one of the constraints again. It thus will constantly grow and scale back to the stable impulse.

5.3 Responses of Prototype Headphones

It has been shown that both, the stability of the LMS adaptation and the feedback stability, depend on the deviation of S from $\hat{S}$. Therefore, the

following sections take a closer look at the variations of the secondary-path once again and review both stability issues with the experimental data.

In chapter 2, secondary-path measurements were presented that cover the extreme cases of completely loose and very tight headphones as well as two further secondary-paths variations between these cases. Fig. 5.3 shows once again the bode plots of the measurements.
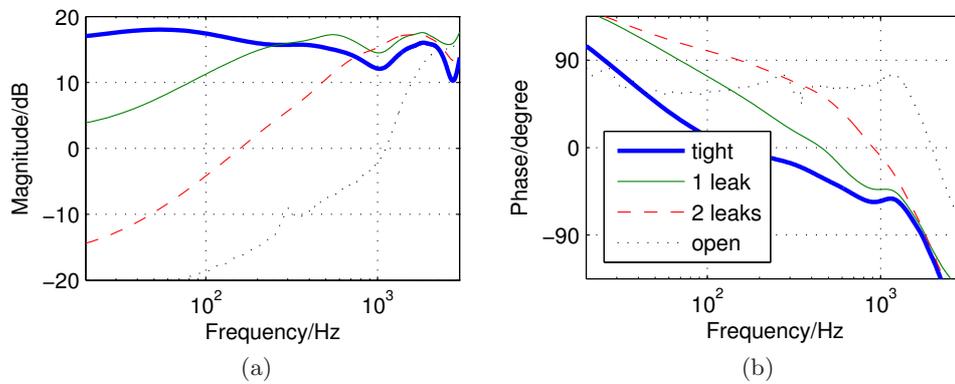
Figure 5.3: (a) Magnitude- and (b) phase response of the secondary-path for tight, leaky and completely loose headphones. The increased leakage leads above all to a magnitude drop-off at low frequencies.

The tight wearing situation is the regular use-case; consequently the secondary-path measure under tight condition will be the nominal model $\hat{S}$. This however implies that there is a large uncertainty on this model if the headphones are lifted as shown in Fig. 5.4.

The maximum uncertainty $U_{\max}(\omega)$ is the one of the three uncertainties that has the largest magnitude at the given ω. As can be seen in Fig. 5.4, $U_{\max}(\omega)$ is mostly determined by the secondary-path of the completely loose (open) headphones. However, below 1000 Hz, the uncertainty of S with two inserted leaks is equally large and even slightly larger around 300 Hz.

In [Wang et al., 2013], the convergence of a narrowband adaptive feedback system is investigated. A constraint on the step size μ is deduced to
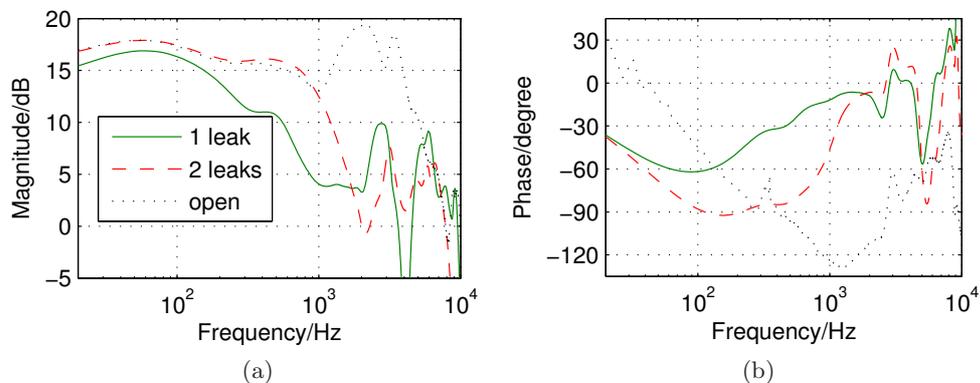
Figure 5.4: (a) Magnitude of the additive uncertainty $U(\omega) = |\hat{S}(j\omega) - S_i(j\omega)|$ and (b) phase error $|\arg \hat{S}(j\omega) - \arg S_i(j\omega)|$ of differently positioned headphones.

keep the system stable, even under imperfect nominal models. The influence of the secondary-path deviations is then examined separately for the FxLMS convergence and for the feedback stability in order to gain a better understanding of the system properties. We will adopt this methodology and treat the FxLMS stability in the following section and the feedback stability in section 5.5, where we will again apply the Nyquist criterion because it is suitable for broadband analysis, too.

5.4 Experimental Considerations for the Stability of Adaptation

5.4.1 Robustness Against Phase Mismatch

In Fig. 5.4b, it can be seen that the phase error between the nominal secondary-path model $\hat{S}$ and the secondary-path of the open headphones is exceeding 90° around 1300 Hz. The usual FxLMS algorithm would thus diverge if the headphones were lifted, as it is also shown in [Wang et al.,

2013]. It is therefore necessary to implement the leaky FxLMS as described in section 5.1.

The leakage factor γ, which is a trade-off between stability and excess error, has to be determined empirically. We thus simulate the normalized leaky FxLMS with the given measurement data in three test set-ups:

1. A narrowband excitation around 1300 Hz with the secondary-path of the open headphones S_{open}. This is the set-up with the largest phase error, as already mentioned.

2. A broadband excitation with S_{open}.

3. A broadband excitation with the secondary-path of the headphones with two leaks S_2 because it also slightly exceeds the phase error of 90°.

The broadband excitation is white noise that is filtered with a cascade of two first order low-pass filters with 500 Hz and 1000 Hz cut-off frequency, respectively. This low-pass cascade simulates the passive attenuation of the headphones as was shown in chapter 4 and yields 33 dB attenuation at 3500 Hz. Since we do not expect more than 30 dB of noise cancellation, aliasing components below -33 dB do not influence the ANC. Thus, no further anti-aliasing filter is required if the sampling frequency f_s is above 7000 Hz. We choose $f_s = 7350$ Hz because it is one sixth of 44.1 kHz, which is a common sampling frequency in audio technology. We set the filter length to L = 6 taps and choose the step size according to the largest eigenvalue of the input autocorrelation matrix as $\mu = \frac{1}{2\lambda_{\text{max}}}$ which yields the fastest convergence [Kuo and Morgan, 1999].

We test the leaky NFxLMS in open loop condition in order to decouple the convergence of the filter from the possible feedback instability as in

[Wang et al., 2013], and find that $\gamma = 0.005$ is the smallest leakage factor that yields a stable update. Fig. 5.5 shows the related error as $e_{\mathrm{dB}} = 10 \log \frac{\bar{e}^2[n]}{\bar{d}^2[n]}$, where $\bar{e}^2[n]$ and $\bar{d}^2[n]$ are smoothed by a moving average over the error and the input signal, respectively.
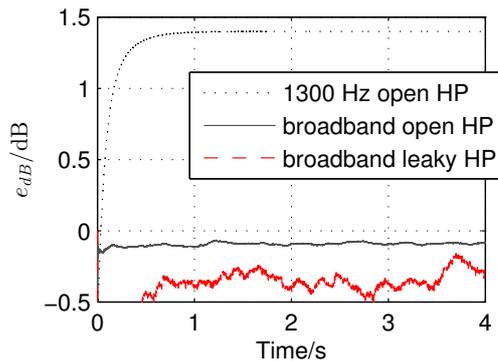
Figure 5.5: The related residual error e_{dB} for 3 worst case scenarios: A narrowband excitation at 1300 Hz for S_{open}, and a broadband excitation for S_{open} and S_2. The excitation at 1300 Hz causes an error which lies 1.4 dB over the excitation level, but the adaption stays stable in all three cases.

In the broadband cases, the filter even yields a small noise reduction. In the narrowband case, the filter causes an amplification of the input noise of 1.4 dB which on the one hand proves that the NFxLMS does not converge to the optimal solution anymore, but on the other hand the simulation also shows that the filter coefficients stay bounded because of the leakage factor γ.

5.4.2 Robustness Against Excessive Amplification

The robustness against phase errors in $\hat{S}$ is not the only advantage of the leaky LMS. It also makes the system more robust against sensor noise and non-linearities in S that might occur at loud playback volumes.

Fig. 5.6 compares the converged filter of the previous leaky-NFxLMS simulation with the optimum filter W_{opt} that yields the minimum mean

square error (MMSE). Since there is hardly any high-frequency excitation, the filter W_{opt} can boost these high frequencies without significantly increasing the error.
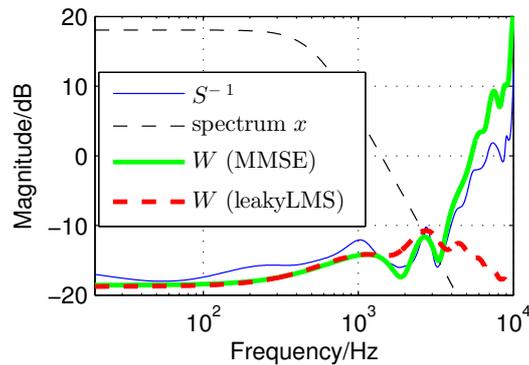
Figure 5.6: The filter W tries to match the inverse of S. Without constraints (MMSE), it has the freedom to heavily boost the high frequencies since there is no noise excitation in this band (due to the passive attenuation of the headphones). (For convenience, the SPL of the excitation d is scaled to approx. 18 dB.) The leaky LMS (with $\gamma = 0.005$) minimizes the filter's energy which leads to a desirable roll-off at high frequencies.

The boost of the high frequencies is detrimental to performance in a real life condition, where e.g. sensor noise or estimation errors of $\hat{d}$ would be strongly amplified. The leaky NFxLMS solution on the other hand prevents the filter from excessively amplifying the high frequencies. It matches the inverse of S less accurately and has a larger excess error below 4000 Hz, but it still yields more than 10 dB ANC between 100 and 500 Hz as can be seen in Fig. 5.7.
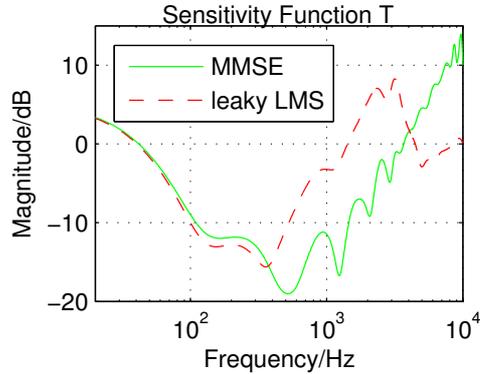
Figure 5.7: Sensitivity function of the NFxLMS for $S = \hat{S}$ with and without leakage factor. Without leakage factor, the NFxLMS converges to the MMSE but it produces a large gain on the high frequencies. With leakage factor, the sensitivity function rolls off at high frequencies but it does not yield the optimum performance below 4000 Hz.

5.5 Experimental Considerations for Feedback Stability

5.5.1 High Frequency Uncertainty

A large uncertainty at high frequencies is inherit in most physical systems since small changes in the plant already lead to large phase differences at high frequencies. In our case, this uncertainty is especially critical around 3000 Hz because the adaptive filter amplifies this frequency band as it was shown in the previous simulation in Fig. 5.6 and 5.7.

It amplifies the middle and high frequency band because the raising slop in W causes a negative group delay which is required to compensate for the delay in the secondary path. The adaptive filter is 'allowed' to amplify the high frequencies because there is hardly any noise excitation in this band, and consequently, the amplification has little influence on the error power.

The amplification can be penalized if the error is weighted stronger at high frequencies. With the weighting filter $H(j\omega)$, the cost function of the

ANC system reads as

$$J(\mathbf{w}) = E\{\mathbf{e}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{e}\}, \qquad (5.12)$$

where $\mathbf{e}$ is a N taps long vector of error samples and $\mathbf{H}$ is the convolution matrix

$$\mathbf{H} = \begin{bmatrix} h_1 & 0 & \ldots & 0 \\ h_2 & h_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_N & h_{N-1} & \ddots & h_1 \\ \vdots & \ddots & \ddots & \vdots \\ h_M & h_{M-1} & \ddots & h_N \\ 0 & h_M & \ddots & h_{N+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & h_M \end{bmatrix}, \qquad (5.13)$$

with h_i being the coefficients of the M taps long impulse response of the weighting filter $H(j\omega)$.

With a neglected secondary path, the cost function can also be written as

$$J(\mathbf{w}) = E\left\{\left(\mathbf{d}^{\mathrm{T}} - \mathbf{w}^{\mathrm{T}}\hat{\mathbf{D}}\right)\mathbf{H}^{\mathrm{T}}\mathbf{H}\left(\mathbf{d} - \hat{\mathbf{D}}\mathbf{w}\right)\right\}, \qquad (5.14)$$

where $\hat{\mathbf{D}}$ is a matrix of estimated input samples

$$\hat{\mathbf{D}} = \begin{bmatrix} \hat{d}_{[n]} & \hat{d}_{[n-1]} & \ldots & \hat{d}_{[n-L+1]} \\ \hat{d}_{[n+1]} & \hat{d}_{[n]} & \ddots & \hat{d}_{[n-L+2]} \\ \vdots & \ddots & \ddots & \vdots \\ \hat{d}_{[n+N-1]} & \hat{d}_{[n+N-2]} & \ldots & \hat{d}_{[n+N-L]} \end{bmatrix}. \qquad (5.15)$$

113

This leads to the update

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \mu \hat{\mathbf{D}}[n] \mathbf{H}^{\mathrm{T}} \mathbf{H} \mathbf{e}[n], \tag{5.16}$$

and by using the convolution operator * instead of the convolution matrix, the normalized FxLMS follows to

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \mu \frac{(\mathbf{h} * e[n])\left(\hat{\mathbf{s}} * \mathbf{h} * \hat{\boldsymbol{d}}[n]\right)}{\hat{\boldsymbol{d}}^{\mathrm{T}}[n]\hat{\boldsymbol{d}}[n]}, \tag{5.17}$$

where $\hat{\mathbf{s}}$ and $\mathbf{h}$ are the impulse responses of $\hat{S}$ and H, respectively.

The block diagram of the new update is shown in Fig. 5.8. The filters H can be seen as whitening filters for the band limited signals $\hat{d}$ and e.
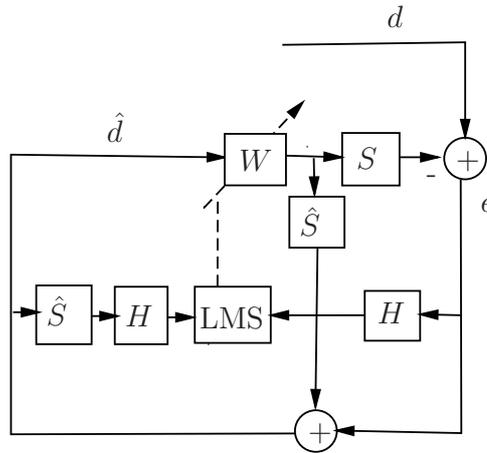
Figure 5.8: Block diagram of the adaptive feedback system with an additional filter H that penalizes the high frequency error.

A complete whitening however would be counterproductive because a band limitation is necessary for the required signal prediction. It is therefore advisable to use a high shelf filter. Fig. 5.9 compares three different shelving filters with corner frequencies at 2000 Hz, 3000 Hz and 4000 Hz, respectively, and shows the consequences onto the adaptive filter gain.
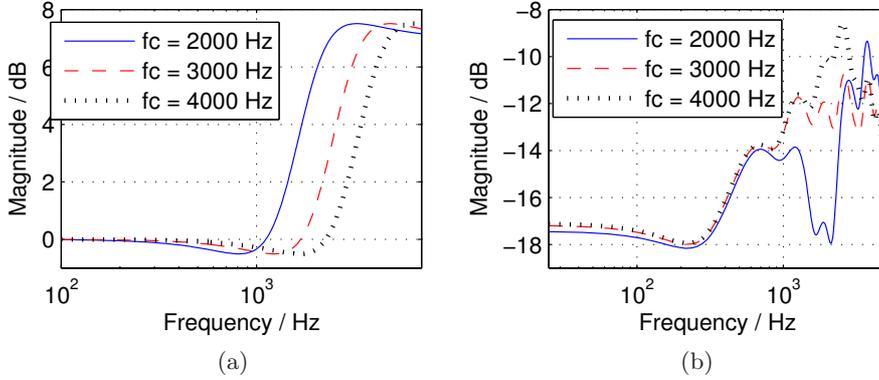
Figure 5.9: (a) Magnitude response of the shelving filter H for three different corner frequencies. (b) Magnitude response of the converged filter W dependent on the used shelving filter.

A shelving filter at 2000 Hz penalizes the frequency band around 2000 Hz very strongly. The shelving filter at 4000 Hz on the other hand has too little influence onto this frequency band. Fig. 5.9b shows that the shelving filter at 3000 Hz is the most suitable to penalize the error signal.

The alternative to penalizing the error in the upper frequency band is to directly penalize the filter gain in this band. The N taps of $\mathbf{w}$ can be weighted with the magnitude response of the high shelf filter $H(j\omega)$ by the convolution operation $\mathbf{Hw}$. The cost function with the frequency dependent penalty term then reads as

$$J(\mathbf{w}) = E\{e^2 + \gamma(\mathbf{w}^T\mathbf{H}^T\mathbf{H}\mathbf{w})\}. \tag{5.18}$$

The frequency dependent penalty on the filter norm results in a frequency dependent leaky FxLMS

$$\mathbf{w}[n+1] = \left(\mathbf{I} - \mu\gamma\mathbf{H}^T\mathbf{H}\right)\mathbf{w}[n] + \mu\frac{e[n]\hat{\mathbf{S}}\hat{\boldsymbol{d}}[n]}{\hat{\boldsymbol{d}}^{\mathrm{T}}[n]\hat{\boldsymbol{d}}[n]}, \tag{5.19}$$

where $\mathbf{I}$ is the identity matrix.

Fig. 5.10 shows the magnitude response of the converged filters W for the frequency dependent leaky FxLMS with the three shelving filters from Fig. 5.9a. The shelving filter at 2000 Hz yields the most suitable penalty for the adaptive filter gain.
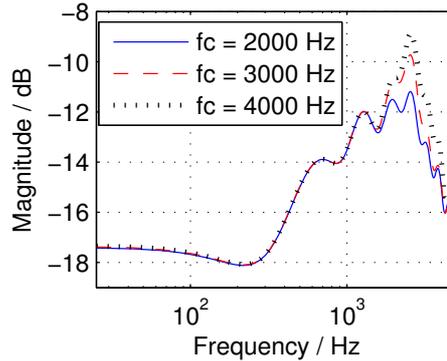
Figure 5.10: Magnitude response of W for the frequency dependent leaky FxLMS dependent on the shelving filters from Fig. 5.9a.

The shelving filter $H(j\omega)$ results in a frequency dependent leakage $(\mathbf{I} - \mu\gamma\mathbf{H}^{\mathrm{T}}\mathbf{H})$ that depends on the step size μ and on the trade-off factor γ. Fig. 5.11 shows this frequency dependent leakage that results from the shelving filter at 2000 Hz with $\gamma = 0.0005$, and compares it with a trade-off factor that is twice as large and half as large, respectively.

Fig. 5.12 compares the ANC results for the two different approaches: (i) The frequency dependent penalty on the error and (ii) the frequency dependent penalty on the filter gain. It can be seen that both approaches yield a very similar performance.

Both, the frequency dependent leaky LMS (with an empirically derived γ) and the additional shelving filters on e and $\hat{d}$ successfully reduce the filter gain above 2000 Hz, but the large uncertainties for the lifted headphones are still a problem and they require further constraints on the adaptive filter W that will be treated in the following section.
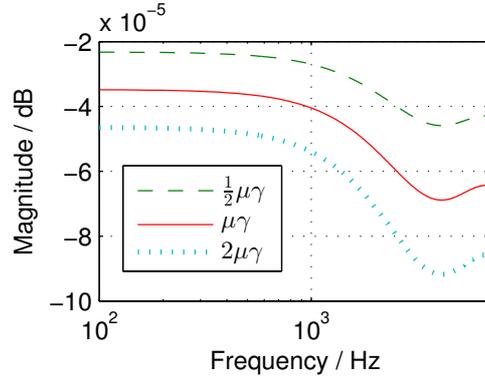
Figure 5.11: The high-shelf penalty on the filter gain results in a frequency dependent leakage on $\mathbf{w}$ that has a low shelf characteristic. The resulting leakage does not only depend on the original shelving filter H but also on the parameters γ and μ.
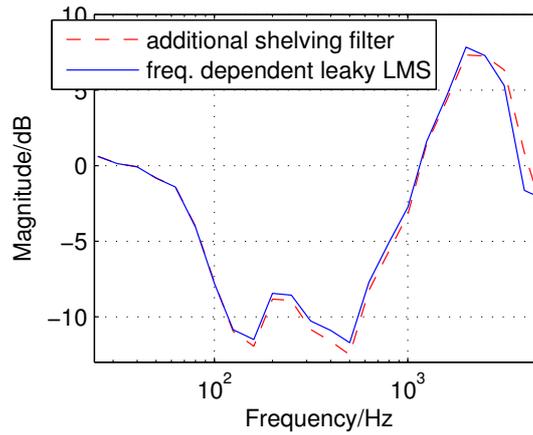
Figure 5.12: Sensitivity function for the FxLMS with an additional shelving filter at 3000 Hz and for the frequency dependent leaky FxLMS with a shelving filter at 2000 Hz.

5.5.2   Low Frequency Uncertainty

The measurements from Fig. 5.3 cover the main variations of the secondary-path. Although this set of assessed secondary-paths S_a is very representative, it has to be assumed that there are secondary-path realizations S_i with magnitude and phase responses somewhere between the ones of S_a. It can furthermore not be guaranteed that the (yet unknown) secondary-paths S_i

117

yield stable feedback just because the set S_a does. The constraint on the real part of the open loop (eq. (5.3)) is therefore not suitable for ANC headphones.

On the other hand, the measurements show that the maximum uncertainty is well described by secondary-paths of the loose and very leaky headphones. The constraint $|W| < \frac{1}{U_{\max}}$ can thus be considered as being robust. However, it is a rather conservative constraint, too. Fig. 5.13 e.g. shows an exemplary case where the feedback would be stable but the inequality of eq. (5.6) is violated.
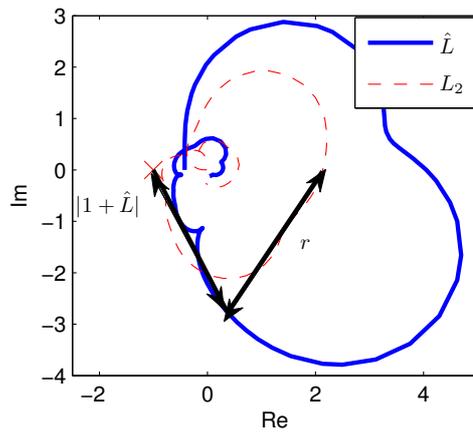
Figure 5.13: Nyquist contour of the open loop L_2 (with the secondary path of two leaks S_2) for a filter realization with the experimental data together with the nominal open loop $\hat{L}$.

The figure shows the Nyquist contour of the open loop L_2 (with the secondary-path of two leaks S_2) for a filter realization with the experimental data together with the nominal open loop $\hat{L}$. Both Nyquist contours are entirely on the right hand side of point (-1,0). Thus, the feedback is stable and the condition on the real part of L (eq. (5.3)) is satisfied, too. However, the radius of uncertainty is larger than the distance of $\hat{L}$ from (-1,0) which means that the condition using $U_{\max}$ (eq. (5.6)) does not hold and W would

be unnecessarily constrained. Nevertheless, for the sake of robust stability, we use the more conservative constraint to handle the remaining uncertainty in the lower frequencies. Moreover, we will show that a constraint on single frequency bins as in (5.8) is sufficient to preserve stability, too.

As stated above, the stability of the feedback loop depends on the denominator of the sensitivity function

$$T = \frac{1 - W\hat{S}}{1 - W(\hat{S} - S)}.$$

As long as $S = \hat{S}$, the feedback is stable and W converges to S^{-1}. For narrowband excitations, W converges to the ideal inverse of S in the corresponding frequency bin. For broadband excitations, W converges to a causal approximation of S^{-1}.

Fig. 5.14 illustrates the scenario when S changes from the tight secondary-path S_{tight} (which equals $\hat{S}$) to the secondary-path with one leak S_1. It shows the magnitude of $\left|W(\hat{S} - S_1)\right|$ and first, it is assumed that W is still an approximation of S_{tight}^{-1}. We distinguish between the ideal inverse of S_{tight} for narrowband analysis and the causal inverse for broadband analysis. The causal version has been derived over the frequency dependent leaky FxLMS algorithm.

When the headphones are lifted, the feedback initially stays stable because $\left|W(\hat{S} - S_1)\right| < 1$, but the product is close to 1 below 300 Hz. Hence, this is the frequency band where a pole outside the unit circle is most likely. This is especially true in the second step when W adapts to the inverse of S_1 because W has to amplify these low frequencies. In this case, $\left|W(\hat{S} - S_1)\right|$ exceeds unity gain below 300 Hz.

Since $\left|W(\hat{S} - S_1)\right| < 1$ is violated in the low frequency band first, it is
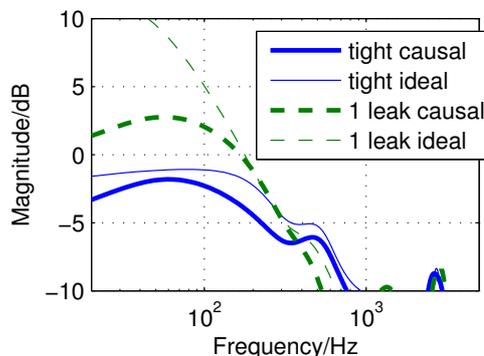
Figure 5.14: Product of the adaptive filter times the additive uncertainty for the secondary path with one leak $\left|W(\hat{S} - S_1)\right|$: The filter W needs to invert S. The blue lines show the case where W is still the inverse of S_{tight} (thus before it has reacted to the new secondary-path S_1). The green lines show the case where W finally changed to the inverse of S_1. For narrowband analysis, W is the ideal inverse of S; for the broadband analysis, W is derived over the frequency dependent leaky FxLMS.

proposed to only check the low frequencies of $WU_{\max}$. It is thus sufficient to determine $W(k)$ at the frequency bins k below 300 Hz instead of doing a full Fourier-transform.

The frequency resolution of W determines how many frequency bins need to be evaluated. W has to approximate the inverse of S up to 1200 Hz, and under regular tight conditions, S has hardly any spectral variation in this band. A short filter with a low frequency resolution is therefore sufficient for feedback ANC (compare [Oppenheim et al., 1994]). If the bin spacing of the filter is larger than 300 Hz, the changes of W below 300 Hz will above all become noticeable in the DC-bin. This is very beneficial because the constraint on the DC-gain of W does not need the decomposition into cosine and sine as in (5.9). It can easily be formulated in the time-domain as

$$\left|\sum_l w_l\right| < \frac{1}{U_{\max,0}}, \tag{5.20}$$

where $U_{\text{max},0}$ is the maximum uncertainty below 300 Hz. Consequently, the entire check for robust stability only requires L multiply-accumulate operations per filter update.

In Fig. 5.4, it can be seen that $U_{\text{max},0}$ is at 17.3 dB. Once the DC-gain exceeds the given threshold, the adaptive filter is changed to the stable default filter $\tilde{\mathbf{w}}$ as in eq. (5.11). With a headroom of $\epsilon = 2.7$ dB, the default filter is scaled to an impulse of -20 dB according to eq. (5.10).

5.6 Experimental Results

5.6.1 Summary of the Algorithm

The whole adaptive feedback-ANC algorithm with DC-constraint can be summarized as follows:

1. Update the coefficients of W via the freqency dependent leaky NFxLMS algorithm as in eq. (5.19).

2. Check if $|\sum_i w_i| < \frac{1}{U_{\text{max},0}}$. If no:

   - Calculate the gradient which leads to the stable default filter $\boldsymbol{\Delta}_w = \tilde{\boldsymbol{w}} - \boldsymbol{w}(n)$.

   - For the next $M = \left|\frac{\boldsymbol{\Delta}_w^T \boldsymbol{\Delta}_w}{\mu}\right|$ samples update the coefficients as $\boldsymbol{w}(m+1) = \boldsymbol{w}(m) + \frac{\boldsymbol{\Delta}_w}{M}$.

   - After M iterations continue the leaky NFxLMS update (jump to step 1).

The development of the algorithm is based on responses of prototype headphones, but all steps are motivated by physical reasons that apply to all headphones of similar design.

1. Small changes in S will always lead to phase deviations in the upper frequency band. Restricting the ANC with an additional high-shelf or with the frequency dependent leaky FxLMS is thus always a good measure to increase the robustness of ANC headphones.

2. A leak between the ear-cups and the ears mainly changes the low frequency response of the secondary-path. Thus, there will always be a large uncertainty at low frequencies that demands a constraint on W. The constraint on the norm of W is always applicable, and the constraint on single frequency bins of W and especially on the DC gain is suitable for a low frequency resolution of W. In general, W does not have to equalize a lot, since a flat frequency response is desired for high-fidelity headphones. Therefore, a low frequency resolution of W and thus a short filter length is generally suitable.

3. The maximum uncertainty that is required to scale the default filter can easily be assessed with preliminary measurements as described in sec. 5.3.

5.6.2   Evaluation of the Algorithm

In the above analysis and deduction of the time-domain constraint, it is assumed that W converges to S^{-1}, but does not exceed the magnitude of S^{-1}. This is most likely because the leaky FxLMS penalizes the norm of W. An extensive numerical analysis is still necessary to prove the robustness of the time-domain constraint. The numerical analysis has to be done through simulations of the adaptive feedback system because firstly the adaptive filter depends on the input signal, and secondly the temporal behaviour of W is important. It has to be examined whether W always exceeds constraint

(5.20) before the feedback loop starts ringing.

The analysis is performed for various input noises:

- Sinusoidal excitation at 50 Hz and in 100 Hz steps from 100 Hz to 1400 Hz.

- Narrowband excitation with white noise passed through a 2nd order peak filter with a quality factor of $Q = 8$ and centre frequencies as before.

- Broad band excitation with white noise.

- Broad band excitation with pink noise.

Each of the excitation signals is filtered with the low-pass characteristic of the passive attenuation. The simulations are not only done for changes from S_{tight} to S_1, but also with

- an initial worst-case S and for

- a sudden change from S_{tight} to the worst case S.

The worst case S is the one with the largest uncertainty in the corresponding frequency band, cf. Fig. 5.4. Thus, the worst case S below 1000 Hz is the secondary-path with two inserted leaks, and above 1000 Hz, it is the open secondary-path. The leaky NFxLMS is run with simulated white sensor noise of -60 dB relative to the excitation level and with the same μ and γ as in section 5.4. With a sampling frequency of 7350 Hz, we choose a 6 taps long filter which results in a sufficient frequency resolution of 1200 Hz per bin.

We calculate the energy of the input noise $E_d(n)$ and the error-energy

$E_e(n)$ over the last M samples as

$$E_d(n) = \sum_{m=n-M}^{n} d(m)^2$$
$$E_e(n) = \sum_{m=n-M}^{n} e(m)^2. \qquad (5.21)$$

The integration time M should be approximately as long as one period of the lowest frequency under consideration, and it is therefore chosen as $M = 167$ samples which corresponds to one period of a 50 Hz tone. The start of ringing is detected when all of the three properties are true:

1. E_e grows such that $E_e(n) > E_e(n - M)$,

2. $E_e(n)$ is larger than twice $E_d(n)$ (i.e. $E_e(n) > 2E_d(n)$), and

3. E_e grows twice as fast as E_d, i.e. $E_e(n) - E_e(n - M) > 2(E_d(n) - E_d(n - M))$

Every time the three properties are detected, $W(n)$ is marked as unstable. In contrast, if

1. E_e decreases such that $E_e(n) < E_e(n - M)$, and

2. $E_e(n)$ is smaller than $E_d(n)$,

W is marked as stable. The condition $E_e(n) < E_e(n-M)$ is not necessary to indicate proper ANC, but we are interested in instances where W is changing and these instances are most probable when ANC is improving.

Fig. 5.15 shows the DC-gain distribution of all filters which are marked as stable and unstable respectively. The DC-gain of the filters which are recorded before ringing always exceeds $\frac{1}{U_{\max,0}} = -17.3$ dB. Also the filters which improve ANC exceed this threshold in some cases. Thus it occurs
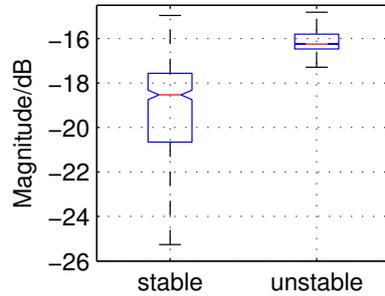
Figure 5.15: DC-gain distribution of W under stable condition and before instabilities occurred. The boxes include 50% of the registered DC-gains and the whiskers show the whole distribution. The bars in the middle of the boxes denote the median value and the notches indicate the confidence interval. The DC-gain of W always exceeds -17.3 dB before instability occurred.

that the adaptive filter is switched to $\tilde{\mathbf{w}}$, although the system would have been stable. However, it is more important to see that the constraint on the DC-gain is a robust constraint to keep the feedback stable. A further test of the algorithm with additional experimental data is applied in the following.

5.6.3 Performance Comparison

The experimental data for the development and evaluation of the algorithm is gathered by measurements on a mannequin. In this section, we test the algorithm with experimental data from real persons and we compare the performance with existing approaches from the literature. As stated in the introduction, there are two main strategies to handle secondary-path uncertainties.

Consideration of uncertainty

The adaptive filter is bounded by the inverse of $U_{\max}$ as in constraint (5.7). In [Rafaely and Elliott, 1999], this constraint is included in the controller design, but the controller is non-adaptive. In [Kinney et al., 2008], an adap-

tive filter is used and constraint (5.7) is checked before each filter update. If the constraint is violated, the filter is not updated. In [Rafaely and Elliott, 2000], the constraint is included in the cost function of the LMS as

$$J(W_{(k)}) = E\left\{e^2 + \sigma \max\left[|W(k)|^2 - \frac{1}{U_{\max}^2(k)}, 0\right]\right\}, \qquad (5.22)$$

where the value of $\max\left[|W(k)|^2 - \frac{1}{U(k)_{\max}^2}, 0\right]$ is zero if constraint (5.7) holds. If it does not hold, the adaptive filter is scaled back by a factor that depends on the weight σ. Rafaely uses the frequency domain LMS. Therefore the adaptive filter can easily be scaled back in the frequency bin where the constraint is violated without affecting the other bins of W. Therefore, this approach can be considered as the most elaborated and it will be used for the following comparison.

Online secondary-path estimation

The second strategy is to obtain a permanent estimate of the current state of S. The most robust approach is to inject an auxiliary signal into the headphones in order to identify the secondary-path [Akhtar et al., 2006, Zhang et al., 2001]. In [Gan et al., 2005], the music playback is used as auxiliary signal, but it only works for slowly changing secondary-paths. Even approaches that inject white noise fail when there are fast and large changes in S. The most advanced method is presented in [Zhang et al., 2003]. It comprises three measures: (i) A third adaptive filter reduces the disturbance of the noise cancelling error onto the secondary-path adaptation. (ii) The auxiliary noise is scaled in dependence of the convergence status in order to keep the play-back volume as low as possible. (iii) A hard constraint on the norm of the adaptive controller prevents divergence, even for sudden large

changes in S. The method is proposed for feedforward ANC where it avoids the divergence of W for the most part. We will include it in the following comparison to investigate if it is extendible to feedback ANC.

Denote that measure (iii) of Zhang's approach actually falls into the first category, too. The threshold for the norm constraint and the parameter σ in Rafaely's approach are determined empirically by tuning them in order to minimize error amplifications. Our approach falls into the first category, too, but we build on the finding that constraint (5.7) is violated around DC, while all other solutions from strategy *1)* analyse the full bandwidth from 0 Hz to $\frac{\text{fs}}{2}$. The benefit of our approach is that no real-time Fourier transform and no auxiliary noise is needed. The drawback is that the frequency resolution has to be very low. The approaches of Rafaely and Zhang do not have this restriction. For their approaches, we choose a frequency resolution that is two times as fine as ours to demonstrate the limitations of our controller. Thus, we set their filter length to $L = 12$ taps, while ours has only 6 taps. The longer filter means a slower convergence (cf. [Kuo and Morgan, 1999] and eq.(5.2)), but the comparison will show that the increased resolution outweighs the slightly slower convergence.

To obtain real-life experimental data, we asked two test subjects to put on the headphones in differently tight, leaky and lifted positions. For the tight measurements, the subjects were also asked to press the headphones to the ears. For all headphone-positions, we measured the secondary-path with a sine sweep yielding 16 different measurements shown in Fig. 5.16. The measurements indicate again that leaks between the ears and the headphones mainly affect the low frequency region.

In all following evaluations, we apply the same measures:

- Every 0.5 s, the measured secondary-paths are replaced by each other
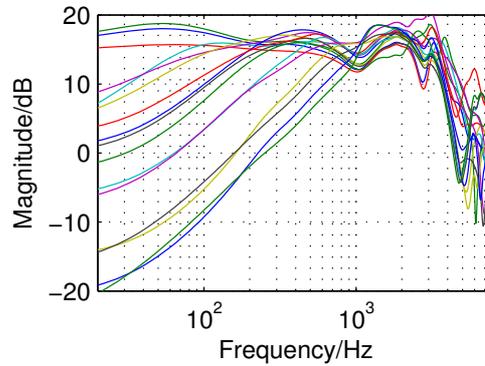
Figure 5.16: Magnitude response of the secondary-paths measured on two persons with differently positioned headphones.

in a way that tight and leaky fittings alternate.

- As (initial) secondary-path model, we still use $\hat{S}$ which we derived from the tight measurement on the dummy head.

- The excitation signals are filtered with a low-pass filter cascade that approximates the passive attenuation of the headphones as in section 5.6.2.

- The high frequency uncertainty is reduced by the frequency dependent leaky FxLMS.

First, pink noise and narrowband noise around 100 Hz were chosen as excitation signals because they lead to instabilities if the normalized leaky FxLMS is run without further constraints (cf. Fig. 5.17).

Fig. 5.18 compares the performance of our algorithm (labelled with dc) with the approaches of Rafaely and Zhang for the same conditions. The performance of all three systems is limited because of the fast and violent changes in S. However, it can be seen that our approach preserves stability for the real-life experimental data, too. Consequently, also Rafaely's approach has to preserve stability because it utilizes the same constraint

128
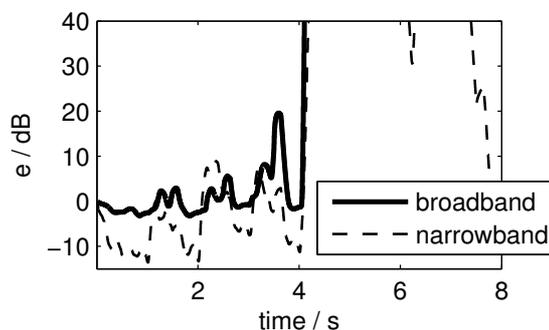
Figure 5.17: The relative residual error e_{dB} over 8 seconds of a normalized leaky FxLMS for pink and narrowband noise around 100 Hz: The secondary-path changes abruptly every 0.5 seconds which leads to a relative error of +10 dB and more in the first 4 seconds and to complete instability afterwards.

only extended to the full bandwidth. Furthermore, it is demonstrated that Zhang's method can be extended to feedback ANC, where it preserves stability, too. However, the performance for broadband noise is deteriorated due to the additionally injected noise and the hard constraint on the norm of W (cf. Fig. 5.18a).

On the other side, the performance of our system is slightly degraded for the low-frequency narrowband noise. There are two reasons for this:

1. As stated in section 5.1, the adaptive filter aims to compensate the phase delay and dynamics of S in the frequency band of the excitation d. As long as the DC-constraint is violated, the adaptive filter is scaled back to an impulse $\tilde{\mathbf{w}}$. The impulse is a suitable broadband compensation for S, but it is a suboptimal narrowband compensation. Fig. 5.19a compares the noise spectrum with the error spectra over the eight seconds of the experiment in third octave bands. Our system yields only 1 dB of noise reduction in average while Rafaely's and Zhang's approach still yield a reduction of 4-6 dB.

129
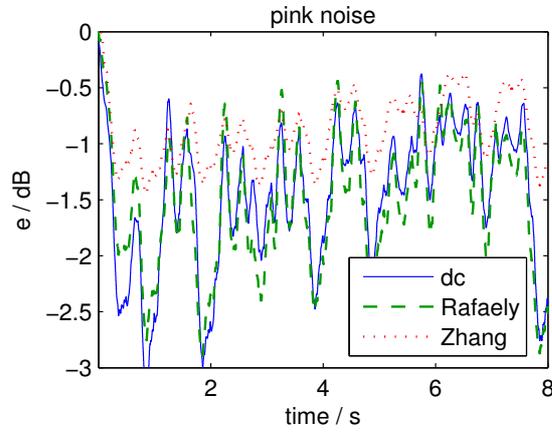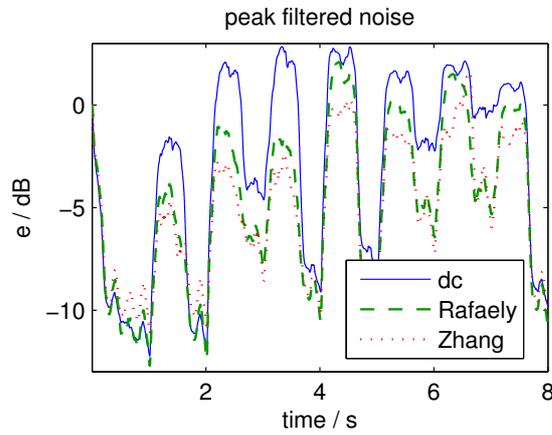
Figure 5.18: e_{dB} for (a) broadband noise and (b) narrowband noise around 100 Hz: The conditions are the same as in Fig. 5.17, but this time, our DC constraint as well as the constraints by Rafaely and Zhang are applied.

2. The filter length is very short. Fig. 5.19b shows the same comparison for a time frame where S is close to $\hat{S}$ (from 1.5 s to 2 s in Fig. 5.18b). Our approach yields 2 dB less noise reduction than Rafaely's, although the filter has not been scaled back to an impulse during this time. Thus, the 2 dB difference occurs because a six taps long filter has been used in our approach, while a longer filter is used in Rafaely's method.
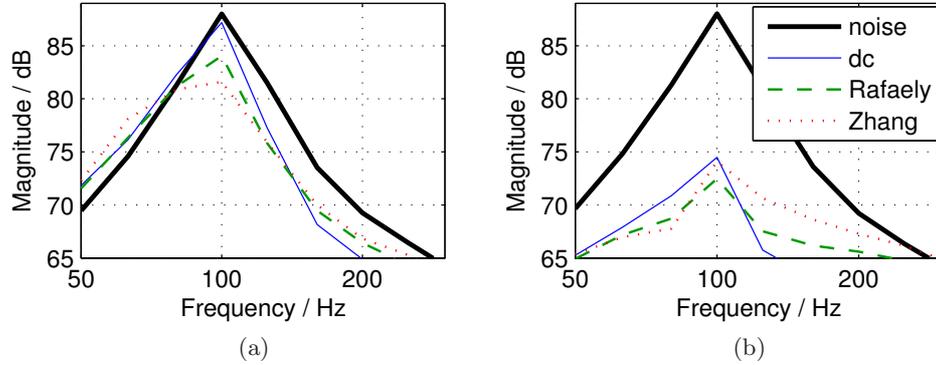
Figure 5.19: Comparison of the narrowband noise power with the error power in third octave bands (a) over the complete 8 seconds (b) during the time where S is close to $\hat{S}$
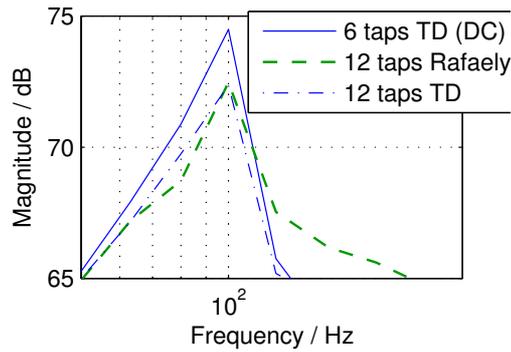
Figure 5.20: Error spectra comparison of our time-domain (TD) approach (6 taps) with Rafaely's approach (12 taps) and our time-domain approach with 12 taps.

Fig. 5.20 indicates that our approach would yield the same noise reduction if a 12 taps long filter would be applied. However, with the increased frequency resolution of the 12 taps long filter, the sinc-shaped low pass of the DC bin decays to approx. -4 dB at 300 Hz. In order to be sure to detect all low frequency changes in W, it might therefore be necessary to extend the frequency analysis to the 600 Hz bin as in eq. (5.8) and (5.9). Instead of 12 multiply and accumulate operations (MACs) to calculate the DC-gain, the frequency analysis would then require additionally 2×12 MACS for

131

the cosine and sine component at 600 Hz and another 2×12 MACS for the square operations yielding 60 MACs in total. Therefore, the 6 taps long filter is preferred since it still yields 14 dB of narrowband noise reduction if the headphones sit regularly tight. The reduced complexity of the shorter filter might outweigh the small loss of performance in most consumer applications.
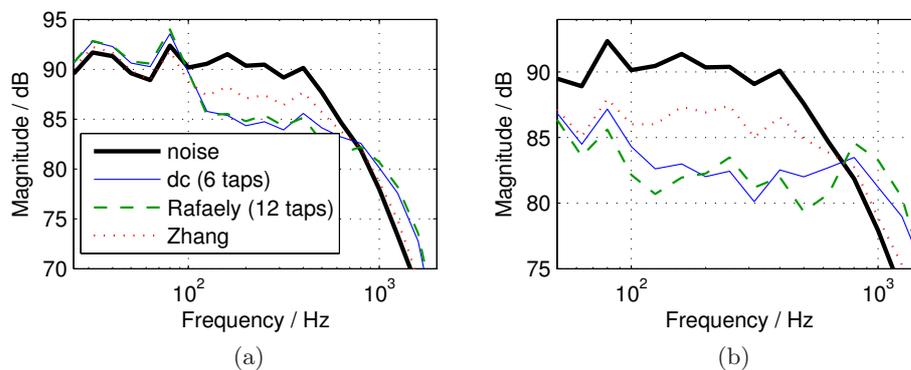

Figure 5.21: Comparison of the broadband noise power with the error power in third octave bands (a) over the complete 8 seconds (b) during the time where S is close to $\hat{S}$

In the broadband case, our approach (with 6 taps) yields the same results as Rafaely's over all secondary-paths as can be seen in Fig. 5.21. However, the latter is computationally far more demanding. Rafaely suggests implementing a frequency-domain LMS with a time-domain adaptive filter. For a filter length of 12 taps, a 24 point FFT of the buffered $\mathbf{d}$ and $\mathbf{e}$, and a 24 point IFFT for $\mathbf{w}$ is required to avoid circular convolution effects [Rafaely and Elliott, 2000, Shynk, 1992]. With a computational complexity of $\mathcal{O}(N \log_2 N)$ per complex-numbered FFT, $2 \times 24 \log_2 24$ real-numbered MACs are required. With two FFTs and one IFFT, at least 661 MACs are required per filter update, while our approach only requires 6 MACs. Even with a filter length of 6 taps, the frequency-domain LMS would require 130 MACs with hardly any advantage over our approach. Zhang's approach can be

implemented in the time-domain, too, but it yields 3-5 dB less noise reduction than Rafaely's and our approach depending on the positioning of the headphones.

To summarize the comparison:

- All approaches under comparison preserve stability of the adaptive feedback system, even after large and sudden changes in S.

- Zhang's and our method can be directly implemented in the time-domain, but Zhang's approach requires three LMS updates, while ours requires only one. Zhang's approach, which does not have a restriction on the frequency resolution, performs slightly better for narrowband excitation, but our approach is clearly superior for broadband excitation.

- Rafaely's approach yields a narrowband noise reduction that is 2-5 dB better than ours and a broadband noise reduction that is comparable to ours. But his approach requires three Fourier transforms per filter update, while ours only requires a single summation for the stability constraint.

- With a 12 taps long filter, our approach performs as well as Rafaely's and Zhang's also for narrowband excitation. However, a 6 taps long filter is good enough for broadband excitation (cf. Fig. 5.21).

Table 5.1 compares the computational complexity of our time domain approach with the frequency domain LMS. Even with the 12 taps long filter, the time domain DC approach requires less MACs than the frequency domain LMS with 6 taps.

Thus, the main contribution of this work is to show that stability of

	6 taps	12 taps
(time domain) DC approach	6	60
frequency domain LMS	130	661

Table 5.1: Required MAC operations for the compared systems.

feedback ANC-headphones can be preserved with a computationally very economical constraint that hardly influences the performance.

5.7   Conclusion

Adaptive feedback-ANC is a very powerful solution for headphones applications, but, as with all feedback systems, it suffers from the risk of instabilities. This chapter presented an algorithm that avoids these instabilities and preserves the benefits of adaptive feedback ANC.

In particular, it has been examined how the secondary-path changes when the headphones are lifted and how these changes affect the stability of the NFxLMS adaption and the feedback loop. It has been proposed to use the leaky NFxLMS to overcome the first stability issue and it has been shown that the leaky NFxLMS keeps the adaption stable even during large changes in the secondary-path.

To overcome the feedback stability issue, an algorithm has been developped that detects changes in the secondary-path. We use the fact that the adaptive filter increases its low-frequency gain if the headphones are lifted. It has been shown that this low-frequency amplification is independent from the excitation-noise characteristic and that it is sufficient to check the filter's DC-gain to identify lifted headphones. Once the lifting is identified, the adaptive filter is gradually changed to a stable default filter from which the leaky NFxLMS adaption starts again.

The algorithm has been tested for various use cases and it has been shown that it yields robust ANC when the headphones sit regularly tight and still preserves stability if the headphones are suddenly lifted. The performance is compared with approaches from literature and the presented approach yields the same ANC results with less computational complexity. The algorithm is based on experimental data of prototype headphones, but all determining factors have physical reason which give the algorithm a general validity.

Chapter 6

General Conclusion

This thesis discusses new approaches for feedforward as well as for feedback ANC headphones. The two main problems that are addressed are the latency in the secondary branch of ANC systems and the variations in the secondary path due to different fittings of the headphones.

Feedforward ANC headphones are powerful if adaptive filters are used because they allow reacting on changes in the sound field as well as on changes in the acoustics of the headphones. Adaptive filters can easily be implemented as digital transversal filters that are updated by the LMS algorithm. However, the conversion into the digital and back into the analogue domain is time consuming which deteriorates the ANC performance. Although analogue implementations of adaptive filters are theoretically possible [Velosoa and Nascimentob, 2005, Johns et al., 1991], their implementation is tricky because of the DC-bias in active analogue circuits [Johns et al., 1991, Shoval et al., 1995]. The effort to handle these offset increases with the number of updated filter coefficients. This makes an analogue implementation of transversal filters rather unsuitable because they generally require much more coefficients to model a given impulse response than IIR

filters.

In chapter 2 of this thesis, it is therefore suggested to use a set of parallel IIR filters and only adapt the gain of each filter output as an adaptive linear combiner (ALC). It is shown that three parallel IIR filters yield equally good results as a 12 taps long transversal filter at 11025 Hz sampling frequency. The ALC of the parallel IIR filters converges faster than the adaptive transversal filter and requires only three instead of twelve updates which clearly reduces the effort to handle the DC-bias.

The linear combination of the three IIR filters has to approximate all possible transfer functions $\frac{P(j\omega)}{S(j\omega)}$ where $P(j\omega)$ is the primary and $S(j\omega)$ is the secondary path. The main variations in $P(j\omega)$ result form different directions of the sound field and the main variations in $S(j\omega)$ result from differently tight fittings of the headphones. All these main variations are determined by preliminary measurements. A principal component analysis and the Gauss-Newton algorithm are then used to extract three IIR filters whose combination best explains the variations in $\frac{P(j\omega)}{S(j\omega)}$.

In [Carusone and Johns, 2003], it is shown that it is advantageous to use analogue filters, but a digital fxLMS update. This strategy can be employed to our approach, too. On the one hand, the three parallel IIRs can be implemented as analogue filters to avoid additional latency. On the other hand digital signal processing can be used for the fxLMS update where the latency only affects the speed of convergence. The benefit of digital signal processing it that it allows easily storing and analysing signals. The fxLMS e.g. requires a model of the secondary path that can be derived online with digital signal processing.

Chapter 3 intensifies the consideration on online secondary path detection. It presents two complementary methods that are more robust than the

online secondary path estimations which have been presented in literature so far. The first method analyses and compares the spectral density of the reference signal and the error signal. If the headphones sit regularly tight, ANC headphones are able to attenuate the low frequencies of the ambient noise by 10 dB. If the comparison detects less ANC, it means that sound power of the anti-noise radiates outside and that the headphones do not sit regularly tight. Then the secondary path model can be changed accordingly. However, this method only works properly if there is enough low frequency content in the ambient noise.

The second method can be used if there is hardly any low frequency noise, but some narrowband excitation. In this case the tightness of the secondary path can be identified with an infrasonic measurement. The infrasound signal is played back by the loudspeakers of the headphones and if it can be detected by the error microphone inside the ear-cups it means that the headphones sit regularly tight. If the infrasound test signal is not sensed by the error microphone it again means that sound power radiates outside and that the headphones do not sit regularly tight.

Chapter 4 treats the problem of the latency for feedback ANC headphones. If the ambient noise arrives at the error microphone earlier than the anti-noise, the ANC system needs to predict the ambient noise. Prediction is possible if the noise signal is band limited. It is therefore suggested to use the feedback structure because the noise which is sensed by the error microphone is naturally band limited due to the passive attenuation of the ear cups. For broadband ambient noise, the spectral characteristic of the penetrated noise is thus determined by the low-pass characteristic of the passive attenuation. Chapter 4 shows how a prediction filter can be designed based on this low pass characteristic. It is further shown that this static

prediction filter yields as good results as adaptive predictors for common ambient noises like aeroplane and engine noises, and that it is more robust against changes in the secondary path.

Chapter 5 on the other hand shows that an adaptive filter can be used to detect changes in the secondary path. As already mentioned, the low-frequency sound power of the headphones gets lost if the headphones do not sit regularly tight. It is shown that the adaptive filter tries to compensate this effect by amplifying the low frequencies and it is further demonstrated that it is sufficient to observe the DC gain of the adaptive filter to detect leaky secondary paths. The DC gain equals the sum of the filter coefficients and can therefore be evaluated very easily on a DSP. It is shown that this approach of detecting changes in the secondary path is as robust as common approaches from literature, and, at the same time, it is much less processor intensive. If a leaky secondary path is detected, the adaptive filter is changed to a stable default filter. In chapter 5, a scaled impulse is suggested as default filter, but it might be advantageous to use a prediction filter as in chapter 4.

Equally, the method to detect changes of $S(j\omega)$ in feedforward headphones which has been presented in chapter 3 can be applied to the adaptive linear combiner of parallel IIR filters (cf. chapter 2). It combines the efficient and (if analogue) quasi delayless parallel filters with a robustly stable fxLMS update. Thus, combinations of the presented feedforward approaches of chapter 2 and 3, as well as the feedback approaches of chapter 4 and 5 have a lot of potential for future work.

Furthermore, the work on the frequency dependent leaky LMS can be intensified. It has been shown that the frequency dependent leakage depends on the weighting filter and on the product of $\mu\gamma$. In chapter 5, the weighing

filter H was chosen to be a high shelf and γ has been derived empirically. In future work, an algorithmic approach should be developed to obtain the ideal weighting filter and to calculate the optimal γ online. Besides the frequency dependent leakage, also a tap dependent leakage could be useful. A larger leakage could for example be used for the later taps similar to a tapering window. This could be advantageous for the adaptive filter of chapter 5 where the short filter sharply cuts the ideal impulse response and causes side lobes in the frequency response of W.

The adaptive ANC filters as well as the fixed prediction filter have been tested on hardware with low latency. This hardware is self assembled and consists of a digital signal processor (DSP) and high speed data converters with a potential sampling frequency of over 500 kHz. Both, DSP and converters are assembled on evaluation boards outside of the headphones. This hardware set-up has been suitable to demonstrate the functionality of digital ANC, but it is far too big sized for being embedded into the ear cups of the headphones. This thesis and others (especially [Foudhaili, 2008]) have demonstrated the benefits of digital broadband ANC, but hardware that is suitable for commercial products still needs to be developed. Such hardware has to satisfy small size, low energy consumption and low costs.

Since these requirements are hard to fulfil, most commercial products still rely on analogue hardware. And although ANC headphones are already well established on the consumer market, their assessment is still a difficult task. The following appendix A deals with this problem. It is shown that different measurement systems, like artificial ears and mannequins, provoke differently leaky positioning, and as a consequence, they lead to different ANC results. It is further shown that even the ranking of different ANC headphones differs depending on the measurement systems. There-

fore, we conducted listening test and compare the subjective ratings with the measurement results of 4 different systems. Finally, a regression of the measurements is developed that robustly represents the user ratings. This result now allows for an easy and inexpensive comparison and assessment of ANC headphones disrespect of their method, be it feedforward or feedback, digital or analogue.

Appendix A

Assessment of Active Noise Cancelling Headphones

based on: M. Guldenschuh et al., 'Assessment of Active Noise Cancelling Headphones', 2012 International Conference on Consumer Electronics - Berlin, pp. 299-303, 3.-5. Sept. 2012

In the last decades, a lot of research on ANC headphones has been conducted [Bartels, 1991, Kuo and Morgan, 1999, Song et al., 2005]. Initially, ANC headphones were developed for air traffic professionals, but during the last few years they also spread on the consumer market. Although various different ANC headphones are commercially available by now, their performance assessment is still a difficult task. This counts for consumer-test and certification centres as well as for manufacturers themselves.

For ANC assessment, headphones are either put on ear simulators that contain a microphone which measures the residual noise inside the earcup or on real persons who have probe microphones inserted in their ear canal [Lancaster and Casali, 2004]. One problem of ANC assessment is that the

noise reduction strongly depends on the tightness of the wearing situation. Headphones sit differently tight on different ear simulators and therefore lead to different ANC results, and also probe microphones influence the wearing situation.

The question is thus, which measurement results are closest to the ANC performance in real life situations.

In this appendix, user ratings of five different ANC headphones are analysed and the ratings are compared with the measured noise reduction on four different ear simulators (including one measurement with probe microphones on a real person). This way, we can identify which measurement procedure on which simulator correlates strongest with the user ratings.

The following section describes the influence of leaky wearing situations on the ANC performance, section A.2 outlines the measurement procedures and the listening test and in section A.3 the regression of the measurements with the user ratings is derived.

A.1 Leakage Influence on ANC headphones

ANC headphones cancel noise that enters the headphone (via the so called primary path) with anti noise that is played back by the secondary path. (See also Fig. A.1.) The secondary path depends on the loudspeaker characteristics and on the acoustic impedance which is determined by the cavity between the headphone and the ear drum. While the loudspeaker characteristics are deterministic and approximately constant, the acoustic impedance depends on the tightness of the headphone and might vary from user to user.

Fig. A.2 shows the changes in the secondary path when small leaks of 1mm radius and 15mm length are introduced in an initially tight sitting
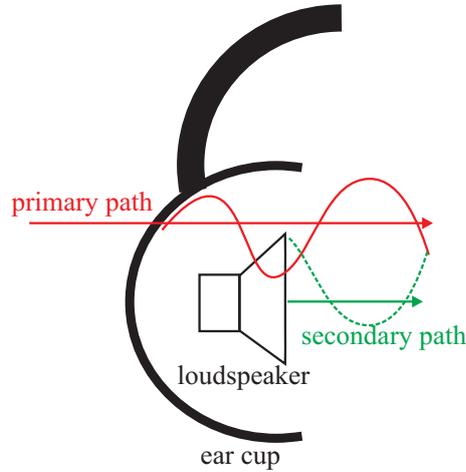
Figure A.1: Ambient noise penetrates the earcup of the headphone via the primary path and the loudspeakers plays back the 'anti' noise via the secondary path.

(a) Magnitude response
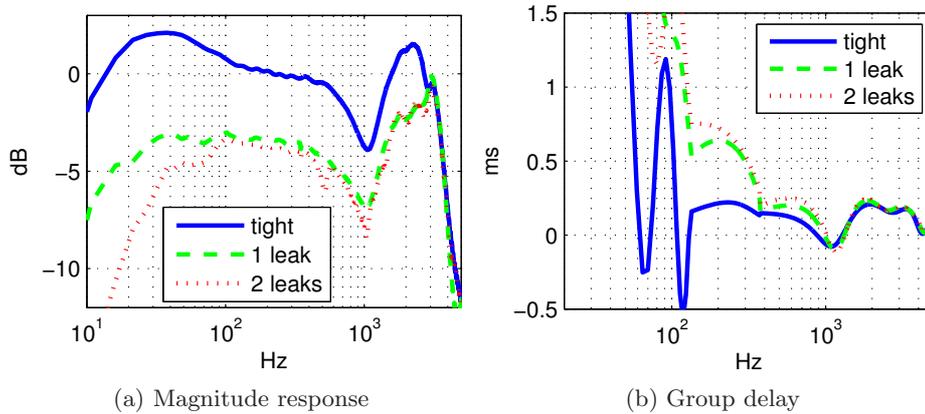
(b) Group delay

Figure A.2: Magnitude response and group delay of the transfer function from the loudspeaker to the ear canal (i.e. the secondary path) for three differently tight wearing situations: With increased leakage, the sound pressure level is diminished and the group delay is increased especially at low frequencies.

headphone. With increased leakage, the sound pressure level (SPL) declines and the group delay increases at low frequencies. Both consequences are disadvantageous for ANC. In leaky wearing situations, the headphone does not produce enough low frequency SPL to cancel the penetrating noise, and

the low frequency anti-noise will arrive too late because of the increased group delay.

Previous work [Snyder and Hansen, 1990,Lopes and Piedade, 2004] investigated the influence of the secondary path delay on the least mean squares (LMS) algorithm (which is often proposed for digital ANC applications), but not on the performance of analogue ANC headphones. Fig. A.3 shows that the leaks degrade the ANC performance by 10 dB.
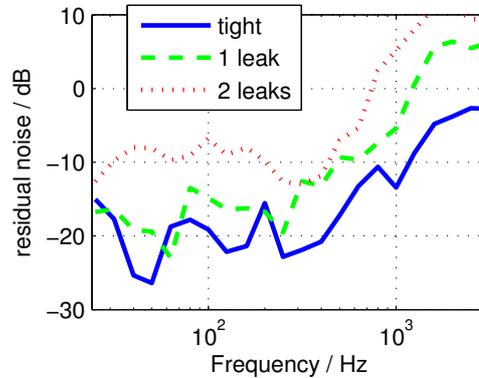
Figure A.3: Comparison of the ANC performance for a tight and leaky ANC headphone. The dB values of the residual noise are related to the penetrated noise without ANC. Under a tight wearing condition the ANC headphone reduces low frequency noise about 20 dB. With two leaks, the ANC amount is degraded to 10 dB.

A.2 Experimental Setup

A.2.1 Objective Measurements

Since ANC headphones do not sit equally tight on all ear simulators, different such simulators lead to different ANC measurement results. We measured the frequency response of the residual noise after ANC for five different headphones on the following four ear simulators:

1. A head and shoulder simulator with an abstract model of the outer

ear and microphones at the end of the ear canal. (Fig. A.4a)

2. A company specific acoustic measurement coupler similar to an artificial ear according to IEC 60318[1] without extra leakage.

3. A head and torso simulator with a detailed model of the outer ear and microphones at the end of the ear canal. (Fig. A.4b)

4. Probe microphones that are inserted 3 mm into the ear canals of a real person. (Fig. A.4c)

(a) A head and shoulder simulator with an abstract model of the outer ear and microphones at the end of the ear canal.

(b) A head and torso simulator with a detailed model of the outer ear and microphones at the end of the ear canal.

(c) Probe microphone inserted 3 mm into the ear canals of a test subject.
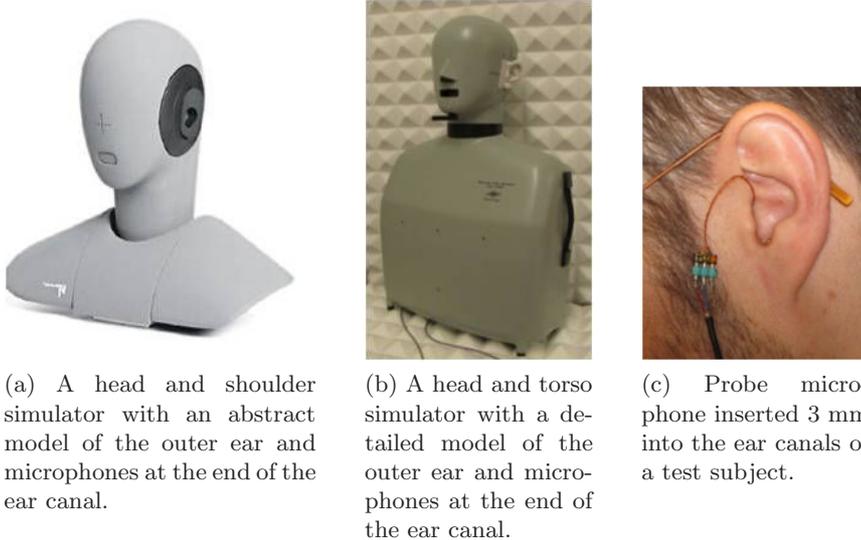
Figure A.4: Measurement tools.

In an semi-anechoic chamber, a loudspeaker played back a sine sweep from a distance of approx. 1m to the ear simulator. For each ear simulator, a reference measurement without ANC headphone is made. The noise reduction is determined by the relation between the measured residual noise when an ANC headphone is put on and the reference measurement without headphone. For each headphone, the noise reduction was measured

---

[1]International Electrotechnical Commission norm:
http://webstore.iec.ch/webstore/webstore.nsf/artnum/043309!opendocument

for 36 horizontal directions of incident sound. Fig. A.5 shows the median broadband power over all measurement directions per headphone and ear simulator.
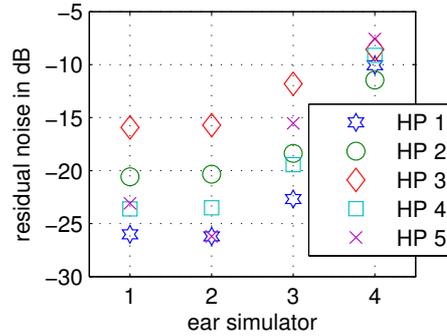

Figure A.5: Median active noise reduction over all measurement directions for all 5 headphones (HPs) on each of the 4 ear simulators. Ear simulators 1 and 2 measure lower SPLs than simulator 3 and 4. Furthermore, the ranking of the headphones changes per ear simulator. Headphone number 5 (HP 5) for example is ranked third best on ear simulator 1, best on simulator 2, second worst on simulator 3, and worst on simulator 4.

The measured attenuation for one and the same headphone can differ more than 15 dB on two different ear simulators. Even more problematic is that the ranking of the five headphones differs from ear simulator to ear simulator.

A.2.2 Subjective Evaluation

To answer the question, which measurement correlates most with user perception, we performed a listening test with 26 subjects (20 male and 6 female) and the five headphones under investigation. The subjects are all related to acoustics and sound engineering, either as students or as professionals. In a blind testing, the subjects were asked to rate the five headphones against each other on a scale which was labelled from 'very strong noise reduction' to 'very low noise reduction'. Three noises that represent

148

distinct sound characteristic and typical situations for ANC were played back from a 5.1 surround system:

- Aeroplane noise (broadband noise)

- Train noise (impulse noise)

- Speech noise (narrow-band noise)

A multifractional analysis of variances (ANOVA) of the subjects' answers however indicates that neither the sex of the subjects nor the type of noise has influence on the rating.

Fig. A.6 shows the distribution of the user ratings for each of the five headphones. As a comparison, Fig. A.7 shows the median broadband measurement results per headphone. It can be seen that the broadband mea-
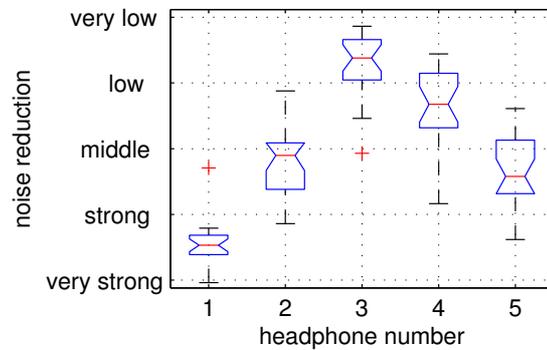
Figure A.6: Distribution of 26 user ratings about the noise reduction of each of the 5 headphones measured before. The boxes include 50% of the answers and the whiskers show the distribution of all 26 answers except for the outliers which are marked as crosses. The bars in the middle of the boxes denote the median values of the user' answers.

surement results match the subjective answers only partly. The idea is therefore to find a regression of measurement parameters that better reproduces the subjective answers.
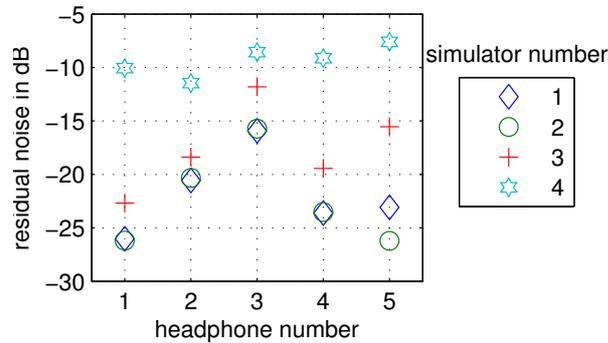
Figure A.7: Median broadband measurement result of the 5 headphones on the 4 ear simulators. A comparison to Fig. A.6 shows that none of the ear simulators delivers the same headphone ranking as the median subject's answers.

A.3 Regression between user ratings and measurement procedures

We determine which measurement procedure comes closest to the subjective perception by regressing the user ratings and the measurement results of each of the four ear simulators. The regression is done for different frequency bands, different measurement directions and psychoacoustically motivated measures like loudness. In detail, for each of the 4 ear simulators, the measurement signals are divided into

- 5 post-processing strategies

    - Logarithm of the measured signal power in 4 frequency bands:
        * Broadband: 20 -   20000   Hz
        * Low band: 20 -   100   Hz
        * Mid band 1: 100 -   700   Hz
        * Mid band 2: 700 -   3000   Hz
    - Loudness of the measured broadband signal in sone, calculated according to [Zwicker, 1990],

150

- and 6 median values over

  - all measurement directions

  - frontal measurement directions: -40° to 40°

  - left lateral measurement directions: 50° to 130°

  - dorsal measurement directions: 140° to 220°

  - right lateral measurement directions: -50° to -130°

  - measurement directions from surround positions: 0°, +/-30°, +/-
    110°,

  whereas the measurement variations for one headphone cover approx-
  imately 28% of the variations between different headphones.

In total, this leads to $J = 4 \times 5 \times 6 = 120$ regression variables $x_{1...J}$. In
a linear regression, the target $y_{k,i}$ (e.g. user judgement k about headphone
number i) is expressed as a weighted sum of these regression variables $x_{1...J,i}$

$$y_{k,i} = \alpha_0 + \alpha_1 x_{1,i} + \alpha_2 x_{2,i} + \cdots + \alpha_J x_{J,i}, \tag{A.1}$$

where the weights $\alpha_{0...J}$ are called regression coefficients. With $i = 1...5$
headphones and the answers of $k = 1...26$ test subjects, eq. (A.1) can be
replicated to a system of 5×26 equations which has to be solved for $\alpha_{0...J}$.
Finally, the objective measures of x_j and the weights α_j are used to derive a
judgement prediction $\tilde{y}$ about an arbitrary headphone analogue to eq. (A.1).
The difference between the prediction $\tilde{y}$ and the known judgement y is called
prediction- or regression error e.

This error is used to define two important quality measures for regres-
sions. The first is the coefficient of determination R^2 which indicates how

good the regression result fits the target. It relates the sum of the squared regression errors to the sum of squared target values

$$R^2 = 1 - \frac{\sum_{k,i} e_{i,k}^2}{\sum_{k,i} y_{i,k}^2}. \tag{A.2}$$

The second is Cook's Distance D which indicates whether a regression can be generalised to new test samples (e.g. new, yet unknown headphones) or not. This general validity is tested by omitting the data y_m and $x_{j,m}$ of headphone number m in the system of equations in eq. (A.1). Cook's distance D then relates the sum of squared differences between the prediction $\tilde{y}$ of the full regression and the prediction $\tilde{y}_{(m)}$ of the regression where headphone m was omitted to the sum of squared regression errors

$$D_m = \frac{\sum_{k,i}(\tilde{y}_{k,i} - \tilde{y}_{k,i,(m)})^2}{\sum_{k,i} e_{i,k}^2}. \tag{A.3}$$

A regression is considered to be robust and generalizable if $D < 1$.

We expect different regression variables to be reliable and hope that summing up those reliable variables will lead to a stronger correlation with the subjective answers, but the problem of a linear regression is that it also allows negative weights α. Instead of summing up the most reliable candidates, unreliable candidates are subtracted from each another. For the evaluated headphones, this leads to a very good fit of the user answers (i.e. a large R^2), but for any new headphone the regression is not reliable (i.e D is very large).

We therefore correlated each of the 120 regression variables separately with the user ratings and identified four variables with $R^2 > 0.7$. The first is the loudness calculated from frontal measurement results on ear simulator 1. Its maximum Cook's distance is $D = 0.3$, which makes it a very robust

regression variable. The other three variables are all related to the logarithm of the signal power between 100 and 700 Hz. We choose ear simulator 2 as a representative and robust candidate, as it's $D = 0.4$.

Both obtained post-processing strategies are very meaningful with respect to ANC. The bandwidth from 100 Hz to 700 Hz is where the active part of noise reduction happens and the calculated loudness corresponds to broadband human perception of SPLs.

Fig. A.8 compares the regression results of the two presented variables with the distribution of the user ratings. The chosen regression variables match the subjective answers clearly better than the simple broadband measures of Fig. A.7.
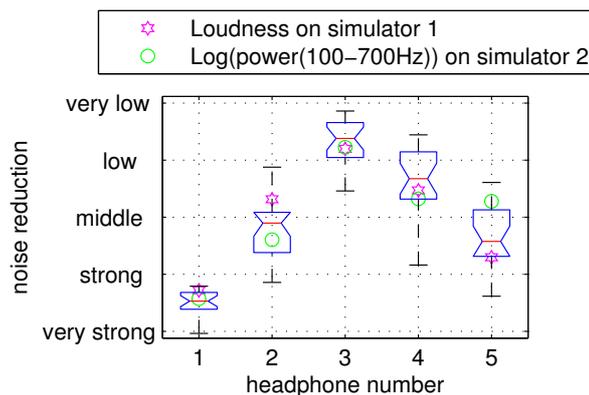
Figure A.8: Distribution of the user ratings and the regression based on the the loudness measured on simulator 1 and on the logarithm of the band-limited signal power measured on simulator 2.

Finally, we also identified the ear simulators with the strongest user correlation for the low band below 100 Hz and the mid band 2 between 700 and 3000 Hz. In the low band, ear simulator 4 has a R^2 of 0.46 and in the mid band 2 ear simulator 2 has $R^2 = 0.58$.

If one now wants to do a combined regression of variables in all different frequency bands and the sone value as broadband measure, the 4 presented

variables can be weighted according to their coefficients of determination R^2. Fig. A.9 shows the weighted combination of these 4 variables. As it can be seen, the combined regression delivers a very good reproduction of the user's headphone ranking. Its R^2 has increased to 76%. This is very much considering that the variations of the users' ratings per headphone already cover 10% to 25% of the judgement variations between all five headphones. Still, it is a robust approach because it is a constructive superposition of 4 reliable single-variable regressions.
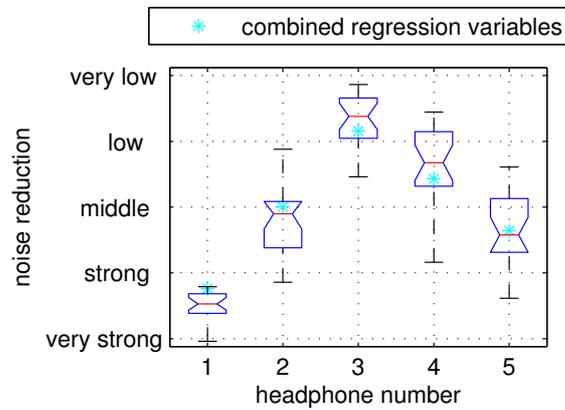
Figure A.9: Distribution of the user ratings and the regression with 4 combined regression variables covering the broadband measure in sone as well as the logarithm of the signal power in 3 different frequency bands. The coefficient of determination is increased to $R^2 = 0.76$.

A.4   Conclusion

ANC headphones can either be assessed via test persons (e.g through user ratings in a listening test or by audiometry) or by measurements on an ear simulator. The first method requires a sufficiently large number of test persons to get statistically significant results which makes the assessment costly and time consuming. Measurements on ear simulators consume far less resources, but we showed that different such ear simulators lead to

different measured attenuation which makes it doubtful if the drawn results are reliable.

The problem is that headphones sit differently tight on different ear simulators, and that the ANC performance strongly depends on the tightness of the wearing situation. The question is thus which measurement result is most trustful. Not necessarily the tightest one, because also on users, the headphones won't sit perfectly tight.

To find a more solid answer, we correlated the measured attenuation of five ANC headphones on four different ear simulators with user ratings of a blind listening test. A regression between the measured attenuation and the user ratings shall predict the user perception of other ANC headphones with similar mechanical characteristics. This way, standard measurement procedures can be used to get inexpensive and fast answers about the ANC performance. We determine single measurement procedures with strong correlation to user ratings and show how their combination leads to a simple and reliable assessment of different ANC headphones.

Bibliography

[Akhtar et al., 2006] Akhtar, M., Abe, M., and Kawamata, M. (2006). A new variable step size lms algorithm-based method for improved online secondary path modeling in active noise control systems. *Audio, Speech, and Language Processing, IEEE Transactions on*, 14(2):720 – 726.

[Ardekani and Abdulla, 2012] Ardekani, I. and Abdulla, W. (2012). Effects of imperfect secondary path modeling on adaptive active noise control systems. *Control Systems Technology, IEEE Transactions on*, 20(5):1252–1262.

[Bank, 2008] Bank, B. (2008). Perceptually motivated audio equalization using fixed-pole parallel second-order filters. *Signal Processing Letters, IEEE*, 15:477–480.

[Bartels, 1991] Bartels, V. (1991). Headset with active noise reduction system for mobile applications. In *Audio Engineering Society Convention 90*.

[Burgess, 1981] Burgess, J. C. (1981). Active adaptive sound control in a duct: A computer simulation. *The Journal of the Acoustical Society of America*, 70:715.

BIBLIOGRAPHY

[Cartes et al., 2002] Cartes, D. A., Ray, L. R., and Collier, R. D. (2002). Experimental evaluation of leaky least-mean-square algorithms for active noise reduction in communication headsets. *The Journal of the Acoustical Society of America*, 111(4):1758–1771.

[Carusone and Johns, 2003] Carusone, A. and Johns, D. (2003). Digital lms adaptation of analog filters without gradient information. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, 50(9):539–552.

[Dennis and Schnabel, 1996] Dennis, Jr., J. E. and Schnabel, R. B. (1996). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations (Classics in Applied Mathematics, 16)*. Soc for Industrial & Applied Math.

[Doyle et al., 1989] Doyle, J. C., Glover, K., Khargonekar, P. P., and Francis, B. A. (1989). State-space solutions to standard h 2 and h? control problems. *Automatic Control, IEEE Transactions on*, 34(8):831–847.

[Erdol and Basbug, 1996] Erdol, N. and Basbug, F. (1996). Wavelet transform based adaptive filters: analysis and new results. *Signal Processing, IEEE Transactions on*, 44(9):2163–2171.

[Farina, 2007] Farina, A. (2007). Advancements in impulse response measurements by sine sweeps. In *Audio Engineering Society Convention 122*.

[Foudhaili, 2008] Foudhaili, H. (2008). *Kombinierte Feedback- und adaptive Feedforward-Regelung fr die aktive Lrmreduktion in einem Kommunikations-Headset*. PhD thesis, Leibniz Universitt Hannover.

[Francis and Zames, 1984] Francis, B. A. and Zames, G. (1984). On h?-optimal sensitivity theory for siso feedback systems. *Automatic Control, IEEE Transactions on*, 29(1):9–16.

[Gan et al., 2005] Gan, W., Mitra, S., and Kuo, S. (2005). Adaptive feedback active noise control headset: implementation, evaluation and its extensions. *Consumer Electronics, IEEE Transactions on*, 51(3):975 – 982.

[Gholami-Boroujeny and Eshghi, 2010] Gholami-Boroujeny, S. and Eshghi, M. (2010). Online secondary path modeling in active noise control system using pbs-lms algorithm. In *Signal Processing (ICSP), 2010 IEEE 10th International Conference on*, pages 2600 –2603.

[Guldenschuh et al., 2013a] Guldenschuh, M., de Callafon, R., and Sontacchi, A. (2013a). Identification of secondary-path irregularities for active-noise-control headphones. In *Internoise 2013*.

[Guldenschuh et al., 2013b] Guldenschuh, M., Sontacchi, A., and Höldrich, R. (2013b). Headphones for active noise suppression. European Patent EP 2677765 A1, US Patent US 2013 0343557

[Guldenschuh et al., 2012] Guldenschuh, M., Sontacchi, A., Perkmann, M., and Opitz, M. (2012). Assessment of active noise cancelling headphones. In *Consumer Electronics - Berlin (ICCE-Berlin), 2012 IEEE International Conference on*, pages 299 –303.

[Hayes et al., 1980] Hayes, M., Lim, J., and Oppenheim, A. (1980). Signal reconstruction from phase or magnitude. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 28(6):672–680.

[Haykin, 2001] Haykin, S. (2001). *Adaptive Filter Theory (4th Edition)*. Prentice Hall.

[Jackson, 1991] Jackson, J. (1991). *A user's guide to principal components*. Wiley series in probability and mathematical statistics. Wiley, New York [u.a.].

[Jain and Ranganath, 1981] Jain, A. and Ranganath, S. (1981). Extrapolation algorithms for discrete signals with application in spectral estimation. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 29(4):830 – 845.

[Jin and Zhang, 2009] Jin, G. and Zhang, H. (2009). An adaptive active vibration and noise control system with online secondary path modeling. In *Measuring Technology and Mechatronics Automation, 2009. ICMTMA '09. International Conference on*, volume 1, pages 582–585.

[Johns et al., 1991] Johns, D., Snelgrove, W., and Sedra, A. S. (1991). Continuous-time lms adaptive recursive filters. *Circuits and Systems, IEEE Transactions on*, 38(7):769–778.

[Johns et al., 1989] Johns, D. A., Snelgrove, W. M., and Sedra, A. S. (1989). Orthonormal ladder filters. *Circuits and Systems, IEEE Transactions on*, 36(3):337–343.

[Kaelin et al., 1993] Kaelin, A., Lindgren, A., and Moschytz, G. (1993). Linear echo cancellation using optimized recursive prefiltering. In *Circuits and Systems, 1993., ISCAS '93, 1993 IEEE International Symposium on*, pages 463–466 vol.1.

[Kamenetsky and Widrow, Nov] Kamenetsky, M. and Widrow, B. (Nov.). A variable leaky lms adaptive algorithm. In *Signals, Systems and Comput-*

ers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on, volume 1, pages 125–128 Vol.1.

[Kannan et al., 2011] Kannan, G., Milani, A., Panahi, I., and Briggs, R. (2011). An efficient feedback active noise control algorithm based on reduced-order linear predictive modeling of fmri acoustic noise. *Biomedical Engineering, IEEE Transactions on*, 58(12):3303 –3309.

[Karjalainen and Paatero, 2006] Karjalainen, M. and Paatero, T. (2006). Equalization of loudspeaker and room responses using kautz filters: Direct least squares design. *EURASIP Journal on Advances in Signal Processing*, 2007.

[Kinney and Callafon, 2009] Kinney, C. and Callafon, R. (2009). Robust estimation for automatic controller tuning with application to active noise control. In Hof, P. M., Scherer, C., and Heuberger, P. S., editors, *Model-Based Control:*, pages 107–124. Springer US.

[Kinney et al., 2008] Kinney, C., Villalta, A., and de Callafon, R. (2008). Active noise control of a cooling fan in a short duct. In *Proc. ASME NoiseCon/NCAD*, pages 1–11.

[Kuo et al., 1996] Kuo, S., Kuo, S., and Morgan, D. (1996). *Active noise control systems: algorithms and DSP implementations*. Wiley series in telecommunications and signal processing. Wiley.

[Kuo and Morgan, 1999] Kuo, S. and Morgan, D. (1999). Active noise control: a tutorial review. *PIEEE*, 87(6):943–973.

[Kuo et al., 2006] Kuo, S. M., Mitra, S., and Gan, W.-S. (2006). Active noise control system for headphone applications. *IEEE Transactions on Control Systems Technology*, 14:331–335.

[Lancaster and Casali, 2004] Lancaster, J. A. and Casali, J. G. (2004). Reat and mire insertion loss comparison for eight headphones of various passive and anr desings, inclusive of nrr and spectral attenuation. Technical report, Auditory Systems Laboratory, Virginia Tech.

[Laroche, 2007] Laroche, J. (2007). Optimal constraint-based loop-shaping in the cepstral domain. *Signal Processing Letters, IEEE*, 14(4):225 –227.

[Liu et al., 2010] Liu, J., Xiao, Y., Sun, J., and Xu, L. (2010). Analysis of online secondary-path modeling with auxiliary noise scaled by residual noise signal. *Audio, Speech, and Language Processing, IEEE Transactions on*, 18(8):1978 –1993.

[Lopes and Piedade, 2004] Lopes, P. and Piedade, M. (2004). The behavior of the modified fx-lms algorithm with secondary path modeling errors. *Signal Processing Letters, IEEE*, 11(2):148 – 151.

[Morari and Zafiriou, 1989] Morari, M. and Zafiriou, E. (1989). *Robust Process Control*. Prentice Hall.

[Morgan and Thi, 1995] Morgan, D. and Thi, J. (1995). A delayless subband adaptive filter architecture. *Signal Processing, IEEE Transactions on*, 43(8):1819 –1830.

[Narayan et al., 1983] Narayan, S., Peterson, A., and Narasimha, M. (1983). Transform domain lms algorithm. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 31(3):609–615.

[Nelson et al., 1993] Nelson, P. A., Elliott, S. J., and Williams, J. E. F. (1993). Active control of sound. *Physics Today*, 46(1):75–76.

[Oppenheim et al., 1994] Oppenheim, A., Weinstein, E., Zangi, K., Feder, M., and Gauger, D. (1994). Single-sensor active noise cancellation. *Speech and Audio Processing, IEEE Transactions on*, 2(2):285 –290.

[Park et al., 2001] Park, S. J., Yun, J. H., Park, Y. C., and Youn, D. H. (2001). A delayless subband active noise control system for wideband noise control. *Speech and Audio Processing, IEEE Transactions on*, 9(8):892 – 899.

[Rafaely and Elliott, 1999] Rafaely, B. and Elliott, S. (1999). H_2/H_∞; active control of sound in a headrest: design and implementation. *Control Systems Technology, IEEE Transactions on*, 7(1):79 –84.

[Rafaely and Elliott, 2000] Rafaely, B. and Elliott, S. (2000). A computationally efficient frequency-domain lms algorithm with constraints on the adaptive filter. *Signal Processing, IEEE Transactions on*, 48(6):1649–1655.

[Rafaely and Jones, 2002] Rafaely, B. and Jones, M. (2002). Combined feedback-feedforward active noise-reducing headset - the effect of the acoustics on broadband performance. *Journal of The Acoustical Society of America*, 112.

[Rife and Vanderkooy, 1987] Rife, D. D. and Vanderkooy, J. (1987). Transfer-function measurement using maximum-length sequences. In *Audio Engineering Society Convention 83*.

[Schumacher et al., 2011] Schumacher, T., Krger, H., Jeub, M., Vary, P., and Beaugeant, C. (2011). Acive noise control in headsets: A new approach for broadband feedback ANC. In *Proceedings of IEEE Inter-*

national Conference on Acoustics, Speech, and Signal Processing, pages 5183–5203, Prague, Czech Republic.

[Shoval et al., 1995] Shoval, A., Johns, D., and Snelgrove, W. (1995). Comparison of dc offset effects in four lms adaptive algorithms. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, 42(3):176–185.

[Shynk, 1989] Shynk, J. (1989). Adaptive iir filtering. *ASSP Magazine, IEEE*, 6(2):4–21.

[Shynk, 1992] Shynk, J. (1992). Frequency-domain and multirate adaptive filtering. *Signal Processing Magazine, IEEE*, 9(1):14–37.

[Snyder and Hansen, 1990] Snyder, S. and Hansen, C. (1990). The influence of transducer transfer functions and acoustic time delays on the implementation of the lms algorithm in active noise control systems. *Journal of Sound and Vibration*, 141(3):409–424.

[Snyder and Hansen, 1994] Snyder, S. and Hansen, C. (1994). The effect of transfer function estimation errors on the filtered-x lms algorithm. *Signal Processing, IEEE Transactions on*, 42(4):950–953.

[Song et al., 2005] Song, Y., Gong, Y., and Kuo, S. (2005). A robust hybrid feedback active noise cancellation headset. *Speech and Audio Processing, IEEE Transactions on*, 13(4):607 – 617.

[Steiglitz and Mcbride, 1965] Steiglitz, K. and Mcbride, L. (1965). A technique for the identification of linear systems. *Automatic Control, IEEE Transactions on*, 10(4):461–464.

[Teschinegg, 2012] Teschinegg, G. (2012). Digitale aktive Geräuschunterdrückung mit geringer Latenz. Master's thesis, Institut für Elektronische Musik und Akustik, Univeristät für Musik und darstellende Kunst Graz.

[Vaidyanathan, 2007] Vaidyanathan, P. (2007). *The Theory of Linear Prediction*. Synthesis lectures on signal processing. Morgan & Claypool.

[Velosoa and Nascimentob, 2005] Velosoa, A. J. and Nascimentob, V. H. (2005). Headphone with active noise control using analog adaptive filters. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 2005, pages 204–211. Citeseer.

[Wang et al., 2013] Wang, L., Gan, W.-S., Khong, A., and Kuo, S. (2013). Convergence analysis of narrowband feedback active noise control system with imperfect secondary path estimation. *Audio, Speech, and Language Processing, IEEE Transactions on*, 21(11):2403–2411.

[Weinzierl et al., 2009] Weinzierl, S., Giese, A., and Lindau, A. (2009). Generalized multiple sweep measurement. In *Audio Engineering Society Convention 126*.

[Widrow et al., 1975] Widrow, B., Glover, J. R., Mccool, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, E., Goodlin, R. C., and Goodlin, R. C. (1975). Adaptive noise cancelling: Principles and applications. *Proceedings of the IEEE*, 63(12):1692–1716.

[Widrow et al., 1976] Widrow, B., McCool, J., Larimore, M., and Johnson, C. (1976). Stationary and nonstationary learning characteristics of the lms adaptive filter. *Proceedings of the IEEE*, 64(8):1151–1162.

BIBLIOGRAPHY

[Widrow and Walach, 1984] Widrow, B. and Walach, E. (1984). On the statistical efficiency of the lms algorithm with nonstationary inputs. *Information Theory, IEEE Transactions on*, 30(2):211–221.

[Widrow et al., 1981] Widrow, M., Shur, D., and Shaffer, S. (1981). On adaptive inverse control. In *Proc. 15th Asilomar Conf.*, pages 185–189.

[Williamson and Zimmermann, 1996] Williamson, G. and Zimmermann, S. (1996). Globally convergent adaptive iir filters based on fixed pole locations. *Signal Processing, IEEE Transactions on*, 44(6):1418–1427.

[Xiang and Schroeder, 2003] Xiang, N. and Schroeder, M. R. (2003). Reciprocal maximum-length sequence pairs for acoustical dual source measurements. *J Acoust Soc Am*, 113(5):2754–61.

[Yu and Hu, 2001] Yu, S.-H. and Hu, J.-S. (2001). Controller design for active noise cancellation headphones using experimental raw data. *Mechatronics, IEEE/ASME Transactions on*, 6(4):483–490.

[Zangi, 1993] Zangi, K. (1993). A new two-sensor active noise cancellation algorithm. In *Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993 IEEE International Conference on*, volume 2, pages 351 –354 vol.2.

[Zeng and de Callafon, 2003] Zeng, J. and de Callafon, R. (2003). Feedforward noise cancellation in an airduct using generalized fir filter estimation. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 6, pages 6392–6397 Vol.6.

[Zhang et al., 2013] Zhang, L., Wu, L., and Qiu, X. (2013). An intuitive approach for feedback active noise controller design. *Applied Acoustics*, 74(1):160 – 168.

[Zhang et al., 2001] Zhang, M., Lan, H., and Ser, W. (2001). Cross-updated active noise control system with online secondary path modeling. *Speech and Audio Processing, IEEE Transactions on*, 9(5):598–602.

[Zhang et al., 2003] Zhang, M., Lan, H., and Ser, W. (2003). A robust online secondary path modeling method with auxiliary noise power scheduling strategy and norm constraint manipulation. *Speech and Audio Processing, IEEE Transactions on*, 11(1):45 – 53.

[Zwicker, 1990] Zwicker, E. (1990). *Psychoacoustics; Facts and models.* Berlin [u.a.].