Analysis, Design, and Evaluation of Acoustic Feedback Cancellation Systems for Hearing Aids

- A Novel Approach to Unbiased Feedback Cancellation

Guo, Meng

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

Analysis, Design, and Evaluation of Acoustic Feedback Cancellation Systems for Hearing Aids

– A Novel Approach to Unbiased Feedback Estimation

Ph.D. Thesis

Meng Guo

Aalborg University
Department of Electronic Systems
Niels Jernes Vej 12
DK-9220 Aalborg, Denmark

Guo, Meng
Analysis, Design, and Evaluation of Acoustic Feedback Cancellation Systems for Hearing Aids – A Novel Approach to Unbiased Feedback Estimation

ISBN 978-87-7152-001-9

Department of Electronic Systems
Aalborg University
Niels Jernes Vej 12
DK-9220 Aalborg, Denmark

This thesis has been prepared using LaTeX. Figures have been created using Ti*k*Z/PGF.
Simulations have been run in Matlab.

Abstract

Acoustic feedback problems occur when the output loudspeaker signal of an audio system is partly returned to the input microphone via an acoustic coupling through the air. This problem often causes significant performance degradations in applications such as public address systems and hearing aids. In the worst case, the audio system becomes unstable and howling occurs.

In this work, first we analyze a general multiple microphone audio processing system, where a cancellation system using adaptive filters is used to cancel the effect of acoustic feedback. We introduce and derive an accurate approximation of a frequency domain measure—the power transfer function—and show how it can be used to predict the convergence rate, system stability bound, and the steady-state behavior of the entire cancellation system across time and frequency without knowing the true acoustic feedback paths. This power transfer function method is also applicable to an acoustic echo cancellation system with a similar structure.

Furthermore, we consider the biased estimation problem, which is one of the most challenging problems for state-of-the-art acoustic feedback cancellation systems. A commonly known approach to deal with the biased estimation problem is adding a probe noise signal to the loudspeaker signal and base the estimation on that. This approach is particularly promising, since it can be shown that, in theory, the biased estimation problem can be completely eliminated. However, we analyze a traditional probe noise approach and conclude that it can not work in most acoustic feedback cancellation systems in practice, due to the very low convergence rate of the adaptive cancellation system when using low level and inaudible probe noise signals.

We propose a novel probe noise approach to solve the biased estimation problem in acoustic feedback cancellation for hearing aids. It utilizes a probe noise signal which is generated with a specific characteristic so that it can facilitate an unbiased adaptive filter estimation with fast tracking of feedback path variations/changes despite its low signal level. We show in a hearing aid application that whereas the traditional and state-of-the-art acoustic feedback cancellation systems fail with significant sound distortions and howling as consequences, the new probe noise approach is able to remove feedback artifacts caused by the feedback path change in no more than a few hundred milliseconds.

Resumé

Akustiske tilbagekoblingsproblemer opstår, når udgangshøjttalersignalet af et lydanlæg vender delvist tilbage til indgangsmikrofonen via en akustisk kobling gennem luften. Dette problem forårsager ofte en betydelig nedsættelse af ydeevnen i applikationer såsom højttaleranlæg og høreapparater. I værste fald bliver lydsystemet ustabilt og hyl opstår.

I dette arbejde analyserer vi først et generelt lydanlæg med flere mikrofoner, hvor et annulleringssystem ved hjælp af adaptive filtre er anvendt til at ophæve akustisk tilbagekoblingsvirkning. Vi introducerer og udleder en præcis tilnærmelse af et frekvensdomæne mål—power transfer function—og viser hvordan det kan bruges til at forudsige konvergenshastighed, system stabilitetsgrænse, og ligevægtstilstandsopførsel af hele annulleringssystemet over tid og frekvens uden at kende de rigtige akustiske tilbagekoblingsveje. Denne power transfer function barserede metode kan også anvendes til et akustisk ekkoannulleringssystem med en lignende struktur.

Desuden arbejder vi med et af de mest udfordrende problemer for de bedste/nyeste akustiske tilbagekoblingsannulleringssystemer, nemlig problemet med det ikke-centrale estimat. En kendt metode til at behandle dette problem på er at tilføje et probestøjsignal til højttalersignalet og derefter baserer estimeringen på dette. Denne metode virker generelt godt, da det kan påvises, at det i teorien giver et centralt estimat. Men vi analyserer en traditionel probestøjsmetode og konkluderer, at den ikke kan fungere i de fleste akustiske tilbagekoblingsannulleringssystemer i praksis som følge af den meget lave konvergensrate af det adaptive annulleringssystem, når der anvendes et ikke hørbart probestøjsignal med lavt niveau.

Vi foreslår en ny probestøjsmetode til at løse problemet med det ikke-centrale estimat i akustisk tilbagekoblingsannullering til høreapparater. Her anvendes et probestøjsignal, som genereres med en bestemt egenskab, så det resulterer i et centralt adaptivt filter estimat med hurtig sporing af tilbagekoblingsvejes variationer/ændringer til trods for sit lave signalniveau. Vi viser i en høreapparatsimulering, at når de traditionelle og de nyeste akustisk tilbagekoblingsannulleringssystemer svigter og betydelige forvrængninger af lyden og hyl derfor opstår, er den nye probestøjsmetode i stand til at fjerne tilbagekoblingsartefakter forårsaget af en ændring i tilbagekoblingsvejen inden for få hundrede millisekunder.

Contents

Preface

This thesis is submitted to the Doctoral School of Engineering and Science at Aalborg University, Denmark, in a partial fulfilment of the requirements for the Ph.D. degree.

The work was carried out as an industrial Ph.D. project, in the period from March 2010 to December 2012, jointly at the Department of Electronic Systems, Aalborg University, and Oticon A/S, Denmark. Some of the work was conducted in collaboration with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology in Rolla, Missouri, USA.

The work deals with acoustic feedback and echo cancellation techniques. The main focus is on analysis of a general acoustic feedback and echo cancellation system, and on design and evaluation of a novel acoustic feedback cancellation system for hearing aids.

This thesis consists of two parts, the introduction and the main body. In the introduction part, we provide an overview of the acoustic feedback problem. More specifically, we present the background, state-of-the-art solutions, and the remaining open problems in the area of acoustic feedback control, before we introduce a novel acoustic feedback cancellation system. The main body of the thesis consists of seven research papers which have been published or accepted to be published in peer-reviewed journals or conferences. The main contributions in this thesis are primarily in these papers.

List of Papers

The main body of this thesis consists of the following papers.

[A] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of Acoustic Feedback/Echo Cancellation in Multiple-Microphone and Single-Loudspeaker Systems Using a Power Transfer Function Method," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5774–5788, Dec. 2011.

[B] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Acoustic Feedback and Echo Cancellation Strategies for Multiple-Microphone and Single-Loudspeaker Systems," in *Proc. 45th Asilomar Conf. Signals, Syst., Comput.*, Nov. 2011, pp. 556–560.

[C] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "On Acoustic Feedback Cancellation Using Probe Noise in Multiple-Microphone and Single-Loudspeaker Systems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 283–286, May 2012.

[D] M. Guo, S. H. Jensen, and J. Jensen, "Novel Acoustic Feedback Cancellation Approaches in Hearing Aid Applications Using Probe Noise and Probe Noise Enhancement," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2549–2563, Nov. 2012.

[E] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "On the Use of a Phase Modulation Method for Decorrelation in Acoustic Feedback Cancellation," in *Proc. 20th European Signal Process. Conf.*, Aug. 2012, pp. 2000–2004.

[F] M. Guo, S. H. Jensen, and J. Jensen, "Evaluation of State-of-the-Art Acoustic Feedback Cancellation Systems for Hearing Aids," *J. Audio Eng. Soc.*, to be published, 2013.

[G] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "Analysis of Closed-Loop Acoustic Feedback Cancellation Systems," in *Proc. 2013 IEEE Int. Conf. Acoust., Speech, Signal Process.*, to be published, May 2013.

In addition to the main papers, the following conference publications and technical reports have also been made.

[H] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of Adaptive Feedback and Echo Cancelation Algorithms in a General Multiple-Microphone and Single-Loudspeaker System," in *Proc. 2011 IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 433–436.

[I] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Comparison of Multiple-Microphone and Single-Loudspeaker Adaptive Feedback/Echo Cancellation Systems," in *Proc. 19th European Signal Process. Conf.*, Aug. 2011, pp. 1279–1283.

[J] M. Guo, S. H. Jensen, and J. Jensen, "An Improved Probe Noise Approach for Acoustic Feedback Cancellation," in *Proc. 7th IEEE Sens. Array Multichannel Signal Process. Workshop*, Jun. 2012, pp. 497–500.

[K] M. Guo, S. H. Jensen, and J. Jensen, "A New Probe Noise Approach for Acoustic Feedback Cancellation in Hearing Aids," in *Conf. Abstract 2012 Int. Hearing Aid Research Conf.*, Aug. 2012, pp. 61.

[L] M. Guo, "Frequency Domain Tracking Characteristics for Time and DFT Domain NLMS Algorithm in a Single Loudspeaker and Multiple Microphone Hearing Aid System," in *Technical Report, Oticon A/S*, 2010.

[M] M. Guo, "Analysis of Acoustic Feedback/Echo Cancellation Algorithms in Multiple-Microphone and Single-Loudspeaker Systems Using Probe Noise and Enhancement Filters," in *Technical Report, Oticon A/S*, 2011.

[N] M. Guo, "On Power Transfer Function in Closed-Loop Acoustic Feedback Cancellation Systems," in *Technical Report, Oticon A/S*, 2012.

Furthermore, the following patents have been filed.

[O] M. Guo, J. Jensen, and T. B. Elmedyb, "A Method of Determining Parameters in an Adaptive Audio Processing Algorithm and an Audio Processing System," European Patent Application, EP 2439958 A1, 2010.

[P] J. Jensen and M. Guo, "Control of an Adaptive Feedback Cancellation System Based on Probe Signal Injection," European Patent Application, EP 11181909, 2011.

Acknowledgment

First, I would like to express my sincere appreciation to my company advisor, Dr. Jesper Jensen, Oticon A/S. He contributed greatly to my work through our frequent, lengthy, and very fruitful technical discussions. I also appreciate his trust in my independent research and his thorough reviewing and proof reading of my work. Secondly, I direct my thankfulness to my academic advisor, Prof. Søren Holdt Jensen, for giving me helpful guidance, the freedom to choose the research directions of my interest, and for supporting me. I would also like to thank Mr. Thomas Bo Elmedyb for our inspiring discussions and his valuable comments on my work. Without their help, I would not be able to achieve the results presented in this thesis so quickly and they certainly deserve my gratitude.

I would like to thank my colleagues at Oticon A/S as well as my fellow Ph.D. students and colleagues at the Multimedia Information and Signal Processing group at the Department of Electronic Systems, Aalborg University, for interesting conversations, discussions, and support. It has been a pleasure to work with all these nice people.

A special thanks goes to Prof. Steven L. Grant for inviting me, as a visiting researcher, to the Department of Electrical and Computer Engineering, Missouri University of Science and Technology (MST) in Rolla, Missouri, USA, from November 2011 to May 2012. I am grateful for our fruitful discussions on acoustic echo and feedback cancellation techniques including various decorrelation methods from which I have benefited tremendously. I really appreciate his great hospitality too. I would also like to thank the people at the MST for making my visit very pleasant.

I thank Oticon A/S, the Oticon Foundation, and the Danish Agency for Science, Technology and Innovation for making this industrial cooperation possible. I also thank all the other people who have assisted me through the project.

Last but not least, I would like to thank my friends and family for their love and support.

Meng Guo
Copenhagen, December 2012

Introduction

1 The Acoustic Feedback Problem

1.1 Acoustic Feedback

Acoustic feedback is an electroacoustic phenomenon which occurs in audio reinforcement systems, such as public address systems and hearing aids. It is also referred to as audio feedback or simply feedback, or the Larsen effect after the Danish physicist Søren Absalon Larsen (1871–1957), who is cited [199, 204] to be first to discover the properties of acoustic feedback described in his work [113]. Acoustic feedback problems occur when microphones of an audio system pick up its output loudspeaker signals, so that an acoustic loop is created. Generally speaking, feedback problems can occur when the signal travelling around this acoustic loop gets stronger for each round trip.

The acoustic feedback problems often cause significant performance degradations in audio systems; in the worst case, the systems become unstable and howling occurs. Although the howling effect due to feedback can be intentionally used to generate a pure tone [113] and certain audio effects in popular music, and feedback is often used on purpose in control systems to improve the performance [143, 176], its appearance is generally undesired in audio reinforcement systems. Fig. 1 illustrates the principle of acoustic feedback in a single microphone system; all signals are real-valued, and we denote all signals as discrete-time signals with time index n for convenience.

In Fig. 1, $x(n)$ denotes the desired incoming signal to an audio processing system consisting of a microphone, a loudspeaker, and the forward path impulse response $\mathbf{f}(n)$ which represents the signal processing applied to the microphone signal $y(n)$ to create the loudspeaker signal $u(n)$; $\mathbf{h}(n)$ denotes the impulse response of the acoustic feedback path from the loudspeaker to the microphone of the audio system, and the microphone signal $y(n)$ consists of the desired incoming signal $x(n)$ and the undesired but unavoidable feedback signal $v(n)$. The presence of the feedback path $\mathbf{h}(n)$ can cause a stability problem in the audio system, and it can significantly degrade sound quality of loudspeaker signal $u(n)$, see e.g. [199] and the references therein.

A measure to determine stability in a linear and time-invariant closed-loop system

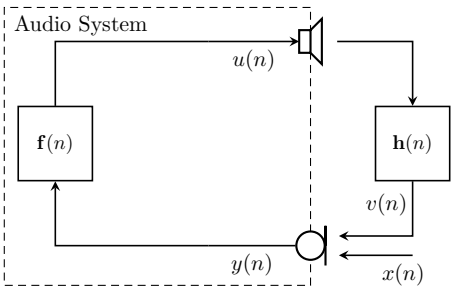


Fig. 1: A single microphone audio system with a feedback path $\mathbf{h}(n)$ from its loudspeaker to microphone. All signals are real-valued, and we denote all signals as discrete-time signals with time index n for convenience.

is the open-loop transfer function $\Theta(\omega, n)$ across discrete frequency ω and time n; for the system shown in Fig. 1, the open-loop transfer function $\Theta(\omega, n)$ is defined as

$$\Theta(\omega, n) = F(\omega, n)H(\omega, n), \tag{1}$$

where $F(\omega, n)$ and $H(\omega, n)$ are the frequency responses of $\mathbf{f}(n)$ and $\mathbf{h}(n)$, respectively. The Nyquist stability criterion [144, 203] states that a linear and time-invariant closed-loop system becomes unstable whenever the following two conditions are both fulfilled:

$$|\Theta(\omega, n)| \geq 1, \tag{2}$$
$$\angle\Theta(\omega, n) = l2\pi, \ l = \mathbb{Z}. \tag{3}$$

That is, the magnitude of the signal travelling around the loop does not decrease for each round trip, and the feedback signal adds up in phase to the microphone signal.

The main functionality of $F(\omega, n)$ in an audio reinforcement system is to amplify sound signals; $|F(\omega, n)|$ typically has a value larger than one for a wide range of ω. Hence, there is a potential risk to violate the condition stated in Eq. (2), and system instability would then occur at the frequencies ω for which the condition stated in Eq. (3) is fulfilled.

Furthermore, even for a stable audio system, feedback can significantly degrade the sound quality of the loudspeaker signal $u(n)$. Let $X(\omega, n)$ and $U(\omega, n)$ denote the spectra of incoming and loudspeaker signals $x(n)$ and $u(n)$, respectively. It can be shown that the magnitude of the input-output transfer function $\Gamma(\omega, n)$ from microphone to loudspeaker of the audio system shown in Fig. 1 is determined by

$$|\Gamma(\omega, n)| = \frac{|U(\omega, n)|}{|X(\omega, n)|} = \frac{|F(\omega, n)|}{|1 - \Theta(\omega, n)|}. \tag{4}$$

In the ideal situation for a system without feedback, i.e. $H(\omega, n) = 0$, we find from Eqs. (1) and (4) that,

$$|\Gamma(\omega, n)| = |F(\omega, n)|, \tag{5}$$

and

$$|U(\omega, n)| = |F(\omega, n)| \cdot |X(\omega, n)|. \tag{6}$$

Thus, the magnitude of the loudspeaker signal spectrum $|U(\omega, n)|$ is desirably obtained as the magnitude of the incoming signal spectrum $|X(\omega, n)|$ shaped by the forward path magnitude function $|F(\omega, n)|$. Otherwise, even for a stable system with feedback, i.e. $0 < |\Theta(\omega, n)| < 1$, undesired modifications of the loudspeaker signal may be introduced. In the limit as $|\Theta(\omega, n)| \to 1$, we get

$$|\Gamma(\omega, n)| \to \begin{cases} \infty \\ \frac{|F(\omega,n)|}{2} \end{cases} \quad \text{for} \quad \begin{array}{l} |\Theta(\omega, n)| \to 1, \text{ and } \angle\Theta(\omega, n) = l2\pi, \ l = \mathbb{Z}. \\ |\Theta(\omega, n)| \to 1, \text{ and } \angle\Theta(\omega, n) = \pi + l2\pi, \ l = \mathbb{Z}. \end{array} \tag{7}$$

This corresponds to an undesired shaping of the loudspeaker signal $u(n)$ depending on the values of $\Theta(\omega, n)$ across frequencies ω. In practice, this undesired signal shaping due to $|\Gamma(\omega, n)| \neq |F(\omega, n)|$ might lead to a significant sound distortion in loudspeaker signal $u(n)$.

Traditionally, studies have focused on controlling the effects of acoustic feedback in public address systems [17, 18, 28, 167]; more attention was specifically paid to the hearing aid application during the recent years, see the example studies in [24, 58, 69, 96, 127, 183]. In this work, we mainly focus on the acoustic feedback problem in a hearing aid application. However, as we will discuss in Sec. 5, our theoretical work is general and may find applications in other areas than hearing aid systems.

1.2 Hearing Aid Systems

In this section, we briefly describe the goal of a hearing aid, its most important functions and limitations.

Deficits in the human auditory system lead to different types of hearing impairments [132, 152]. The topic of hearing impairment and compensation is beyond the scope of this work, therefore, we refer to [34, 101, 165] for details. However, for many people, a hearing impairment can be a major barrier in their everyday life. It might reduce the ability to communicate with other people leading to social isolation, it might induce safety problems should a person not be able to hear alarms or understand security instructions, it might also delay or even affect the language development and learning of children, etc.

There exist different ways for helping people with hearing impairments. One well-known and probably the most commonly used method is by means of a hearing aid. A

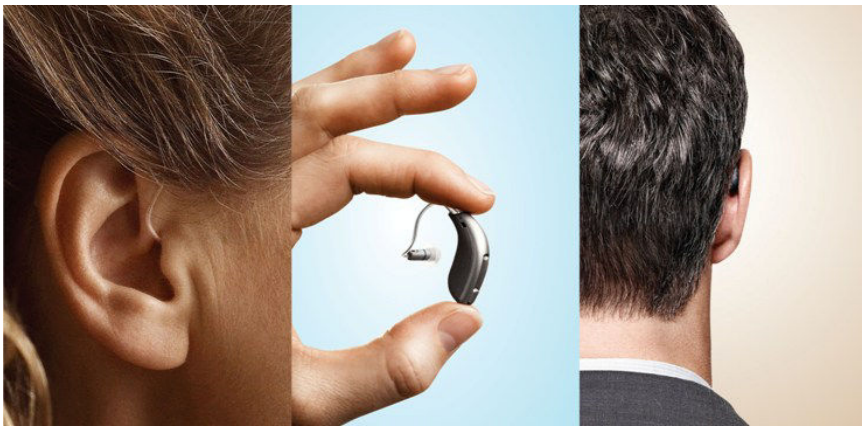


Fig. 2: An example of a modern behind-the-ear hearing aid.

hearing aid is a small electroacoustic device mainly used to amplify a sound signal for a user with hearing impairments [34, 101, 165]. A modern hearing aid is small in size, and it typically fits behind the ear or even in the ear canal of its user. Fig. 2 shows an example of it.

A modern hearing aid typically has many functions, the far most important one is to provide an amplified sound signal to its user; the technique used is typically referred to as compression [34, 101, 165]. Another important function in a hearing aid is noise reduction, which is used to increase the desired signal to noise signal ratio and thereby reduce listening effort, and improve sound quality and speech intelligibility. Other functions make the use of the hearing aid more convenient, such as automatic program selections for ensuring that a hearing aid is always working appropriately in different environments, the ability to connect to other electronic devices such as mobile phones or televisions to improve the sound quality in these situations, to mention a few examples.

Unfortunately, there are also some drawbacks by using a hearing aid. Probably the most significant one is the already mentioned acoustic feedback problem. If not properly treated, the feedback problem limits the maximum available amplification/gain in hearing aids due to the stability and sound quality degradation reasons discussed above, and a user would not benefit sufficiently from his/her hearing aid. It was e.g. shown in [107, 108], that the feedback problem is still one of the main factors causing hearing aid user dissatisfaction. Another cause for feedback problems in hearing aids is the mechanical coupling between loudspeakers and microphones. In this work, we do not focus on the mechanical feedback problem, since it is already largely reduced in hearing aids after the introduction of modern electret microphones which are less sensitive to mechanical vibration [2].

1.3 Acoustic Feedback in Hearing Aids

The acoustic feedback problem in hearing aids is difficult to avoid, since the hearing aid microphone should ideally be placed next to ear drum in order to preserve recorded sound signal as it would have been perceived naturally by the user. Furthermore, the output signal of the hearing aid loudspeaker[1] must be presented to the ear drum of the user. In practice, although compromises are made with respect to microphone and loudspeaker locations, the microphone would normally pick up the loudspeaker signal as a feedback signal.

In the following, we discuss some characteristics of feedback paths for hearing aids. In many hearing aid styles, an ear mold or a hearing aid shell is placed in the user's ear canal, and there are typically acoustic leaks between these and the ear canal. Furthermore, to reduce the unpleasant feeling of a closed ear canal which is also referred to as occlusion effect [34], a ventilation channel referred to as vent is typically used on the ear mold or the hearing aid shell, or a so-called open dome solution is applied which allows the ear to remain partially open. The impulse response of feedback paths in a hearing aid application is mainly determined by the acoustic leaks, the vent size, the open dome, the effects of the pinna, combined with microphones and loudspeakers including their amplification circuits, see more details in [34, 94, 101, 154].

Fig. 3 illustrates different characteristics of hearing aid feedback paths. In general, the impulse response $\mathbf{h}(n)$ of a feedback path is short in duration, typically in the order of a few milliseconds, especially when compared to the feedback paths of public address systems, in which the length of the impulse response could easily be hundreds and even thousands of milliseconds depending on the room acoustics. Fig. 3(a) illustrates two example impulse responses of feedback paths measured from the same user using two hearing aid styles (In-the-ear and Behind-the-ear), at a sampling rate of 20 kHz. In both cases, the numerical values of taps above time index 50 are close to zero, indicating that the effective duration of the impulse responses is roughly 2.5 ms.

We also observe from Fig. 3(a), that different hearing aid styles contribute greatly to variations in feedback path impulse responses, even for the same user. In different hearing aid styles, microphones and loudspeakers are placed at different positions; they are typically placed either both behind the ear, both in the ear canal, or microphones behind the ear and loudspeakers in the ear canal.

Moreover, the feedback paths are time-varying, especially when the hearing aid user is eating, chewing, or yawning [81, 99]. Other situations might cause an almost momentary change of feedback paths, such as when a user puts on a hat or places a telephone next to his/her ear [81, 156]. Fig. 3(b) shows magnitude responses $|H(\omega)|$ of two feedback path measurements for the same user wearing the same hearing aid with and without a telephone placed next to his ear. In this example, the difference is more than

[1] In hearing aid terminologies, the loudspeaker is generally referred to as the receiver. However, we refer it to as the loudspeaker in this work.

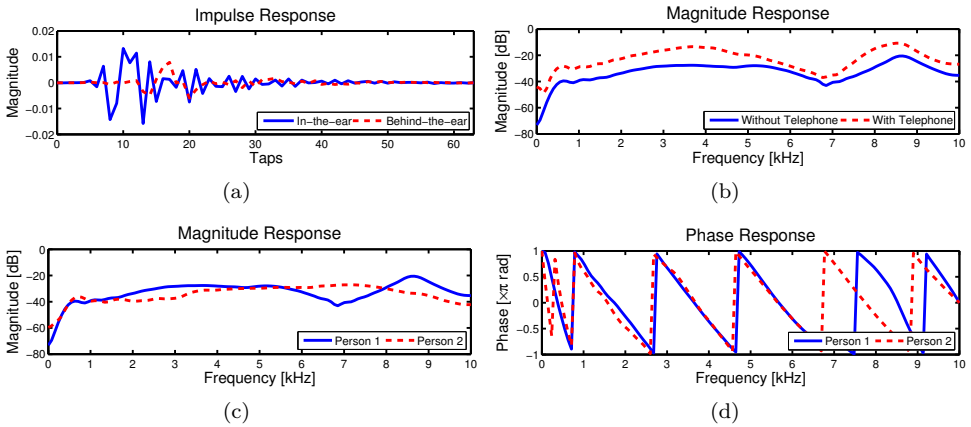


Fig. 3: Measured acoustic feedback paths at a sampling rate of 20 kHz. (a) Impulse responses for two hearing aid styles measured on the same user. (b) Magnitude responses with and without a telephone placed next to the ear of a hearing aid user. (c)-(d) Magnitude and phase responses of feedback paths for two hearing aid users.

15 dB at some frequencies, and it can be even more than 25 dB in some cases [81].

Another interesting characteristic of feedback paths in hearing aids is that the feedback problem is more probable to occur at higher frequencies above typically 3-4 kHz [81], due to the hearing aid styles and the surrounding geometry of hearing aids, see e.g. the magnitude responses in Fig. 3(b). Unfortunately, the desired amplification in the hearing aid forward path $\mathbf{f}(n)$ is often higher at high frequencies [34], making the feedback problem even more probable to occur.

Furthermore, the acoustic feedback in a hearing aid application depends highly on each individual user. Figs. 3(c) and 3(d) show magnitude and phase responses of feedback paths measured on two different users wearing the same hearing aid. The frequency regions with high risk of feedback are different for these two users; they are around 8-9 kHz for the first user and 6-8 kHz for the second user, respectively.

Moreover, hearing aid loudspeakers and microphones are essentially nonlinear devices [101], which become part of the acoustic feedback paths; this makes the feedback control even more challenging. However, the nonlinearity can often be modeled and compensated as e.g. discussed in [44, 95]. In this work, we do not pay attention to the nonlinearity in acoustic feedback paths.

All these specific characteristics and variations in acoustic feedback paths make the feedback control in hearing aids unique and difficult to solve.

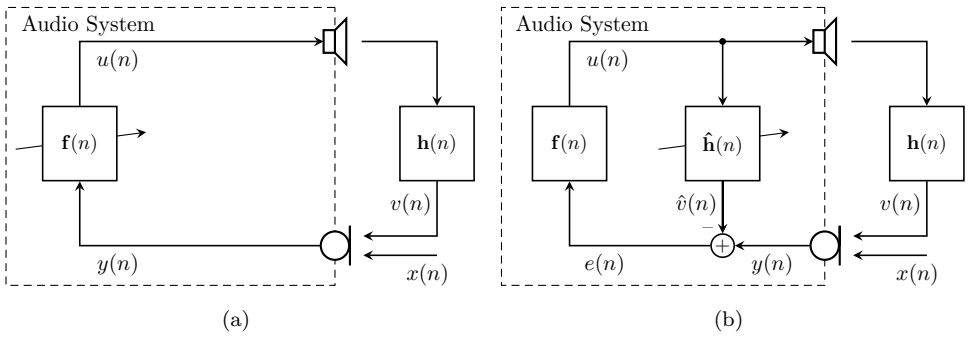


Fig. 4: A single channel acoustic feedback control system. The arrows through blocks indicate modifications made specifically for feedback control. (a) Feedforward suppression by modifying $\mathbf{f}(n)$. (b) Feedback cancellation by estimating $\hat{\mathbf{h}}(n)$.

2 State-of-the-Art Feedback Control Systems

In order to minimize the effects of acoustic feedback, a vast range of techniques for feedback control have been developed in the past. In this section, we provide an overview of these and discuss some remaining challenges and possible solutions.

Feedback problems can be reduced via electroacoustic design, e.g. by reducing vent diameter, ensuring a tight seal of ear mold in the ear, etc. [2]. However, even the best electroacoustic design may not be sufficient to avoid feedback problems. One way of dealing with these remaining feedback problems is by means of signal processing in hearing aids.

There are different ways to categorize signal processing based feedback control techniques [77, 180, 199]. In this work, we divide these techniques into two categories as suggested in [77, 180], the feedforward suppression and feedback cancellation techniques.

Fig. 4 illustrates how these two types of techniques control acoustic feedback. The feedforward suppression techniques in Fig. 4(a) modify the forward path $\mathbf{f}(n)$ from the microphone signal $y(n)$ to the loudspeaker signal $u(n)$ for suppressing the feedback effect, whereas the feedback cancellation techniques in Fig. 4(b) make an estimation $\hat{\mathbf{h}}(n)$ of the acoustic feedback path $\mathbf{h}(n)$ to create a signal $\hat{v}(n)$ to cancel the feedback signal $v(n)$.

The motivation for both categories is to ensure that the conditions of the Nyquist stability criterion in Eqs. (2) and (3) are not satisfied. Recall the definition of the open-loop transfer function $\Theta(\omega, n) = F(\omega, n)H(\omega, n)$ in Eq. (1). The feedforward suppression techniques in Fig. 4(a) modify the forward path transfer function $F(\omega, n)$ to avoid that $\Theta(\omega, n)$ fulfills the conditions of the Nyquist stability criterion by e.g. carrying out a gain reduction. For the cancellation system shown in Fig. 4(b), the open-loop transfer

Table 1: Categorized feedforward suppression methods.

Category	Feedforward Suppression Method
Gain reduction	Fullband gain reduction Automatic equalization Notch-filter-based gain reduction Spatial filtering
Phase modification	Frequency shifting/compression Phase modulation

function $\Theta_c(\omega, n)$ is expressed by

$$\Theta_c(\omega, n) = F(\omega, n)\left(H(\omega, n) - \hat{H}(\omega, n)\right), \tag{8}$$

where $\hat{H}(\omega, n)$ is the frequency response of $\hat{\mathbf{h}}(n)$. The feedback cancellation techniques minimize the contribution of $H(\omega, n) - \hat{H}(\omega, n)$ in Eq. (8).

Clearly, an ideal feedback cancellation system is better than a feedforward suppression system, since it removes the feedback contribution of $H(\omega, n)$ completely and provides an unmodified forward path transfer function $F(\omega, n)$. In the following, we provide a brief overview of feedforward suppression techniques, before we focus on feedback cancellation techniques and their remaining challenges.

2.1 Feedforward Suppression

The feedforward suppression techniques can be further divided into two categories: gain reduction and phase modification methods. In both cases, the goal is to alter $F(\omega, n)$ so that ideally $|\Theta(\omega, n)| \ll 1$ and/or $\angle\Theta(\omega, n) \neq l2\pi \ \forall \ \omega, l \in \mathbb{Z}$, respectively. Table 1 provides an overview of different methods within these two categories of feedforward suppression techniques.

2.1.1 Gain Reduction

Obviously, the gain reduction can be simply and effectively performed by the users of audio systems using a volume control to reduce $|F(\omega, n)|$. More sophisticated automatic gain reduction methods exist. For example in [150], a fullband gain reduction in $|F(\omega, n)|$ is carried out based on the detection of system instability. However, the fullband gain reduction is often not necessary for stabilizing audio systems, and it might unnecessarily reduce gain in frequency regions without stability or sound quality problems.

Therefore, automatic equalization in frequency subbands, e.g. in auditory critical bands, have been proposed [6, 135, 136, 190]; the gain reduction is only carried out in the frequency regions where $|\Theta(\omega, n)|$ is close to unity. A further refinement of the

equalization methods are the notch-filter-based techniques [24, 41, 55, 111, 198], where an instability frequency is detected and a notch filter is constructed for this particular frequency. In this way, the gain reduction is performed in a minimal frequency region, ideally leading to fewer distortions in the loudspeaker signal. To further reduce artifacts, a gain reduction scheme based on audibility of signal components is suggested in [148].

An attempt to minimize undesired gain reductions is carried out in [35] using spatial filters by assuming that the feedback and desired signals are coming from different spatial directions. A microphone array beamformer [22, 52] is used to let the desired sound signals from certain directions pass through unmodified, whereas the beamformer places its spatial nulls in the directions of feedback signals. Hence, the gain provided to the desired signals is ideally unchanged, whereas the feedback signal is attenuated. In [35], both loudspeaker and microphone arrays are suggested to achieve this.

However, in all these methods, the general problem is that large values of $|\Theta(\omega, n)|$ must be detected and the gain reduction may therefore be applied at times or frequencies where no instability is present, leading to sound quality degradations for the user. Furthermore, with gain reductions the audio system might only provide a less-than-desired amplification in $|F(\omega, n)|$.

2.1.2 Phase Modification

The second group of feedforward suppression techniques involves phase modifications of $F(\omega, n)$, and it can e.g. be performed by modulating $F(\omega, n)$ with an exponential function $e^{j\varphi(\omega,n)}$, as

$$F_{\mathrm{m}}(\omega, n) = F(\omega, n) \cdot e^{j\varphi(\omega,n)}, \tag{9}$$

to form the modified forward path frequency response $F_{\mathrm{m}}(\omega, n)$. Frequency shifting [23, 166, 167] and phase modulation [128] methods are two well-known phase modification methods which use a linear and sinusoidal phase function of $\varphi(\omega, n)$, respectively. In addition to these methods, a delay modulation method is discussed in [142]. Implementations of these methods are relatively simple, see e.g. [202]. Generally, modifying the signal phase by $e^{j\varphi(\omega,n)}$ causes the system to become a time-varying system, and, strictly speaking, the Nyquist stability criterion does not apply anymore.

Frequency shifting and phase modulation methods break the acoustic feedback loop by moving feedback sound to a different frequency. Furthermore, they have a smoothing effect on the open-loop magnitude function $|\Theta(\omega, n)|$ so that the maximum value of $|\Theta(\omega, n)|$ would generally be smaller [13, 167]. A similar smoothing effect on $|\Theta(\omega, n)|$ can also be obtained using spatial filtering [39].

However, the frequency shifting and phase modulation methods do generally not preserve harmonic structures found in voiced part of speech signals [126] and many audio signals. The consequence could therefore be a sound quality degradation in the loudspeaker signals $u(n)$. A frequency compression [4] technique can be used to preserve

the harmonic frequency structure and still being able to compensate acoustic feedback, although the frequency compressed signal sounds clearly different than the original signal. In general, an additional gain of about 6 dB can be obtained in $|F(\omega, n)|$, when limiting the sound quality degradation to a low level in phase modification methods [166].

2.1.3 Summary

Although somewhat effective for feedback control, the feedforward suppression techniques have significant limitations. The gain reduction techniques limit the amplification in the forward path $\mathbf{f}(n)$, which is contradicting the main purpose of audio reinforcement systems including hearing aids. Phase modification techniques can lead to severe sound quality distortions in loudspeaker signals $u(n)$. In the next section, we give an overview of feedback cancellation techniques which generally allow a higher forward path gain $|F(\omega, n)|$ and better sound quality in $u(n)$.

2.2 Feedback Cancellation Using Adaptive Filters

In contrast to the feedforward suppression approaches, the feedback control in a feedback cancellation system using adaptive filters [76, 163, 164, 207] is not based on modifications of the forward path $\mathbf{f}(n)$. In Fig. 4(b), the adaptive filter $\hat{\mathbf{h}}(n)$ estimates the true acoustic feedback path $\mathbf{h}(n)$ in a system identification configuration [116, 151]. Ideally, $\hat{\mathbf{h}}(n) = \mathbf{h}(n)$ and the feedback cancellation signal $\hat{v}(n)$ is thereby identical to the feedback signal $v(n)$; when subtracting $\hat{v}(n)$ from the microphone signal $y(n)$, we obtain a completely feedback canceled signal, that is $e(n) = x(n)$.

2.2.1 Basics of Adaptive Filters

There are different ways to estimate the adaptive filter $\hat{\mathbf{h}}(n)$ which we assume to be of order $L-1$. A general class of adaptive filters referred to as Wiener filters minimizes the cost function $J_{\mathrm{MSE}}(n)$ in terms of the mean square error of $e(n)$,

$$J_{\mathrm{MSE}}(n) = E\left[e^2(n)\right], \tag{10}$$

where $e(n) = y(n) - \hat{\mathbf{h}}^T(n)\mathbf{u}(n)$, $\mathbf{u}(n) = [u(n), u(n-1), \ldots, u(n-L+1)]^T$ is the loudspeaker signal vector, $E[\cdot]$ is the expectation operator, and the signals $u(n)$ and $y(n)$ are considered realizations of the underlying stochastic processes.

The Wiener filter is derived based on ensemble averages, so that the filter is statistically optimal on average across all realizations of the underlying stochastic processes. Minimizing Eq. (10) with respect to $\hat{\mathbf{h}}(n)$, we find the Wiener-Hopf equation as, see e.g. [76],

$$\mathbf{h}_{\mathrm{o}}(n) = \mathbf{R}_{uu}^{-1}(n)\mathbf{r}_{uy}(n), \tag{11}$$

where $\mathbf{h}_{\mathrm{o}}(n)$ is the Wiener-Hopf solution, $\mathbf{R}_{uu}(n)$ and $\mathbf{r}_{uy}(n)$ are the autocorrelation matrix and the cross-correlation vector of signals $u(n)$ and $y(n)$, defined as

$$\mathbf{R}_{uu}(n) = E\left[\mathbf{u}(n)\mathbf{u}^T(n)\right], \tag{12}$$

$$\mathbf{r}_{uy}(n) = E\left[\mathbf{u}(n)y(n)\right], \tag{13}$$

respectively.

To avoid inverting the correlation matrix $\mathbf{R}_{uu}(n)$ in Eq. (11), a deterministic gradient approach such as the steepest decent algorithm can be used to recursively compute the Wiener-Hopf solution $\mathbf{h}_{\mathrm{o}}(n)$. The gradient with respect to $\hat{\mathbf{h}}(n)$ can be shown to be

$$\begin{aligned}\frac{\partial J_{\mathrm{MSE}}(n)}{\partial \hat{\mathbf{h}}(n)} &= -2\left(\mathbf{r}_{uy}(n) - \mathbf{R}_{uu}(n)\hat{\mathbf{h}}(n)\right)\\ &= -2E\left[\mathbf{u}(n)e(n)\right],\end{aligned} \tag{14}$$

and the update of $\hat{\mathbf{h}}(n)$ would be

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu(n)E\left[\mathbf{u}(n)e(n)\right], \tag{15}$$

with the step size parameter $\mu(n)$.

Furthermore, a widely used stochastic gradient approach is the least mean square (LMS) algorithm, firstly proposed in the area of telecommunication [206], due to its simplicity and it does not require the knowledge of $\mathbf{R}_{uu}(n)$ and $\mathbf{r}_{uy}(n)$. In the LMS algorithm, the adaptive filter estimation of $\hat{\mathbf{h}}(n)$ is carried out using the stochastic gradient vector $\mathbf{u}(n)e(n)$ and the step size parameter $\mu(n)$, as

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu(n)\mathbf{u}(n)e(n). \tag{16}$$

Other well-known stochastic gradient algorithms are normalized least mean square (NLMS) and affine projection (AP) algorithms. The NLMS differs from the LMS algorithm by utilizing a step size parameter scaled by the signal power/energy estimate of $u(n)$ in terms of $\mathbf{u}^T(n)\mathbf{u}(n) + \delta$, where δ is a regularization parameter for ensuring numerical stability [76]. The AP algorithm can be considered as a generalization of the NLMS algorithm, which involves the loudspeaker signal matrix $\mathbf{A}(n) = [\mathbf{u}(n), \mathbf{u}(n-1), \ldots, \mathbf{u}(n-N+1)]$, of order $N-1$, instead of the loudspeaker signal vector $\mathbf{u}(n)$. In this way, the NLMS algorithm is an AP algorithm with $N = 1$. Both algorithms improve the convergence rate of the original LMS algorithm at the cost of increased computational complexity.

Another class of adaptive filters is based on a deterministic approach referred to as the method of least squares (LS). In contrast to the Wiener filter which is derived from the mean square error $E[e^2(n)]$ to be optimal on average across all realizations of the underlying stochastic process (ensemble average), the LS approach is based on averages

of deterministic data samples over time. More specifically, it minimizes the cost function $J_{\mathrm{LS}}(n)$ in terms of the sum of squares of the error signal $e(n)$ as

$$J_{\mathrm{LS}}(n) = \sum_{i=0}^{n} e^2(i). \tag{17}$$

The basic LS approach requires a potentially computationally complex matrix inversion. Therefore, the recursive least squares (RLS) algorithm was developed based on the matrix inversion lemma to bypass the matrix inversion [76]. The RLS algorithm typically increases convergence rate compared to the AP and NLMS algorithms, depending on the signal properties of $u(n)$. Furthermore, the RLS can be shown to be a special case of the Kalman filter framework, which has a recursive solution based on the latest data samples and its state estimate [76].

The NLMS type algorithms for adaptive estimation of $\hat{\mathbf{h}}(n)$ are typically preferred in a hearing aid application [101, 165] due to their simplicity and tracking property. In this work, we mainly focus on the NLMS type algorithms.

2.2.2 Advances in Adaptive Filtering

A lot of specific improvements for different applications have been proposed in the past, such as the algorithms choosing optimal step size and regularization parameters such as [1, 12, 110, 112, 147, 170, 200], the so-called filtered-X algorithms with a fixed filter to model a (typically known) part of the unknown impulse response in series with the adaptive filter [15, 78, 98, 129, 207], the proportionate algorithms [10, 38, 50, 104, 119, 120, 211] for long and sparse impulse response estimations, the robust algorithms with slow divergence properties as discussed in [9], the signed regressor algorithms for better implementation simplicity [37, 53, 109, 169], other computationally efficient algorithms such as [32, 37, 51, 168, 178], etc.

The adaptive filter can also be estimated in subbands [46, 103, 209, 210], in the frequency domain [93, 174, 185], or in a band-limited frequency region [29]. The main advantages are typically increased convergence rate, more control flexibilities, and computational complexity reductions for adaptive filters with many taps. One of the biggest drawbacks for real-time applications in traditional subband and frequency domain approaches is that they introduce an additional delay in the forward path $\mathbf{f}(n)$. However, the delayless subband structure introduced in [134] and refined in [83] eliminates this additional delay.

Moreover, a combination of feedback cancellation and feedforward suppression in terms of a gain reduction is suggested in [157]. The motivation for this was that the feedback cancellation in practice would not be perfect, an adaptive gain limit is therefore computed based on the tracking ability of the feedback cancellation system, and it is applied to the actual audio system gain to further ensure system stability. Some similar gain processing approaches are suggested in [16, 61]. Other studies have been carried

out for analyzing and determining optimal strategies for combined feedback cancellation and beamformer or noise suppression systems in hearing aids, see e.g. [89, 114, 121, 186].

Furthermore, much work has been done that analyze [40, 59, 60, 62, 63, 72, 105, 117, 122, 138, 177, 205, 208] and improve [7, 44, 191] adaptive algorithms in terms of e.g. robustness, stability bounds, convergence rate, and steady-state behavior.

2.2.3 Summary

Feedback cancellation using adaptive filters is generally more effective to control feedback than feedforward suppression methods, and it provides better sound quality [180, 199]. However, one of the biggest remaining problems is the biased estimation problem for closed-loop systems such as hearing aids. In the next section, we study this problem in more details.

2.3 The Biased Estimation Problem

One of the biggest problems remaining in using adaptive filters for acoustic feedback cancellation is the biased estimation of $\hat{\mathbf{h}}(n)$ [175, 180, 185]. It can be easily shown by inserting Eqs. (12) and (13) in the Wiener-Hopf equation in Eq. (11), that

$$\begin{aligned}\mathbf{h}_{\mathrm{o}}(n) &= \mathbf{R}_{uu}^{-1}(n)\mathbf{r}_{uy}(n) \\ &= E\left[\mathbf{u}(n)\mathbf{u}^T(n)\right]^{-1} E\left[\mathbf{u}(n)y(n)\right] \\ &= E\left[\mathbf{u}(n)\mathbf{u}^T(n)\right]^{-1} E\left[\mathbf{u}(n)\left(\mathbf{u}^T(n)\mathbf{h}(n)+x(n)\right)\right] \\ &= \underbrace{\mathbf{h}(n)}_{\text{Desired}} + \underbrace{E\left[\mathbf{u}(n)\mathbf{u}^T(n)\right]^{-1} E\left[\mathbf{u}(n)x(n)\right]}_{\text{Bias}}. \end{aligned} \tag{18}$$

Eq. (18) shows that the Wiener-Hopf solution $\mathbf{h}_{\mathrm{o}}(n)$ consists of two terms, the true feedback path $\mathbf{h}(n)$ and the product of the inverse correlation matrix $E\left[\mathbf{u}(n)\mathbf{u}^T(n)\right]^{-1}$ and the cross-correlation vector $E[\mathbf{u}(n)x(n)]$ between the loudspeaker signal $u(n)$ and the incoming signal $x(n)$. To obtain an unbiased estimation, the cross-correlation vector must satisfy $E[\mathbf{u}(n)x(n)] = \mathbf{0}$ in Eq. (18). In the following, we show that this is generally not the case in practice, so that $\mathbf{h}_{\mathrm{o}}(n)$ becomes biased.

In general audio reinforcement systems including hearing aids, the loudspeaker signal $u(n)$ is ideally an amplified version of the incoming signal $x(n)$, see Fig. 1. Furthermore, there is a processing delay d through the audio processing system. To demonstrate the problem, we simply model the loudspeaker signal $u(n)$ as

$$u(n) = c \cdot x(n-d), \tag{19}$$

where c is a gain factor. The cross-correlation vector $E[\mathbf{u}(n)x(n)]$ can be written as

$$E\left[\mathbf{u}(n)x(n)\right] = c \cdot E\left[\mathbf{x}(n-d)x(n)\right], \tag{20}$$

where $\mathbf{x} = [x(n), x(n-1), \ldots, x(n-L+1)]^T$ denotes the incoming signal vector. Let $r_{xx}(k) = E[x(n)x(n-k)]$ denote the autocorrelation function of $x(n)$, such that Eq. (20) can be rewritten as

$$E\left[\mathbf{u}(n)x(n)\right] = c \cdot \begin{bmatrix} r_{xx}(d) \\ r_{xx}(d+1) \\ \vdots \\ r_{xx}(d+L-1) \end{bmatrix}. \tag{21}$$

Eq. (21) reveals that the autocorrelation $r_{xx}(k)$ of incoming signals $x(n)$ is the key factor in obtaining an unbiased estimation of $\hat{\mathbf{h}}(n)$. For an incoming signal with a short correlation time compared to the audio system latency d so that $r_{xx}(|k|) = 0 \; \forall \; |k| \geq d$, the cross-correlation vector $E[\mathbf{u}(n)x(n)]$ in Eq. (21) would be a null vector, and unbiased estimation, i.e. $\mathbf{h}_\mathrm{o}(n) = \mathbf{h}(n)$, is obtained in Eq. (18). Otherwise, the cross-correlation vector in Eq. (21) would consist of nonzero values and the consequence is an biased estimation, i.e. $\mathbf{h}_\mathrm{o}(n) \neq \mathbf{h}(n)$.

Unfortunately, for many everyday sound signals like speech signals and tonal signals such as most musical signals and alarm tones, the signal correlation time is longer than the audio system latency, especially in a hearing aid application, where the system latency is typical between $4 - 8$ ms [21, 182]. Consequently, the estimate $\hat{\mathbf{h}}(n)$ becomes biased, and the cancellation performance is reduced and/or howling occurs.

2.4 Towards Unbiased Estimation

2.4.1 Methods for Unbiased Estimation

Different methods have been proposed to minimize the effect of biased estimation. Some studies take the physical feedback path into consideration to limit the estimation of $\hat{\mathbf{h}}(n)$ to avoid bias. For example in [97, 155], the estimation of $\hat{\mathbf{h}}(n)$ is constrained by the a priori knowledge of true acoustic feedback paths $\mathbf{h}(n)$, typically as a feedback path model, to minimize the influence from the correlation between $u(n)$ and $x(n)$. Some studies to determine acoustic feedback path characteristics and models can e.g. be found in [80, 81, 123, 125, 156].

Band-limited adaptations [29, 100] have been suggested for carrying out the estimation in the frequency regions without strong autocorrelation of $x(n)$. In [137], an additional microphone is used to obtain an incoming signal estimate, which is removed from the signals for the adaptive estimation to minimize the biased estimation problem.

Furthermore, a detection can be used to control the adaptive algorithms. In [102], a slow adaptation of the feedback path estimate is chosen to avoid fast divergence due to autocorrelation in $x(n)$, and the adaptation speed is increased when system instability or feedback path changes are detected. In [106], when strong autocorrelation is detected in the incoming signal $x(n)$, the adaptation is slowed down or frozen. However, these

Table 2: Some well-known decorrelation techniques.

Decorrelation Method	Decorrelation Path
Delay	Loudspeaker signal and/or filter estimation
Phase modification	Loudspeaker signal
Pre-whitening	Filter estimation
Probe noise injection	Loudspeaker signal and filter estimation

detect-and-react strategies do not work well when there is a strong autocorrelation in $x(n)$ and the feedback path changes at the same time.

2.4.2 Traditional Decorrelation Methods for Unbiased Estimation

Unbiased estimation of $\hat{\mathbf{h}}(n)$ can be obtained via methods which decorrelate the loudspeaker signal $u(n)$ from the incoming signal $x(n)$, so that the term $E[\mathbf{u}(n)x(n)] \rightarrow \mathbf{0}$ in Eq. (18). The decorrelation can be performed either in the loudspeaker signal path or in the adaptive filter estimation path in the cancellation system. The advantage of carrying out the decorrelation in the adaptive estimation path is that this does not modify the loudspeaker signal, so that no sound quality degradation is introduced due to decorrelation. Table 2 provides an overview of some well-known decorrelation techniques and indicates whether the decorrelation is performed in the loudspeaker signal path or in the adaptive filter estimation path. Fig. 5 illustrates these decorrelation techniques.

Delay

Inserting a delay in the forward path $\mathbf{f}(n)$ and/or in the adaptive filter estimation path as suggested in [25, 175] is the simplest decorrelation method, see Fig. 5(a). It is used to partly bypass the typically strong signal correlation at lower time lags and model the initial delay in acoustic feedback path impulse response due to the distance between loudspeaker and microphone.

However, only relatively short delays can be used in real-time applications and to correctly model the initial delay in feedback paths $\mathbf{h}(n)$. Therefore, although delay is effective to decorrelate many signals with relatively short correlation times, its use is limited in practice.

Phase Modification Methods

Fig. 5(b) shows a feedback cancellation system with phase modification in the forward path $\mathbf{f}(n)$. As mentioned in Sec. 2.1.2, frequency shifting and phase modulation can be used for feedforward suppression in the forward path. However, they are also effective for decorrelation between $u(n)$ and $x(n)$. Frequency compression and shifting have been discussed in e.g. [92, 153], and phase modification has been studied in e.g. [19, 71, 149],

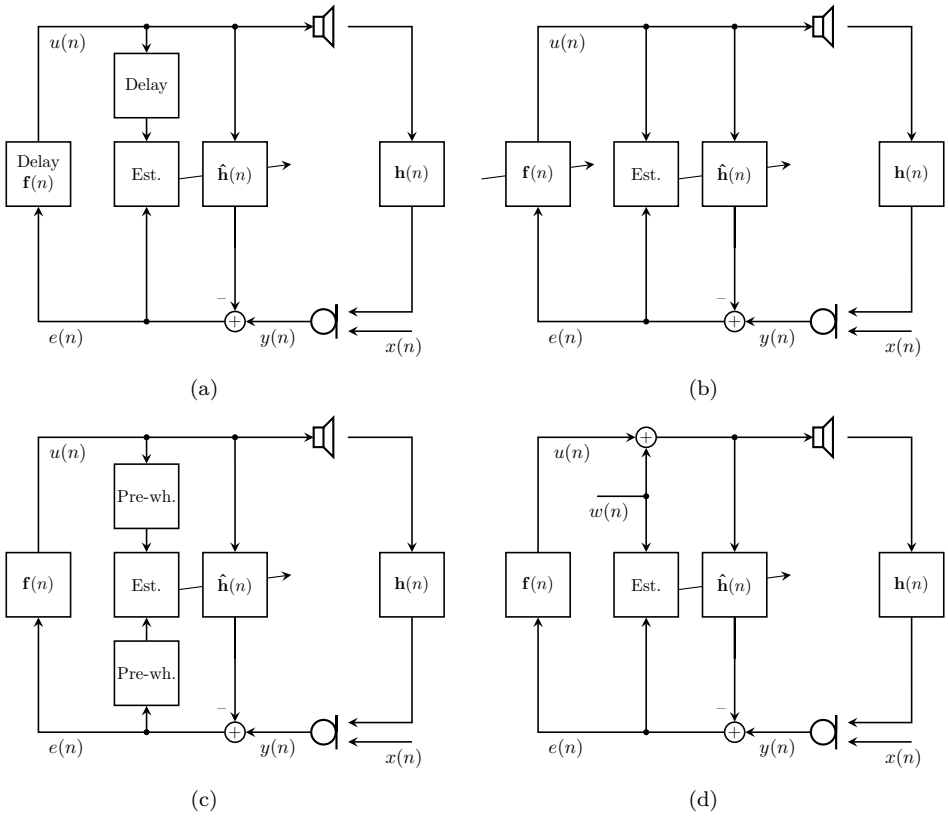


Fig. 5: Some decorrelation techniques for an unbiased estimation of $\hat{\mathbf{h}}(n)$. (a) A delay in the loudspeaker signal path and/or in the adaptive filter estimation path. (b) Phase modification in the loudspeaker signal path. (c) Pre-whitening in the adaptive filter estimation path. (d) Probe noise signal injection of $w(n)$ in the loudspeaker signal path and in the adaptive filter estimation path.

for improving feedback cancellation performance by decorrelating $u(n)$ from $x(n)$. Generally, the cancellation performance improvement is relatively large for phase modification methods, at the price of audible artifacts due to the modifications of loudspeaker signal $u(n)$.

Pre-whitening Approaches

In the pre-whitening approach, the decorrelation is carried out on the signals used for the estimation of $\hat{\mathbf{h}}(n)$, see Fig. 5(c). In this way, the forward path $\mathbf{f}(n)$ is unmodified, and no artifacts are introduced to the loudspeaker signal $u(n)$ due to decorrelation.

Some simple pre-whitening approaches are suggested by removing parts of the sig-

nals, which are strongly correlated, from the adaptive filter estimation. For example in [162], notch filters are estimated and used to remove the frequencies with strong signal correlation from the signals $e(n)$ and $u(n)$ to form $e_{\mathrm{p}}(n)$ and $u_{\mathrm{p}}(n)$, which are used to obtain an unbiased estimation of $\hat{\mathbf{h}}(n)$.

Another well-known pre-whitening method is the prediction error method, which is known from closed-loop identification [5, 42, 73, 192], and it was analyzed and suggested for a hearing aid system in e.g. [78, 79, 183, 184], or for applications with long feedback path impulse responses such as public address systems and automotive speech reinforcement systems in e.g. [30, 158, 159].

Many prediction error method based approaches whiten the signals for the adaptive filter estimation by assuming that the incoming signal $x(n)$ can be modeled well as a white noise sequence $\epsilon(n)$ filtered through an all-pole model $A(n, z)$,

$$A(n, z) = \frac{1}{1 + \sum_{k=1}^{L_p-1} p_k(n) z^{-k}}, \tag{22}$$

where z^{-1} is the unit delay operator.

Let $\mathbf{p}(n) = [1, p_1(n), \ldots, p_{L_p-1}(n)]^T$. The prefilter $\hat{\mathbf{p}}(n) = [1, \hat{p}_1(n), \ldots, \hat{p}_{L_p-1}(n)]^T$ is then jointly estimated with the cancellation filter $\hat{\mathbf{h}}(n)$. Furthermore, the prefilters are applied to the signals $u(n)$ and $e(n)$ entering the adaptive filter estimation and they approximately whiten the incoming signal components in these signals and thereby compensate for the biased estimation of $\hat{\mathbf{h}}(n)$. Ideally, the signal component $x(n)$ in the error signal $e(n)$, filtered by the prefilter $\hat{\mathbf{p}}(n)$, would be the white noise excitation sequence $\epsilon(n)$ due to the assumption of an autoregressive incoming signal $x(n)$ and it would no longer cause a biased estimation. In this way, the prediction error method would ideally provide an unbiased estimation of $\hat{\mathbf{h}}(n)$. In practice it works well for most speech signals, since the unvoiced part of speech signals can be modeled well as a white noise sequence filtered through the all-pole model in Eq. (22) [126].

However, these all-pole model based prediction error methods do not perform well for e.g. music signals, because the assumption that $x(n)$ is autoregressive is violated. Several studies have suggested modifications to improve the performance for music signals by using instead a frequency-warped all-pole model [194], a cascade of a conventional all-pole linear prediction model and one of the alternative linear prediction models [196] described in [195], and sinusoidal models [139, 140], or the prediction error method in a transform domain [54].

Traditional Probe Noise Approaches

Fig. 5(d) illustrates a feedback cancellation system using a probe noise signal $w(n)$, where $w(n)$ is uncorrelated with $x(n)$ and $u(n)$ by construction, and an unbiased estimation of $\hat{\mathbf{h}}(n)$ can be obtained based on $w(n)$. Alternatively, the mixture of $u(n)$ and $w(n)$ can be used for the estimation [199]. However, in that case the estimation of $\hat{\mathbf{h}}(n)$ can be shown to be a mixture of a biased and an unbiased part.

A class of closed-loop identification methods utilizes a probe noise signal $w(n)$ added to the original loudspeaker signal $u(n)$, e.g. the indirect and joint input-output approaches [42, 43, 172, 173, 193]. The goal of the probe noise signal $w(n)$ is to estimate $\hat{\mathbf{h}}(n)$ indirectly in an open-loop setup.

It is also possible to deploy the added probe noise signal $w(n)$ in a different way, where the estimation of $\hat{\mathbf{h}}(n)$ is directly based on $w(n)$. A non-continuous adaptation is studied in [96], the adaptive estimation of $\hat{\mathbf{h}}(n)$ is only performed when the system is detected to be close to instability, and the loudspeaker signal $u(n)$ is then muted and the probe noise signal $w(n)$ is presented as the loudspeaker signal. In this way, the adaptive estimation is directly driven by the probe noise signal $w(n)$ in an open-loop configuration. By looking at Eq. (18), this corresponds to replacing the loudspeaker signal vector $\mathbf{u}(n)$ with the probe noise signal vector $\mathbf{w}(n) = [w(n), w(n-1), \ldots, w(n-L+1)]^T$, and $\mathbf{h}_o(n)$ can be shown to be,

$$\mathbf{h}_o(n) = \mathbf{R}_{ww}^{-1}(n)\mathbf{r}_{wy}(n) \tag{23}$$

$$= \mathbf{h}(n) + \underbrace{E\left[\mathbf{w}(n)\mathbf{w}^T(n)\right]^{-1} E\left[\mathbf{w}(n)x(n)\right]}_{=\mathbf{0}}. \tag{24}$$

Since $w(n)$ is uncorrelated with $x(n)$, we get an unbiased estimation, i.e. $\mathbf{h}_o(n) = \mathbf{h}(n)$. In [127], a similar non-continuous adaptation with a different decision criterion for starting the adaptation is proposed; specifically, probe noise insertion and adaptive filter estimation is only performed during quiet intervals; the attempt is made to reduce the audible artifacts introduced by a high-level probe noise signal. Nevertheless, the cancellation performance in these systems is highly dependent on the detectors for the non-continuous adaptation.

Although the probe noise approach in principle eliminates the biased estimation problem, the perhaps biggest drawback in using it in feedback cancellation systems in general is that the probe noise signal must often be powerful compared to the loudspeaker signal $u(n)$, for achieving a noticeable improvement in acoustic feedback cancellation systems. Unfortunately, powerful probe noise signals become clearly audible [79, 173, 197]. In [66], it is shown theoretically that when the probe noise level is adjusted to be inaudible, the probe noise to disturbing signal ratio is generally low and the convergence rate of the adaptive system is decreased, by as much as a factor of 30 in practice compared to a traditional cancellation system. Hence, the decreased convergence rate and/or clearly audible artifacts limit the practical use of the probe noise approach.

Studies exist that minimize the sound quality degradation due to decorrelation by using specifically generated probe noise signals, such as in [124], where the high frequency part of the loudspeaker signal $u(n)$ is replaced by a synthetic signal, and the replacement signal is perceptually similar to $u(n)$ but it is statistically uncorrelated with $x(n)$.

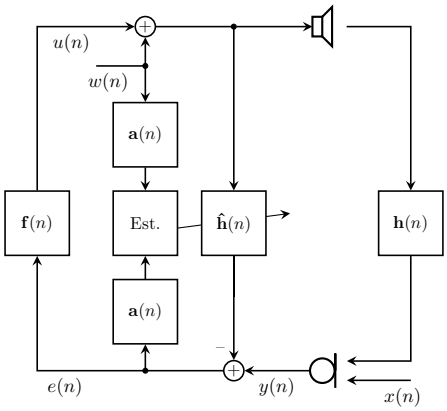


Fig. 6: A multiple microphone channel hearing aid acoustic feedback cancellation system using a probe noise signal $w(n)$ and probe noise enhancement filters $\hat{\mathbf{a}}_i(n)$.

Generally, a compromise exists in traditional probe noise approaches, between sound quality degradation in the loudspeaker signal $u(n)$ and improved feedback cancellation performance.

2.4.3 A Novel Probe Noise Based Approach for Unbiased Estimation

In [69], we introduced a novel probe noise approach which provides an unbiased estimation of $\hat{\mathbf{h}}(n)$ without introducing perceptually significant sound quality degradations. Fig. 6 illustrates the structure of this novel probe noise based approach, which looks somewhat like a combination of the pre-whitening approach and the traditional probe noise approach shown in Figs. 5(c) and 5(d), respectively. However, the goal of the enhancement filters $\mathbf{a}(n)$ is to increase the probe noise to disturbing signal ratio, in contrast to the pre-whitening approach that decorrelates the incoming signal $x(n)$ and the loudspeaker signal $u(n)$.

We desire a probe noise signal $w(n)$ with the highest possible signal power at all frequencies while being inaudible in the presence of the original loudspeaker signal $u(n)$. Therefore, we generate the probe noise signal $w(n)$ using a spectral masking model based on e.g. [91, 146]. This kind of probe noise generation method was firstly introduced for system identification and feedback cancellation applications in [33, 90]; however, the low probe noise signal power leads to a decreased convergence rate in the adaptive systems. Hence, further improvements of the system convergence rate are needed.

The improvements found in the proposed approach are obtained by using so-called enhancement filters $\mathbf{a}(n)$ to reduce the influences of the disturbing signals, e.g. the incoming signals $x(n)$, for the estimation of $\hat{\mathbf{h}}(n)$. At the same time, the filters $\mathbf{a}(n)$ are specifically designed, in coordination with the probe noise signal $w(n)$, so that they are

statistically transparent for $w(n)$ in the estimation of $\hat{\mathbf{h}}(n)$. This characteristic of $\mathbf{a}(n)$ increases the probe noise to disturbing signal ratio, and it leads to an increased convergence rate compared to the traditional probe signal approach, without compromising the steady-state behavior in the cancellation system. We refer to [69] for more details.

A comparison between different state-of-the-art feedback cancellation systems in [70] shows that this novel probe noise approach outperforms other feedback cancellation systems, including a prediction error method based system and a frequency shifting based system, since this novel probe noise approach is very robust against biased estimation problem for all types of incoming signals without introducing perceptual sound quality degradation. Furthermore, the computational complexity increase in this novel approach is less than a factor of three compared to traditional feedback cancellation systems [70].

3 Evaluation of Feedback Cancellation Systems

An evaluation of feedback cancellation systems needs to cover various aspects, and the final assessment is typically based on a trade-off between these aspects. We consider generally three major aspects: feedback cancellation performance, sound quality distortion, and computational complexity.

3.1 Feedback Cancellation Performance

3.1.1 Evaluation Methods

An objective evaluation of feedback cancellation performance can be carried out either in computer simulations or by physical measurements. Objective evaluation is essentially useful for system analysis and design, because it is reproducible, and based on computer simulation experiments it can evaluate many complicated test scenarios with different test situations, parameter settings, etc. quickly compared to subjective tests based on opinions from test subjects. Objective cancellation performance is typically evaluated in terms of convergence rate, stability bounds, steady-state behavior including steady-state error and tracking error [76].

Generally, it is straightforward to evaluate feedback cancellation performance in simulations. The convergence rate and steady-state error can easily be determined using a static feedback path $\mathbf{h}$. To evaluate the tracking ability of feedback cancellation systems, the feedback paths must undergo variations, this can be achieved using different feedback path variation models [117].

Physical measurements in a static feedback environment can be performed e.g. on a mannequin, designed for sound quality testing. However, to make feedback path variations reproducible in physical measurements is more challenging. A robotic mechanism is suggested for this purpose in [187], particularly to measure the influence of a fast

variation/change in feedback paths due to a telephone handset brought next to user's ear and hearing aid.

3.1.2 Performance Measures

There are generally two types of performance evaluation measures for feedback cancellation systems.

The first type is based on measurements carried out on signals in the cancellation system. Traditional evaluation measures are often formulated in terms of the mean square error $E[e^2(n)]$, e.g. its decay which indicates the convergence, and its steady-state value. Other measures are typically based on comparisons of different signals such as $e(n)$, $u(n)$, and/or the amplifications $|F(\omega, n)|$ in the systems as described and applied in evaluations of commercial hearing aids in [45, 171, 181, 182]. The greatest advantage of this kind of evaluation measures is that they can be used both in physical measurements and simulations for design purpose, since all signals are in principle accessible, although it might be necessary to perform additional reference measurements in a system without feedback. Some of these cancellation performance measures based on measured signals are also somewhat related to sound quality [181].

The other type of evaluation measures is based on a distance measure between the true and estimated feedback path, such as the mean square deviation $E[\|\mathbf{h}(n)-\hat{\mathbf{h}}(n)\|^2]$ [76]. In contrast to the signal based measures, the use of these measures is more limited since the true acoustic feedback path $\mathbf{h}(n)$ must be known a priori. However, this is always the case in computer simulations and, in principle, in physical measurements with a well-controlled acoustic feedback environment, but never in a real application. The greatest advantage of using this kind of evaluation measures is that the evaluation results are directly linked to system stability via the open-loop transfer function, which $\mathbf{h}(n) - \hat{\mathbf{h}}(n)$ is a part of.

There are different variants of the mean square deviation measure such as its frequency domain version $E[|H(\omega, n) - \hat{H}(\omega, n)|^2]$. An example measure of this kind referred to as the power transfer function is introduced in [63]. Furthermore, some commonly used performance measures are based on the mean square deviation such as the maximum stable gain and the added stable gain etc. [181].

In this work, we mainly evaluate cancellation performance, in terms of convergence rate and steady-state behavior, in simulation experiments using the distance based measures.

3.2 Sound Quality

3.2.1 Listening Test

Sound quality evaluation of feedback cancellation systems can be performed using a listening test. In order to assess sound quality of feedback cancellation systems, typically

Table 3: Descriptions of mean-opinion-scores (MOS) scores [87].

MOS	Quality	Description of Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

a paired comparison or an absolute rating of quality is performed [34]. In a paired comparison, test subjects simply choose which of the two presented test signals they prefer. This method is simple and it is often used to improve hearing aid fittings [34]. In tests with absolute ratings of quality, test subjects rate each test signal on a sound quality scale, such as the mean-opinion-scores (MOS) in the range $1-5$ [87], as given in Table 3.

Listening tests based on absolute ratings of quality are more complicated to design and conduct than paired comparison based tests. However, it is easy to assess, based on absolute ratings, how much better is the preferred test signal. Tests based on absolute ratings are often used for evaluation of sound quality in feedback cancellation systems. In [58, 127], test subjects rated hearing aid loudspeaker signals on a scale from 1 to 10, indicating unacceptable and excellent quality, respectively. However, it is also possible to rate directly the difference between two test signals as performed in [29], where test subjects should rate two different hearing aid loudspeaker signals on a scale from 0 to 5, which indicate no difference or one of them is much better, respectively. In our work, we use the absolute ratings of quality method.

Another concern in a listening test for hearing aid applications regards the choice between normal hearing and hearing impaired test subjects. In our work, we choose to evaluate the sound quality using normal hearing test subjects. Our assumption is that if the sound distortion is acceptable for normal hearing people, then it will also be acceptable for hearing impaired people. In this way, we expect the results of normal hearing test subjects provide a sound quality acceptance lower bound. The study in e.g. [20] shows that this assumption is realistic.

3.2.2 Evaluation of Feedback Cancellation System Introduced Artifacts

The overall sound quality degradations in a feedback cancellation system consists of sound distortions due to feedback and sound distortions introduced by the cancellation system, such as the effects of decorrelation, which is used in an attempt to get better cancellation performance and overall sound quality. A severe sound distortion due to decorrelation could in principle lead to better cancellation performance and thereby

improve the overall sound quality. Therefore, it is preferable to evaluate the overall sound quality.

In practice, however, it is generally too complicated to evaluate the overall sound quality subjectively, since the listening test must cover many different aspects such as acoustic situations, system parameter settings, etc. and it will be very time consuming to conduct. Therefore, much work has focused on maximizing cancellation performance and improve the overall sound quality by keeping sound distortions introduced by the cancellation system at a low level. For example in [19], a pairwise comparison test based on the International Telecommunication Union (ITU) recommendation BS.1116-1 [84] is performed to evaluate sound quality distortion due to a time-varying all-pass filter processing in hearing aids. In [70], a multiple comparison based on the ITU recommendation BS.1534-1 [86] (also referred to as a MUSHRA test) is carried out for assessing sound quality distortion introduced by frequency shifting and probe noise injection.

3.2.3 Objective Sound Quality Evaluation

Sound quality evaluation via listening tests is generally complicated and time consuming, and it requires proper preparations and post-processing [8, 131]. Therefore, robust objective predictions of sound quality is useful as a supplement.

Objective sound quality evaluation is typically carried out by comparing test signals to a reference signal without sound distortions. The frequency weighted log-spectral signal distortion [57] is a distance measure between the test signal and the reference signal spectra, and it is e.g. used for sound quality evaluation in feedback cancellation systems [197]. Furthermore, in order to verify that the initial parameter choices for decorrelation only introduce insignificant sound quality distortions, the perceptual evaluation of speech quality (PESQ) and perceptual evaluation of audio quality (PEAQ) models can be applied, see e.g. [69–71]. The PESQ and PEAQ are standardized algorithms for objectively measuring perceived speech and audio quality, described in the ITU recommendations [88] and [85]. In [70], it is shown that there is actually a reasonable agreement between PESQ/PEAQ predictions and the obtained subjective sound quality scores in the listening test for evaluating frequency shifting and probe noise artifacts.

Unfortunately, in contrast to the PESQ and PEAQ scores which are widely accepted for objectively predicting/evaluating speech and audio qualities for coding and/or communication channel artifacts, there is so far no well-established objective sound quality measure, which is verified to be reliable, for evaluating acoustic feedback artifacts.

3.3 Computational Complexity

Computational complexity in terms of required arithmetic operations is an important design parameter in any digital signal processing algorithm. This becomes even more

important for the hearing aid application, since one particular limitation in hearing aids is a very short battery time [101], besides the less powerful processing unit, compared to other electronic mobile devices such as mobile phones or laptops.

Specifically for the feedback cancellation systems in hearing aids, the trade-off is often between cancellation performance, sound quality, and computational complexity when choosing a particular system. In our work, we typically count the number of multiplications (and additions/subtractions) for a particular algorithm to make a rough estimate of computational complexity, as e.g. reported in [70]. On the other hand, we do not focus on divisions, although they are computationally expensive in hardware implementations, since they are rare and very often realized in different ways based on multiplication, subtraction, and table look-up [145].

Moreover, it should be mentioned that another important power consumption concern regards the memory requirement in different algorithms/systems [27, 130]. A less demanding algorithm in terms of arithmetic operations might not be necessarily the most power efficient one, should it require excessive memory usage. Furthermore, numerical robustness in fixed-point implementations is another main concern [31, 115, 118] in general signal processing algorithm design.

4 Acoustic Echo Cancellation

The acoustic feedback problem is very similar in structure to the acoustic echo problem, which also involves an audio system which loudspeaker signal is recaptured by its microphone. In this section, we give a short introduction to the acoustic echo cancellation problem, before we relate the echo and feedback cancellation systems.

4.1 Acoustic Echo and Echo Cancellation

4.1.1 Background

The acoustic echo problem occurs typically in hands-free telephony and teleconferencing situations. Fig. 7 illustrates an echo situation and an echo cancellation system using adaptive filters. The far-end and near-end denote the transmitting and receiving ends over a communication channel, such as a telephone line, where two users, one at the far-end and the other at the near-end side, are communicating.

Both ends can be considered as mirrored copies of each other, therefore, we only focus on the near-end in the following. Ideally, only the near-end speech signal $x(n)$ should be transmitted to the far-end. However, in practice, the signal $v(n)$ is also transmitted to the far-end where it is perceived as an echo.

For convenience, we consider the near-end loudspeaker signal $u(n)$ as an unprocessed speech signal spoken by the far-end user. Very similar to the acoustic feedback problem, the loudspeaker signal $u(n)$ is modified by the near-end echo path $\mathbf{h}_{\mathrm{N}}(n)$ to produce

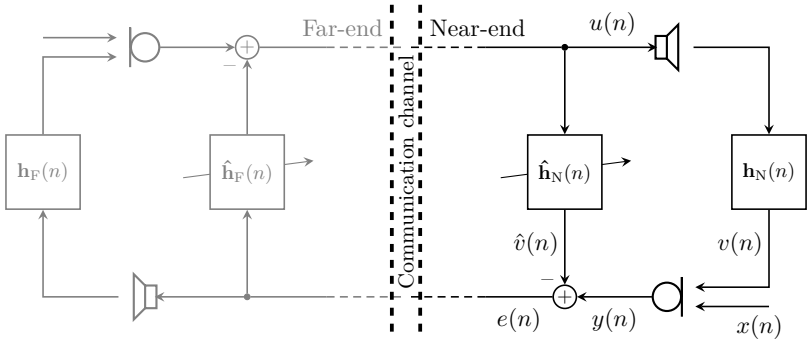


Fig. 7: An acoustic echo cancellation system. The far-end and near-end systems can be considered as mirrored copies of each other.

an echo signal $v(n)$ which is recorded by the microphone. Since the microphone signal $y(n)$ is transmitted back to the far-end, the far-end user would hear a delayed and distorted version of his/her own voice as an echo. This is the basic acoustic echo problem. The echo path $\mathbf{h}_\mathrm{N}(n)$ depends on the acoustic properties of the near-end room such as reflective surfaces and movements of the user [74].

4.1.2 Echo Cancellation Using Adaptive Filters

Echo cancellation using adaptive filters is an effective method to control echoes, see e.g. [9]. In Fig. 7, an adaptive filter $\hat{\mathbf{h}}_\mathrm{N}(n)$ is used to model the echo path $\mathbf{h}_\mathrm{N}(n)$ and create the cancellation signal $\hat{v}(n)$. It is clear that the near-end echo cancellation system in Fig. 7 is very similar in structure to the acoustic feedback cancellation system in Fig. 4(b). The only difference is that there is an additional forward path denoted by $\mathbf{f}(n)$ in the acoustic feedback cancellation system.

In echo cancellation systems, the far-end impulse response difference $\mathbf{h}_\mathrm{F}(n) - \hat{\mathbf{h}}_\mathrm{F}(n)$ can be considered as the counterpart to $\mathbf{f}(n)$, and it is often neglected in the near-end echo cancellation of $\mathbf{h}_\mathrm{N}(n)$, since $\mathbf{h}_\mathrm{F}(n) - \hat{\mathbf{h}}_\mathrm{F}(n)$ typically does not contain significant amplification in contrast to $\mathbf{f}(n)$ in feedback cancellation systems, especially when a relatively accurate estimate $\hat{\mathbf{h}}_\mathrm{F}(n)$ is obtained at the far-end; it leads to a relatively low magnitude of the open-loop transfer function $\Theta_\mathrm{e}(\omega, n)$,

$$\Theta_\mathrm{e}(\omega, n) = \Big(H_\mathrm{F}(\omega, n) - \hat{H}_\mathrm{F}(\omega, n)\Big)\Big(H_\mathrm{N}(\omega, n) - \hat{H}_\mathrm{N}(\omega, n)\Big), \tag{25}$$

where $H_\mathrm{F}(\omega, n)$, $\hat{H}_\mathrm{F}(\omega, n)$, $H_\mathrm{N}(\omega, n)$, and $\hat{H}_\mathrm{N}(\omega, n)$ are frequency responses of $\mathbf{h}_\mathrm{F}(n)$, $\hat{\mathbf{h}}_\mathrm{F}(n)$, $\mathbf{h}_\mathrm{N}(n)$, and $\hat{\mathbf{h}}_\mathrm{N}(n)$, respectively. Hence, whereas the acoustic feedback cancellation is a closed-loop system identification problem, the echo cancellation is generally considered as an open-loop problem.

4.2 Some Relations to Feedback Cancellation

The echo cancellation problem is similar to the feedback cancellation problem. In this section, we present two difficult-to-handle echo cancellation problems, the double-talk problem and the non-uniqueness problem, and we relate them to feedback cancellation systems.

4.2.1 The Double-Talk Problem

One of the main challenges in echo cancellation is the so-called double-talk situation, see e.g. [9]. It occurs when both the far-end and the near-end users are talking simultaneously, such that the near-end loudspeaker signal $u(n)$ and the near-end speech signal $x(n)$ are active at the same time. Unfortunately, adaptive algorithms adjusted to a high convergence rate usually diverge quickly in this situation. A well-known procedure to limit the divergence of adaptive filters is based on a double-talk detector, e.g. the Geigel detector [36], which controls the adaptive algorithm by e.g. freezing or slowing down the adaptation when a double-talk situation is detected. Extensive studies have been carried out to deal with the double-talk situation, see e.g. [48, 49, 201] and the references therein.

Although the double-talk situation is difficult to handle, it is typically not always present in an echo cancellation situation, and it is possible to carry out a normal single-talk adaptation of adaptive filters most of the time. In contrast, in a feedback situation "double-talk" is actually unavoidable and occurs all the time; moving back to Fig. 4(b), we observe that the loudspeaker signal $u(n)$ is a processed version of $x(n)$. Without the presence of an incoming signal $x(n)$, there is no loudspeaker signal $u(n)$, and they will always appear in pairs.

Hence, in feedback cancellation, we need to actively deal with double-talk situation always, and freezing the adaptive filters when detecting a double-talk situation is generally not an option. Traditionally, a relatively slow adaptation is very often needed in order to handle the double-talk situation. From the double-talk problem's point of view, the acoustic feedback cancellation problem is more difficult to solve than the acoustic echo cancellation problem.

4.2.2 The Non-Uniqueness Problem

Another major challenge in echo cancellation arises when stereo or multichannel audio systems are used to provide spatial perception of sound signals. Fig. 8 shows a stereophonic echo cancellation situation with adaptive filters $\hat{\mathbf{h}}_1(n)$ and $\hat{\mathbf{h}}_2(n)$ for echo cancellation in each microphone channel, and the far-end source signal $s(n)$ is modified by the room impulse responses $\mathbf{g}_1(n)$ and $\mathbf{g}_2(n)$ to form the far-end microphone incoming signals $u_1(n)$ and $u_2(n)$.

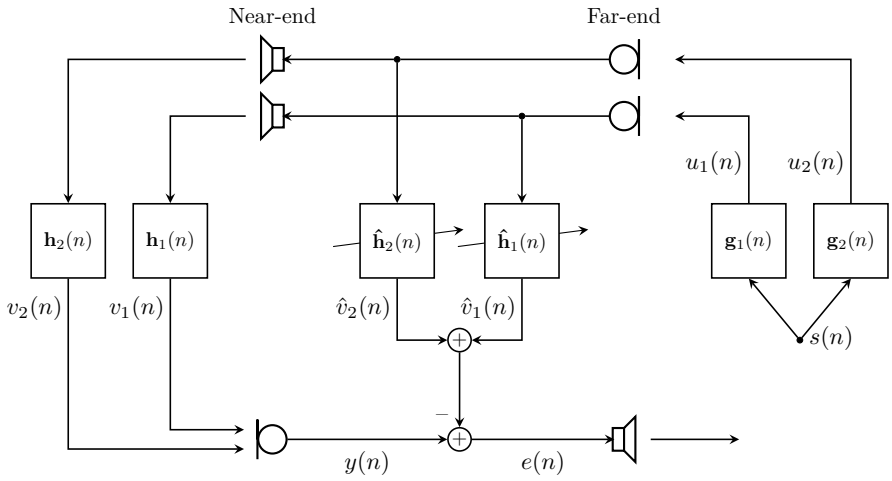


Fig. 8: A stereophonic acoustic echo cancellation system.

The stereophonic echo cancellation situation suffers from a non-uniqueness problem, see e.g. [179], so that there exist infinitely many solutions of $\hat{\mathbf{h}}_1(n)$ and $\hat{\mathbf{h}}_2(n)$ that lead to perfect echo cancellation of $\mathbf{h}_1(n)$ and $\mathbf{h}_2(n)$, and typically $\hat{\mathbf{h}}_1(n) \neq \mathbf{h}_1(n)$ and $\hat{\mathbf{h}}_2(n) \neq \mathbf{h}_2(n)$. Unfortunately, all solutions but the true one, where $\hat{\mathbf{h}}_1(n) = \mathbf{h}_1(n)$ and $\hat{\mathbf{h}}_2(n) = \mathbf{h}_2(n)$, depend on the far-end room impulse responses $\mathbf{g}_1(n)$ and $\mathbf{g}_2(n)$, and any change of far-end room impulse responses, e.g. the movement of talkers, would affect or even destroy the near-end echo cancellation.

In [11], it is shown that an effective solution to the non-uniqueness problem is to reduce the cross-correlation between the signals $u_1(n)$ and $u_2(n)$, and a nonlinear method of half-wave rectification for decorrelation is proposed. Since then, the stereophonic echo cancellation problem has been extensively studied, different suggestions for decorrelation have been proposed, such as perceptually shaped noise signals [47, 56], time-varying all-pass filters [3, 188, 189], time-reversal of signals [141], different types of nonlinearities [133], phase modification based methods [14, 82], and methods [26, 160, 161] based on the psychoacoustic phenomenon of the missing fundamental [75].

Hence, although the underlying reasons are different, decorrelation techniques are useful in acoustic echo cancellation as well as in acoustic feedback cancellation. Unfortunately, methods which are effective for decorrelation for one system might not be effective for the other. It is shown in [197], that the half-wave rectification which is effective against the non-uniqueness problem in stereophonic echo cancellation is not effective against the biased estimation problem in acoustic feedback cancellation. Hence, decorrelation should be designed and verified specifically for each cancellation system.

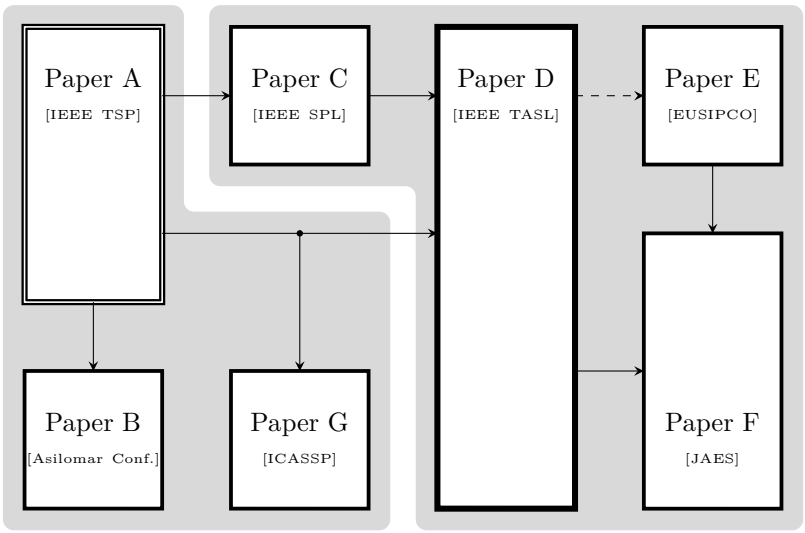


Fig. 9: An overview of relations between the included papers. Shaded areas indicate the two main topics in this work. Arrows with solid/dashed lines indicate direct/indirect connections between papers. Paper A is the initial point for our work (indicated by the double lined box), and Paper D is the most significant contribution (indicated by the thick solid box). Paper outline: A. Power transfer function method [63]. B. Estimation Strategies [62]. C. Biased estimation and traditional probe noise approach [66]. D. A novel probe noise approach for unbiased estimation [69]. E. Frequency shifting and phase modulation methods for unbiased estimation [71]. F. Comparison of different state-of-the-art cancellation systems [70]. G. Power transfer function refinement [72].

5 Topics of the Thesis

The main part of this thesis consists of a collection of selected papers, contributing to the development of acoustic feedback and echo cancellation systems. In this section, first we provide an overview of the relations between these papers, then we describe each paper in more details and highlight the main contributions.

5.1 Outline

Fig. 9 shows the relations between the included papers. There are two main topics in this work. First, we analyze an acoustic feedback and/or echo cancellation system in a multiple microphone audio processing system to predict cancellation system behavior. Secondly, we use this analytical method to evaluate the biased estimation problem, which is perhaps the biggest remaining problem in the field of acoustic feedback cancellation. Based on this evaluation result, we propose a novel probe noise based acoustic

feedback cancellation system for a hearing aid application to solve the biased estimation problem.

Paper A is our initial work, where we propose a frequency domain evaluation measure referred to as the power transfer function (PTF) to predict system behavior for both feedback and echo cancellation systems. The PTF expression is derived as a function of different signal properties and the applied adaptive algorithms, but it does not require knowledge of true acoustic feedback/echo path. In Paper B, the PTF method is used to determine optimal adaptive cancellation strategies for feedback and echo cancellation systems. The PTF is further refined in Paper G for a more accurate evaluation and prediction of system behavior in acoustic feedback cancellation systems.

The PTF method is used in Papers C and D to analyze a traditional probe noise approach and design a novel probe noise based cancellation system, respectively. The novel probe noise approach is able to provide an unbiased estimation, and a significantly increased convergence rate compared to the traditional probe noise approach, while only introducing minor noticeable but not annoying artifacts. In Paper E, we further analyze and compare the frequency shifting and phase modulation methods with perceptually motivated parameter setups, which ensure minimal sound distortions, to deal with the biased estimation problem. In Paper F, we conduct a comparison of different state-of-the-art feedback cancellation systems in a hearing aid application including the methods described in Papers D and E. At an insignificant sound quality degradation level, the method proposed in Paper D turns out to have the best overall cancellation performance, with only a relatively small computational complexity increase.

5.2 Summary of Contributions

Paper A – Analysis of Acoustic Feedback/Echo Cancellation in Multiple-Microphone and Single-Loudspeaker Systems Using a Power Transfer Function Method

In Paper A, we analyze a general multiple-microphone and single-loudspeaker audio processing system, where a multichannel adaptive system is used to cancel the effect of acoustic feedback/echo, and a beamformer processes the feedback/echo canceled signals. We introduce and derive an accurate approximation of a frequency domain measure—the power transfer function—and show how it can be used to predict the convergence rate, steady-state behavior, and the system stability bound of the entire cancellation system across frequency and time. Furthermore, we derive expressions to determine the step size parameter in the adaptive algorithms to achieve a desired system behavior, e.g. convergence rate at a specific frequency. Different parts of this work have been published in preliminary form in [64, 65].

Paper B – Acoustic Feedback and Echo Cancellation Strategies for Multiple-Microphone and Single-Loudspeaker Systems

This work is motivated by the fact that often acoustic feedback/echo cancellation in a multiple-microphone and single-loudspeaker system is carried out using a cancellation filter for each microphone channel, and the filters are adaptively estimated, independently of each other. Hence, we consider another strategy by estimating all the cancellation filters jointly and assess if an improved cancellation performance is achievable compared to the independent estimation strategy.

We show, using the power transfer function method introduced in Paper A, that under certain reasonable assumptions the independent estimation strategy is statistically identical to the joint estimation strategy. Hence, there is no performance advantages by using the computationally more complex joint estimation strategy in the considered system. Furthermore, we relate the joint estimation strategy to a stereophonic acoustic echo cancellation system and provide analytic expressions for its system behavior.

Paper C – On Acoustic Feedback Cancellation Using Probe Noise in Multiple-Microphone and Single-Loudspeaker Systems

In Paper C, we focus on a traditional probe noise approach to prevent biased estimation in a feedback cancellation system, as discussed in Sec. 2.3. Although the traditional probe noise approach is effective against the bias problem, practical experiences and simulation results indicated that whenever a low-level and inaudible probe noise signal is used, the convergence rate of the adaptive estimation is significantly decreased when keeping the steady-state error unchanged.

In this work, we show theoretically how different system parameters and signal properties affect the cancellation performance, and the results explain the decreased convergence rate from a theoretical point of view. Understanding this was important for making further improvements (as presented in Paper D) to the traditional probe noise approach.

Paper D – Novel Acoustic Feedback Cancellation Approaches in Hearing Aid Applications Using Probe Noise and Probe Noise Enhancement

In Paper D, based on the knowledge obtained from Paper C, we propose and study analytically two new probe noise approaches utilizing a combination of specifically designed probe noise signals and probe noise enhancement filters. Despite using low-level and inaudible probe noise signals, both approaches significantly improve the convergence behavior of the cancellation system compared to the traditional probe noise approach.

In particular, we utilize a simple spectral masking model to generate a probe noise signal $w(n)$, which is inaudible in the presence of the original loudspeaker signal $u(n)$. The improvements in convergence rate are obtained by processing the signals entering

the adaptive algorithms using enhancement filters specifically designed as long-term prediction error filters, so that the disturbance from the incoming signals $x(n)$ is reduced, whereas the probe noise signal properties are unmodified, and it thereby increases the probe noise to disturbing signal ratio. Part of this work has been published in preliminary form in [67], whereas its application in hearing aids has been presented in [68].

Paper E – On the Use of a Phase Modulation Method for Decorrelation in Acoustic Feedback Cancellation

In Paper E, we consider an otherwise well-known phase modulation decorrelation method for reducing the biased estimation problem in feedback cancellation systems. However, we configure the parameter setting for the phase modulation over frequency in a perceptually motivated way so that the sound quality degradation is at a very low level. We determine if this configuration is effective for decorrelation in acoustic feedback cancellation systems by comparing it to a structurally similar frequency shifting decorrelation method. We show that the phase modulation method with the specific perceptually motivated parameter choices is suitable for decorrelation in a hearing aid acoustic feedback cancellation system, although the frequency shifting method in general is slightly more effective.

Paper F – Evaluation of State-of-the-Art Acoustic Feedback Cancellation Systems for Hearing Aids

In Paper F, we evaluate four state-of-the-art acoustic feedback cancellation systems used and/or proposed for a hearing aid application including the novel probe noise based system, described in Paper D, in terms of their abilities to cancel acoustic feedback, additional sound quality degradations they might introduce to hearing aid output signals due to decorrelation, and their computational complexity.

All these four state-of-the-art cancellation systems outperform the traditional cancellation system which has significant limitations due to the biased estimation problem, and the computational complexity increases are no more than a factor of three. Furthermore, we show that especially the novel probe noise based system is most effective for cancellation and robust against the biased estimation in the case of highly correlated incoming signals like music.

Paper G – Analysis of Closed-Loop Acoustic Feedback Cancellation Systems

In Paper G, we propose a refinement to the power transfer function described in Paper A. The analysis in Paper A is derived in an open-loop acoustic echo cancellation system, and it provides inaccurate predictions in closed-loop acoustic feedback cancellation systems if there is a *strong* correlation between the loudspeaker signal and the signals entering the microphones.

This work extends the power transfer function performance analysis by studying the effects of the nonzero signal correlation on adaptive filters, and the extension provides accurate performance predictions in closed-loop acoustic feedback cancellation systems.

6 Conclusions and Future Directions

This thesis treats the analysis, design, and evaluation of acoustic feedback and/or echo cancellation systems. The first main contribution is the analysis of a general multiple-microphone and single-loudspeaker acoustic feedback/echo cancellation system using an introduced frequency domain evaluation measure referred to as the power transfer function in Paper A. This measure can be used to predict system behaviors such as convergence rate and/or steady-state error in acoustic feedback and echo cancellation systems, and it can be used to set system parameters to obtain specifically desired cancellation properties.

The second main contribution is the design and evaluation of a novel probe noise based approach for hearing aid acoustic feedback cancellation described in Paper D and Paper F, respectively. We have shown in Paper F that the proposed probe noise based feedback cancellation system outperforms other state-of-the-art hearing aid feedback cancellation systems, especially in the most difficult-to-handle situation.

Although the contributions in this thesis solve some of the major problems in acoustic feedback cancellation for hearing aids, some challenges still remain. For example, the novel probe noise based system has an increase in computational complexity by a factor of roughly three compared to the traditional acoustic feedback cancellation system. Due to the limited computing power available in hearing aids, a complexity reduction is therefore preferable before it is realized for practical use. Furthermore, it is also important to verify how the proposed probe noise based system perform in practice, e.g. in a more complicated acoustic environment, which makes the enhancement filter estimation more challenging, as the incoming signal is very probably a mixture of different signals including background noise etc. Another interesting research question regards the possible interactions between the proposed probe noise based cancellation system and other hearing aid signal processing algorithms such as adaptive beamformer/noise reduction algorithms. These are topics for future work.

Moreover, the proposed probe noise system is specifically designed for hearing aids, utilizing their short acoustic feedback paths. Therefore, we expect that it is useful for other applications with short acoustic feedback/echo paths such as headsets. Another topic for future work is to investigate if this system is useful (or should be modified) for acoustic feedback cancellation in public address systems and/or acoustic echo cancellation in general, where the feedback/echo paths are typically much longer.

References

[1] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: Analysis and simulations," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 631–639, Mar. 1997.

[2] J. Agnew, "Acoustic feedback and other audible artifacts in hearing aids," *Trends in Amplification*, vol. 1, no. 2, pp. 45–82, Jun. 1996.

[3] M. Ali, "Stereophonic acoustic echo cancellation system using time-varying all-pass filtering for signal decorrelation," in *Proc. 1998 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, May 1998, pp. 3689–3692.

[4] J. Alisobhani and S. Knorr, "Improvement of acoustic-feedback stability by bandwidth compression," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 6, pp. 636–644, Dec. 1980.

[5] B. D. O. Anderson and M. R. Gevers, "Identifiability of linear stochastic systems operating under linear feedback," *Elsevier Automatica*, vol. 18, no. 2, pp. 195–213, Mar. 1982.

[6] S. Ando, "Howling detection and prevention circuit and a loudspeaker system employing the same," United States Patent, US 6,252,969 B1, Jun. 2001.

[7] Y. Avargel and I. Cohen, "Adaptive system identification in the short-time Fourier transform domain using cross-multiplicative transfer function approximation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 1, pp. 162–173, Jan. 2008.

[8] S. Bech and N. Zacharov, *Perceptual Audio Evaluation – Theory, Method and Application.* Hoboken, NJ, US: Wiley, Jun. 2006.

[9] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation.* Berlin, Heidelberg, Germany: Springer, May 2001.

[10] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. 2002 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, May 2002, pp. 1881–1884.

[11] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 2, pp. 156–165, Mar. 1998.

[12] J. Benesty, C. Paleologu, and S. Ciochină, "On regularization in adaptive filtering," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1734–1742, Aug. 2011.

[13] E. Berdahl and D. Harris, "Frequency shifting for acoustic howling suppression," in *Proc. 13th Int. Conf. Digit. Audio Effects*, Sep. 2010, pp. 174–177.

[14] B. C. Bispo and D. da S. Freitas, "Hybrid pre-processor based on frequency shifting for stereophonic acoustic echo cancellation," in *Proc. 20th European Signal Process. Conf.*, Aug. 2012, pp. 2447–2451.

[15] E. Bjarnason, "Analysis of the filtered-X LMS algorithm," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 6, pp. 504–514, Nov. 1995.

[16] J. B. Boldt and T. B. Elmedyb, "Online anti-feedback system for a hearing aid," European Patent Application, EP 2003928 A1, Jun. 2007.

[17] C. P. Boner and C. R. Boner, "Minimizing feedback in sound systems and room-ring modes with passive networks," *J. Acoust. Soc. Am.*, vol. 37, no. 1, pp. 131–135, Jan. 1965.

[18] ——, "A procedure for controlling room-ring modes and feedback modes in sound systems with narrow-band filters," *J. Audio Eng. Soc.*, vol. 13, no. 4, pp. 297–299, Oct. 1965.

[19] C. Boukis, D. P. Mandic, and A. G. Constantinides, "Toward bias minimization in acoustic feedback cancellation systems," *J. Acoust. Soc. Am.*, vol. 121, no. 3, pp. 1529–1537, Mar. 2007.

[20] L. Bramsløw, "An objective estimate of the perceived quality of reproduced sound in normal and impaired hearing," *Acta Acustica united with Acustica*, vol. 90, no. 6, pp. 1007–1018, Nov./Dec. 2004.

[21] ——, "Preferred signal path delay and high-pass cut-off in open fittings," *Int. J. Audiology*, vol. 49, no. 9, pp. 634–644, Sep. 2010.

[22] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Heidelberg, Germany: Springer, Jun. 2001.

[23] M. D. Burkhard, "A simplified frequency shifter for improving acoustic feedback stability," *J. Audio Eng. Soc.*, vol. 11, no. 3, pp. 234–237, Jul. 1963.

[24] D. K. Bustamante, T. L. Worrall, and M. J. Williamson, "Measurement and adaptive suppression of acoustic feedback in hearing aids," in *Proc. 1989 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, May 1989, pp. 2017–2020.

[25] H. Cao, J. Liu, and W. Zhang, "A combined de-correlation method for acoustic feedback cancellation in hearing aids," in *Proc. 2009 World Congress Comput. Sci. Inform. Eng.*, vol. 7, Mar. 2009, pp. 220–224.

[26] S. Cecchi, L. Romoli, P. Peretti, and F. Piazza, "A combined psychoacoustic approach for stereo acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1530–1539, Aug. 2011.

[27] A. P. Chandrakasan and R. W. Brodersen, "Minimizing power consumption in digital CMOS circuits," *Proc. IEEE*, vol. 83, no. 4, pp. 498–523, Apr. 1995.

[28] C. Chen, "Detection of acoustical feedback in a public address system," in *Proc. 1977 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, May 1977, pp. 385–388.

[29] H.-F. Chi, S. X. Gao, S. D. Soli, and A. Alwan, "Band-limited feedback cancellation with a modified filtered-X LMS algorithm for hearing aids," *Elsevier Speech Commun.*, vol. 39, no. 1-2, pp. 147–161, Jan. 2003.

[30] S. Cifani, L. C. Montesi, R. Rotili, E. Principi, S. Squartini, and F. Piazza, "A PEM-AFROW based algorithm for acoustic feedback control in automotive speech reinforcement systems," in *Proc. 6th Int. Symp. Image Signal Process. Anal.*, Sep. 2009, pp. 656–661.

[31] J. Cioffi, "Limited-precision effects in adaptive filtering," *IEEE Trans. Circuits Syst.*, vol. 34, no. 7, pp. 821–833, Jul. 1987.

[32] J. Cioffi and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 304–337, Apr. 1984.

[33] R. B. Coleman, E. F. Berkman, and B. G. Watters, "Optimal probe-noise generation for on-line plant identification within filtered-X LMS controllers," *Active Control Vibration Noise*, vol. 75, pp. 1–6, 1994.

[34] H. Dillon, *Hearing Aids.* Stuttgart, Germany: Thieme, May 2001.

[35] T. K. Duong, E. Lefort, and M. G. Bellanger, "Acoustic feedback cancelling electro-acoustic transducer network," United States Patent, US 4,485,272, Nov. 1984.

[36] D. L. Duttweiler, "A twelve-channel digital echo canceler," *IEEE Trans. Commun.*, vol. 26, no. 5, pp. 647–653, May 1978.

[37] ——, "Adaptive filter performance with nonlinearities in the correlation multiplier," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 30, no. 4, pp. 578–586, Aug. 1982.

[38] ——, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, Sep. 2000.

[39] G. W. Elko and M. M. Goodwin, "Beam dithering: Acoustic feedback control using a modulated-directivity loudspeaker array," in *Proc. 1993 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, Apr. 1993, pp. 173–176.

[40] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2937–2944, Nov. 1994.

[41] J. B. Foley, "Adaptive periodic noise cancellation for the control of acoustic howling," in *Proc. IEE Colloq. Adaptive Filters*, Mar. 1989, pp. 7/1–7/4.

[42] U. Forsell and L. Ljung, "Closed-loop identification revisited," *Elsevier Automatica*, vol. 35, no. 7, pp. 1215–1241, Jul. 1999.

[43] ——, "A projection method for closed-loop identification," *IEEE Trans. Autom. Control*, vol. 45, no. 11, pp. 2101–2106, Nov. 2000.

[44] D. J. Freed, "Adaptive feedback cancellation in hearing aids with clipping in the feedback path," *J. Acoust. Soc. Am.*, vol. 123, no. 3, pp. 1618–1626, Mar. 2008.

[45] D. J. Freed and S. D. Soli, "An objective procedure for evaluation of adaptive antifeedback algorithms in hearing aids," *Ear Hearing*, vol. 27, no. 4, pp. 382–398, Aug. 2006.

[46] I. Furukawa, "A design of canceller for broad band acoustic echo," in *Proc. Int. Teleconf. Symp.*, 1984, pp. 232–239.

[47] T. Gänsler and P. Eneroth, "Influence of audio coding on stereophonic acoustic echo cancellation," in *Proc. 1998 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, May 1998, pp. 3649–3652.

[48] T. Gänsler, S. L. Gay, M. M. Sondhi, and J. Benesty, "Double-talk robust fast converging algorithms for network echo cancellation," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 6, pp. 656–663, Nov. 2000.

[49] T. Gänsler, M. Hansson, C.-J. Ivarsson, and G. Salomonsson, "A double-talk detector based on coherence," *IEEE Trans. Commun.*, vol. 44, no. 11, pp. 1421–1427, Nov. 1996.

[50] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," in *Proc. 32nd Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Nov. 1998, pp. 394–398.

[51] S. L. Gay and S. Tavathia, "The fast affine projection algorithm," in *Proc. 1995 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, May 1995, pp. 3023–3026.

[52] S. L. Gay and J. Benesty, *Acoustic Signal Processing for Telecommunication*. Berlin, Heidelberg, Germany: Springer, Mar. 2000.

[53] A. Gersho, "Adaptive filtering with binary reinforcement," *IEEE Trans. Inf. Theory*, vol. 30, no. 2, pp. 191–199, Mar. 1984.

[54] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Transform domain prediction error method for improved acoustic echo and feedback cancellation," in *Proc. 20th European Signal Process. Conf.*, Aug. 2012, pp. 2422–2426.

[55] P. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Regularized adaptive notch filters for acoustic howling suppression," in *Proc. 17th European Signal Process. Conf.*, Aug. 2009, pp. 2574–2578.

[56] A. Gilloire and V. Turbin, "Using auditory properties to improve the behaviour of stereophonic acoustic echo cancellers," in *Proc. 1998 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, May 1998, pp. 3681–3684.

[57] A. H. Gray, Jr. and J. D. Markel, "Distance measures for speech processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 5, pp. 380–391, Oct. 1976.

[58] J. E. Greenberg, P. M. Zurek, and M. Brantley, "Evaluation of feedback-reduction algorithms for hearing aids," *J. Acoust. Soc. Am.*, vol. 108, no. 5, pp. 2366–2376, Nov. 2000.

[59] S. Gunnarsson, "On the quality of recursively identified FIR models," *IEEE Trans. Signal Process.*, vol. 40, no. 3, pp. 679–682, Mar. 1992.

[60] S. Gunnarsson and L. Ljung, "Frequency domain tracking characteristics of adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 1072–1089, Jul. 1989.

[61] M. Guo, "Hearing device with adaptive feedback suppression," European Patent Application, EP 2217007 A1, Feb. 2009.

[62] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Acoustic feedback and echo cancellation strategies for multiple-microphone and single-loudspeaker systems," in *Proc. 45th Asilomar Conf. Signals, Syst., Comput.*, Nov. 2011, pp. 556–560.

[63] ——, "Analysis of acoustic feedback/echo cancellation in multiple-microphone and single-loudspeaker systems using a power transfer function method," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5774–5788, Dec. 2011.

[64] ——, "Analysis of adaptive feedback and echo cancelation algorithms in a general multiple-microphone and single-loudspeaker system," in *Proc. 2011 IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 433–436.

[65] ——, "Comparison of multiple-microphone and single-loudspeaker adaptive feedback/echo cancellation systems," in *Proc. 19th European Signal Process. Conf.*, Aug. 2011, pp. 1279–1283.

[66] ——, "On acoustic feedback cancellation using probe noise in multiple-microphone and single-loudspeaker systems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 283–286, May 2012.

[67] M. Guo, S. H. Jensen, and J. Jensen, "An improved probe noise approach for acoustic feedback cancellation," in *Proc. 7th IEEE Sensor Array Multichannel Signal Process. Workshop*, Jun. 2012, pp. 497–500.

[68] ——, "A new probe noise approach for acoustic feedback cancellation in hearing aids," in *Abstract 2012 Int. Hearing Aid Research Conf.*, Aug. 2012, p. 61.

[69] ——, "Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancements," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2549–2563, Nov. 2012.

[70] ——, "Evaluation of state-of-the-art acoustic feedback cancellation systems for hearing aids," *J. Audio Eng. Soc.*, to be published, 2013.

[71] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "On the use of a phase modulation method for decorrelation in acoustic feedback cancellation," in *Proc. 20th European Signal Process. Conf.*, Aug. 2012, pp. 2000–2004.

[72] ——, "Analysis of closed-loop acoustic feedback cancellation systems," in *Proc. 2013 IEEE Int. Conf. Acoust., Speech, Signal Process.*, to be published, May 2013.

[73] I. Gustavsson, L. Ljung, and T. Söderström, "Identification of processes in closed loop – identifiability and accuracy aspects," *Elsevier Automatica*, vol. 13, no. 1, pp. 59–75, Jan. 1977.

[74] E. Hänsler, "The hands-free telephone problem – an annotated bibliography," *Elsevier Signal Process.*, vol. 27, no. 3, pp. 259–271, Jun. 1992.

[75] W. M. Hartmann, *Signals, Sound, and Sensation*. Berlin, Heidelberg, Germany: Springer, Jan. 1997.

[76] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, US: Prentice Hall, Sep. 2001.

[77] J. Hellgren, "Compensation for hearing loss and cancellation of acoustic feedback in digital hearing aids," Ph.D. dissertation, Linköpings Universitet, 2000.

[78] ——, "Analysis of feedback cancellation in hearing aids with filtered-X LMS and the direct method of closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 2, pp. 119–131, Feb. 2002.

[79] J. Hellgren and U. Forssell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 906–913, Nov. 2001.

[80] J. Hellgren, T. Lunner, and S. Arlinger, "System identification of feedback in hearing aids," *J. Acoust. Soc. Am.*, vol. 105, no. 6, pp. 3481–3496, Jun. 1999.

[81] ——, "Variations in the feedback of hearing aids," *J. Acoust. Soc. Am.*, vol. 106, no. 5, pp. 2821–2833, Nov. 1999.

[82] J. Herre, H. Buchner, and W. Kellermann, "Acoustic echo cancellation for surround sound using perceptually motivated convergence enhancement," in *Proc. 2007 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, Apr. 2007, pp. 17–20.

[83] J. Huo, S. Nordholm, and Z. Zang, "New weight transform schemes for delayless subband adaptive filtering," in *Proc. IEEE Global Telecommunications Conf.*, vol. 1, Nov. 2001, pp. 197–201.

[84] ITU-R Recommendation BS.1116-1, *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*, Int. Telecommun. Union, 1994.

[85] ITU-R Recommendation BS.1387-1, *Method for objective measurements of perceived audio quality*, Int. Telecommun. Union, 1998.

[86] ITU-R Recommendation BS.1534-1, *Method for subjective assessment of intermediate quality level of coding systems*, Int. Telecommun. Union, 2001.

[87] ITU-T Recommendation P.800, *Methods for subjective determination of transmission quality*, Int. Telecommun. Union, 1996.

[88] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, Int. Telecommun. Union, 2001.

[89] C. P. Janse and H. J. W. Belt, "Sound reinforcement system having an echo suppressor and loudspeaker beamformer," United States Patent, US 7,054,451 B2, May 2006.

[90] C. P. Janse and C. C. Tchang, "Acoustic feedback suppression," Int. Patent Application, WO 2005/079109 A1, Aug. 2005.

[91] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 2, pp. 314–323, Feb. 1988.

[92] H. A. L. Joson, F. Asano, Y. Suzuki, and T. Sone, "Adaptive feedback cancellation with frequency compression for hearing aids," *J. Acoust. Soc. Am.*, vol. 94, no. 6, pp. 3248–3254, Dec. 1993.

[93] A. Kaelin, A. Lindgren, and S. Wyrsch, "A digital frequency-domain implementation of a very high gain hearing aid with compensation for recruitment of loudness and acoustic echo cancellation," *Elsevier Signal Process.*, vol. 64, no. 1, pp. 71–85, Jan. 1998.

[94] J. M. Kates, "A computer simulation of hearing aid response and the effects of ear canal size," *J. Acoust. Soc. Am.*, vol. 83, no. 5, pp. 1952–1963, May 1988.

[95] ——, "A time-domain digital simulation of hearing aid response," *J. Rehabil. Research Develop.*, Summer 1990.

[96] ——, "Feedback cancellation in hearing aids: Results from a computer simulation," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 553–562, Mar. 1991.

[97] ——, "Constrained adaptation for feedback cancellation in hearing aids," *J. Acoust. Soc. Am.*, vol. 106, no. 2, pp. 1010–1019, Aug. 1999.

[98] ——, "Feedback cancellation apparatus and methods," United States Patent, US 6,072,884, Jun. 2000.

[99] ——, "Room reverberation effects in hearing aid feedback cancellation," *J. Acoust. Soc. Am.*, vol. 109, no. 1, pp. 367–378, Jan. 2001.

[100] ——, "Feedback cancellation in a hearing aid with reduced sensitivity to low-frequency tonal inputs," United States Patent, US 6,831,986 B2, Dec. 2004.

[101] ——, *Digital Hearing Aids*. San Diego, CA, US: Plural Publishing Inc., Mar. 2008.

[102] T. Kaulberg, "A hearing aid with an adaptive filter for suppression of acoustic feedback," European Patent, EP 1191814 B1, Mar. 2002.

[103] W. Kellermann, "Kompensation akustischer Echos in Frequenzteilbändern," *Frequenz*, vol. 39, no. 7-8, pp. 209–215, Jul. 1985.

[104] A. W. H. Khong and P. A. Naylor, "Efficient use of sparse adaptive filters," in *Proc. 40th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2006, pp. 1375–1379.

[105] ——, "Selective-tap adaptive filtering with performance analysis for identification of time-varying systems," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 5, pp. 1681–1695, Jul. 2007.

[106] K. T. Klinkby, P. M. Norgaard, and H. P. Foeh, "Hearing aid, and a method for control of adaptation rate in anti-feedback systems for hearing aids," Int. Patent Application, WO 2007/113282 A1, Oct. 2007.

[107] S. Kochkin, "Marketrak v: "Why my hearing aids are in the drawer": The consumers' perspective," *Hearing J.*, vol. 53, no. 2, pp. 34–41, Feb. 2000.

[108] ——, "Marketrak viii: Consumer satisfaction with hearing aids is slowly increasing," *Hearing J.*, vol. 63, no. 1, pp. 19–27, Jan. 2010.

[109] S. Koike, "Analysis of adaptive filters using normalized signed regressor LMS algorithm," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2710–2723, Oct. 1999.

[110] ——, "A class of adaptive step-size control algorithms for adaptive filters," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1315–1326, Jun. 2002.

[111] S. M. Kuo and J. Chen, "New adaptive IIR notch filter and its application to howling control in speakerphone system," *IEEE Electron. Lett.*, vol. 28, no. 8, pp. 764–766, Apr. 1992.

[112] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1633–1642, Jul. 1992.

[113] A. Larsen, "Ein akustischer Wechselstromerzeuger mit regulierbarer Periodenzahl für schwache Ströme," *Elektrotech. Z., ETZ 32*, pp. 284–285, Mar. 1911.

[114] H.-W. Lee and M.-Y. Jeon, "A combined feedback and noise cancellation algorithm for binaural hearing aids," *J. Advances Elect. Comput. Eng.*, vol. 11, no. 3, pp. 35–40, 2011.

[115] F. Ling and J. G. Proakis, "Numerical accuracy and stability: Two problems of adaptive estimation algorithms caused by round-off error," in *Proc. 1984 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 9, Mar. 1984, pp. 571–574.

[116] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ, US: Prentice Hall, Dec. 1998.

[117] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification – a survey," *Elsevier Automatica*, vol. 26, no. 1, pp. 7–21, Jan. 1990.

[118] S. Ljung and L. Ljung, "Error propagation properties of recursive least-squares adaptation algorithms," *Elsevier Automatica*, vol. 21, no. 2, pp. 157–167, Mar. 1985.

[119] P. Loganathan, E. A. P. Habets, and P. A. Naylor, "Performance analysis of IPNLMS for identification of time-varying systems," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 317–320.

[120] P. Loganathan, A. W. H. Khong, and P. A. Naylor, "A class of sparseness-controlled algorithms for echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 8, pp. 1591–1601, Nov. 2009.

[121] A. Lombard, K. Reindl, and W. Kellermann, "Combination of adaptive feedback cancellation and binaural adaptive filtering in hearing aids," *EURASIP J. Advances Signal Process.*, vol. 2009, pp. 1–15, Apr. 2009.

[122] G. Long, F. Ling, and J. G. Proakis, "The LMS algorithm with delayed coefficient adaptation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 9, pp. 1397–1405, Sep. 1989.

[123] G. Ma, F. Gran, F. Jacobsen, and F. Agerkvist, "Using a reflection model for modeling the dynamic feedback path of digital hearing aids," *J. Acoust. Soc. Am.*, vol. 127, no. 3, pp. 1458–1468, Mar. 2010.

[124] ——, "Adaptive feedback cancellation with band-limited LPC vocoder in digital hearing aids," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 4, pp. 677–687, May 2011.

[125] ——, "Extracting the invariant model from the feedback paths of digital hearing aids," *J. Acoust. Soc. Am.*, vol. 130, no. 1, pp. 350–363, Jul. 2011.

[126] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, Apr. 1975.

[127] J. A. Maxwell and P. M. Zurek, "Reducing acoustic feedback in hearing aids," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 4, pp. 304–313, Jul. 1995.

[128] L. N. Mishin, "A method for increasing the stability of sound amplification systems," *Sov. Phys. Acoust.*, vol. 4, pp. 64–71, 1958.

[129] S. Miyagi and H. Sakai, "Performance comparison between the filtered-error LMS and the filtered-X LMS algorithms," in *Proc. 2001 IEEE Int. Symp. Ciruits Syst.*, vol. 2, May 2001, pp. 661–664.

[130] S. P. Mohanty, N. Ranganathan, E. Kougianos, and P. Patra, *Low-Power High-Level Synthesis for Nanoscale CMOS Circuits*. Berlin, Heidelberg, Germany: Springer, Jul. 2008.

[131] D. C. Montgomery, *Design and Analysis of Experiments*, 6th ed. Hoboken, NJ, US: Wiley, Dec. 2004.

[132] B. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. Bingley, UK: Emerald Group Publishing Limited, Apr. 2003.

[133] D. R. Morgan, J. L. Hall, and J. Benesty, "Investigation of several types of nonlinearities for use in stereo acoustic echo cancellation," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 6, pp. 686–696, Sep. 2001.

[134] D. R. Morgan and J. C. Thi, "A delayless subband adaptive filter architecture," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1819–1830, Aug. 1995.

[135] Y. Nagata, S. Suzuki, M. Yamada, M. Yoshida, M. Kitano, K. Kuroiwa, and S. Kimura, "Howling remover composed of adjustable equalizers for attenuating complicated noise peaks," United States Patent, US 5,729,614, Mar. 1998.

[136] ——, "Howling remover having cascade connected equalizers suppressing multiple noise peaks," United States Patent, US 5,710,823, Jan. 1998.

[137] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Dual microphone solution for acoustic feedback cancellation for assistive listening," in *Proc. 2012 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2012, pp. 149–152.

[138] V. H. Nascimento, M. T. M. Silva, L. A. Azpicueta-Ruiz, and J. Arenas-García, "On the tracking performance of combinations of least mean squares and recursive least squares adaptive filters," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 3710–3713.

[139] K. Ngo, T. van Waterschoot, M. G. Christensen, M. Moonen, S. H. Jensen, and J. Wouters, "Adaptive feedback cancellation in hearing aids using a sinusoidal near-end signal model," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 181–184.

[140] ——, "Prediction-error-method-based adaptive feedback cancellation in hearing aids using pitch estimation," in *Proc. 18th European Signal Process. Conf.*, Aug. 2010, pp. 40–44.

[141] D.-Q. Nguyen, W.-S. Gan, and A. W. H. Khong, "Time-reversal approach to the stereophonic acoustic echo cancellation problem," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 2, pp. 385–395, Feb. 2011.

[142] J. L. Nielsen and U. P. Svensson, "Performance of some linear time-varying systems in control of acoustic feedback," *J. Acoust. Soc. Am.*, vol. 106, no. 1, pp. 240–254, Jul. 1999.

[143] N. S. Nise, *Control Systems Engineering*, 6th ed. Hoboken, NJ, US: Wiley, Apr. 2011.

[144] H. Nyquist, "Regeneration theory," *Bell System Tech. J.*, vol. 11, pp. 126–147, 1932.

[145] S. F. Oberman and M. J. Flynn, "Division algorithms and implementations," *IEEE Trans. Comput.*, vol. 46, no. 8, pp. 833–854, Aug. 1997.

[146] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. IEEE*, vol. 88, no. 4, pp. 451–513, Apr. 2000.

[147] C. Paleologu, J. Benesty, and S. Ciochină, "A variable step-size affine projection algorithm designed for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 8, pp. 1466–1478, Nov. 2008.

[148] A. Pandey, V. J. Mathews, and M. Nilsson, "Adaptive gain processing to improve feedback cancellation in digital hearing aids," in *Proc. 2008 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2008, pp. 357–360.

[149] Y.-C. Park, I.-Y. Kim, and S.-M. Lee, "An efficient adaptive feedback cancellation for hearing aids," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2, Sep. 2003, pp. 1647–1650.

[150] E. T. Patronis, Jr., "Electronic detection of acoustic feedback and automatic sound system gain control," *J. Audio Eng. Soc.*, vol. 26, no. 5, pp. 323–326, May 1978.

[151] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*. Hoboken, NJ, US: Wiley, May 2001.

[152] C. J. Plack, *The Sense of Hearing*. Hove, UK: Psychology Press, Jun. 2005.

[153] M. A. Poletti, "The stability of multichannel sound systems with frequency shifting," *J. Acoust. Soc. Am.*, vol. 116, no. 2, pp. 853–871, Aug. 2004.

[154] B. Rafaely and J. L. Hayes, "On the modeling of the vent path in hearing aid systems," *J. Acoust. Soc. Am.*, vol. 109, no. 4, pp. 1747–1749, Apr. 2001.

[155] B. Rafaely and M. Roccasalva-Firenze, "Control of feedback in hearing aids – a robust filter design approach," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 6, pp. 754–756, Nov. 2000.

[156] B. Rafaely, M. Roccasalva-Firenze, and E. Payne, "Feedback path variability modeling for robust hearing aids," *J. Acoust. Soc. Am.*, vol. 107, no. 5, pp. 2665–2673, May 2000.

[157] B. Rafaely, N. A. Shusina, and J. L. Hayes, "Robust compensation with adaptive feedback cancellation in hearing aids," *Elsevier Speech Commun.*, vol. 39, no. 1-2, pp. 163–170, Jan. 2003.

[158] G. Rombouts, T. van Waterschoot, and M. Moonen, "Robust and efficient implementation of the PEM-AFROW algorithm for acoustic feedback cancellation," *J. Audio Eng. Soc.*, vol. 55, no. 11, pp. 955–966, Nov. 2007.

[159] G. Rombouts, T. van Waterschoot, K. Struyve, and M. Moonen, "Acoustic feedback cancellation for long acoustic paths using a nonstationary source model," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3426–3434, Sep. 2006.

[160] L. Romoli, S. Cecchi, L. Palestini, P. Peretti, and F. Piazza, "A novel approach to channel decorrelation for stereo acoustic echo cancellation based on missing fundamental theory," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 329–332.

[161] L. Romoli, S. Cecchi, P. Peretti, and F. Piazza, "A mixed decorrelation approach for stereo acoustic echo cancellation based on the estimation of the fundamental frequency," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 2, pp. 690–698, Feb. 2012.

[162] H. Sakai, "Analysis of an adaptive algorithm for feedback cancellation in hearing aids for sinusoidal signals," in *Proc. 18th European Conf. Circuit Theory Design*, Aug. 2007, pp. 416–419.

[163] A. H. Sayed, *Fundamentals of Adaptive Filtering.* Hoboken, NJ, US: Wiley, Jun. 2003.

[164] ——, *Adaptive Filters.* Hoboken, NJ, US: Wiley, May 2008.

[165] A. Schaub, *Digital Hearing Aids.* Stuttgart, Germany: Thieme, Jul. 2008.

[166] M. R. Schroeder, "Improvement of acoustic-feedback stability by frequency shifting," *J. Acoust. Soc. Am.*, vol. 36, no. 9, pp. 1718–1724, Sep. 1964.

[167] ——, "Improvement of feedback stability of public address systems by frequency shifting," *J. Audio Eng. Soc.*, vol. 10, no. 2, pp. 108–109, Apr. 1962.

[168] C. Schuldt, F. Lindstrom, and I. Claesson, "A low-complexity delayless selective subband adaptive filtering algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 5840–5850, Dec. 2008.

[169] W. A. Sethares, "Adaptive algorithms with nonlinear data and error functions," *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2199–2206, Sep. 1992.

[170] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 132–135, Feb. 2004.

[171] M. Shin, S. Wang, R. A. Bentler, and S. He, "New feedback detection method for performance evaluation of hearing aids," *Elsevier J. Sound Vibration*, vol. 302, no. 1-2, pp. 350–360, Apr. 2007.

[172] N. A. Shusina and B. Rafaely, "Feedback cancellation in hearing aids based on indirect closed-loop identification," in *Proc. 3rd IEEE Benelux Signal Process. Symp.*, Mar. 2002, pp. 177–180.

[173] ——, "Unbiased adaptive feedback cancellation in hearing aids by closed-loop identification," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 658–665, Mar. 2006.

[174] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

[175] M. G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 4, pp. 443–453, Jul. 2000.

[176] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*, 2nd ed. Hoboken, NJ, US: Wiley, Sep. 2005.

[177] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2811–2825, Sep. 1993.

[178] D. T. M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 92–114, Jan. 1991.

[179] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic acoustic echo cancellation – an overview of the fundamental problem," *IEEE Signal Process. Lett.*, vol. 2, no. 8, pp. 148–151, Aug. 1995.

[180] A. Spriet, S. Doclo, M. Moonen, and J. Wouters, "Feedback control in hearing aids," *Springer Handbook of Speech Processing*, pp. 979–999, Nov. 2007.

[181] A. Spriet, K. Eneman, M. Moonen, and J. Wouters, "Objective measures for real-time evaluation of adaptive feedback cancellation algorithms in hearing aids," in *Proc. 16th European Signal Process. Conf.*, Aug. 2008, pp. SS3–2.

[182] A. Spriet, M. Moonen, and J. Wouters, "Evaluation of feedback reduction techniques in hearing aids based on physical performance measures," *J. Acoust. Soc. Am.*, vol. 128, no. 3, pp. 1245–1261, Sep. 2010.

[183] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.

[184] ——, "An instrumental variable method for adaptive feedback cancellation in hearing aids," in *Proc. 2005 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, Mar. 2005, pp. 129–132.

[185] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Elsevier J. Franklin Inst.*, vol. 343, no. 6, pp. 545–573, Sep. 2006.

[186] ——, "Combined feedback and noise suppression in hearing aids," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 6, pp. 1777–1790, Aug. 2007.

[187] M. R. Stinson and G. A. Daigle, "Effect of handset proximity on hearing aid feedback," *J. Acoust. Soc. Am.*, vol. 115, no. 3, pp. 1147–1156, Mar. 2004.

[188] A. Sugiyama, Y. Joncour, and A. Hirano, "A stereo echo canceler with correct echo-path identification based on an input-sliding technique," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2577–2587, Nov. 2001.

[189] A. Sugiyama, Y. Mizuno, A. Hirano, and K. Nakayama, "A stereo echo canceller with simultaneous input-sliding and sliding-period control," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 325–328.

[190] Y. Terada and A. Murase, "Howling control device and howling control method," United States Patent, US 7,190,800 B2, Mar. 2007.

[191] L. S. Theverapperuma and J. S. Kindred, "Continuous adaptive feedback canceller dynamics," in *Proc. 49th IEEE Int. Midwest Symp. Circuits Syst.*, vol. 1, Aug. 2006, pp. 605–609.

[192] P. M. J. Van Den Hof, "Closed-loop issues in system identification," *Annu. Reviews in Control*, vol. 22, pp. 173–186, 1998.

[193] P. M. J. Van Den Hof and R. J. P. Schrama, "An indirect method for transfer function estimation from closed loop data," *Elsevier Automatica*, vol. 29, no. 6, pp. 1523–1527, Nov. 1993.

[194] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio signals using a warped all-pole near-end signal model," in *Proc. 2008 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2008, pp. 269–272.

[195] ——, "Comparison of linear prediction models for audio signals," *EURASIP J. Audio, Speech, Music Process.*, vol. 2008, p. Article ID 706935, 2008.

[196] ——, "Adaptive feedback cancellation for audio applications," *Elsevier Signal Process.*, vol. 89, no. 11, pp. 2185–2201, Nov. 2009.

[197] ——, "Assessing the acoustic feedback control performance of adaptive feedback cancellation in sound reinforcement systems," in *Proc. 17th European Signal Process. Conf.*, Aug. 2009, pp. 1997–2001.

[198] ——, "Comparative evaluation of howling detection criteria in notch-filter-based howling suppression," *J. Audio Eng. Soc.*, vol. 58, no. 11, pp. 923–940, Nov. 2010.

[199] ——, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

[200] T. van Waterschoot, G. Rombouts, and M. Moonen, "Optimally regularized adaptive filtering algorithms for room acoustic signal enhancement," *Elsevier Signal Process.*, vol. 88, no. 3, pp. 594–611, Mar. 2008.

[201] T. van Waterschoot, G. Rombouts, P. Verhoeve, and M. Moonen, "Double-talk-robust prediction error identification algorithms for acoustic echo cancellation," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 846–858, Mar. 2007.

[202] S. Wardle, "A Hilbert-transform frequency shifter for audio," in *Proc. Digit. Audio Effects Workshop*, Nov. 1998, pp. 25–29.

[203] R. V. Waterhouse, "Theory of howlback in reverberant rooms," *J. Acoust. Soc. Am.*, vol. 37, no. 5, pp. 921–923, 1965.

[204] R. L. Weaver and O. I. Lobkis, "On the linewidth of the ultrasonic Larsen effect in a reverberant body," *J. Acoust. Soc. Am.*, vol. 120, no. 1, pp. 102–109, Jul. 2006.

[205] A. Weiss and D. Mitra, "Digital adaptive filters: Conditions for convergence, rates of convergence, effects of noise and errors arising from the implementation," *IEEE Trans. Inf. Theory*, vol. 25, no. 6, pp. 637–652, Nov. 1979.

[206] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *IRE WESCON Conv. Record, Part 4*, pp. 96–104, 1960.

[207] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Upper Saddle River, NJ, US: Prentice Hall, Mar. 1985.

[208] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, no. 8, pp. 1151–1162, Aug. 1976.

[209] S. Wyrsch and A. Kaelin, "Adaptive feedback cancelling in subbands for hearing aids," in *Proc. 1999 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, Mar. 1999, pp. 921–924.

[210] ——, "Subband signal processing for hearing aids," in *Proc. 1999 IEEE Int. Symp. Ciruits Syst.*, vol. 3, Jul. 1999, pp. 29–32.

[211] Z. Yang, Y. R. Zheng, and S. L. Grant, "Proportionate affine projection sign algorithms for network echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 8, pp. 2273–2284, Nov. 2011.

Paper A

Analysis of Acoustic Feedback/Echo Cancellation in Multiple–Microphone and Single–Loudspeaker Systems Using a Power Transfer Function Method

Meng Guo, Thomas Bo Elmedyb, Søren Holdt Jensen, and Jesper Jensen

Published in
IEEE Trans. Signal Process., vol. 59, no. 12, pp. 5774–5788, Dec. 2011.

The layout has been revised.

Analysis of Acoustic Feedback/Echo Cancellation in Multiple-Microphone and Single-Loudspeaker Systems Using a Power Transfer Function Method

Meng Guo, Thomas Bo Elmedyb, Søren Holdt Jensen, and Jesper Jensen

Abstract

In this work, we analyze a general multiple-microphone and single-loudspeaker audio processing system, where a multichannel adaptive system is used to cancel the effect of acoustic feedback/echo, and a beamformer processes the feedback/echo canceled signals. We introduce and derive an accurate approximation of a frequency domain measure—the power transfer function—and show how it can be used to predict the convergence rate, system stability bound and the steady-state behavior of the entire cancellation system across frequency and time. We consider three example adaptive algorithms in the cancellation system: the least mean square, normalized least mean square and the recursive least squares algorithms. Furthermore, we derive expressions to determine the step size parameter in the adaptive algorithms to achieve a desired system behavior, e.g. convergence rate at a specific frequency. Finally, we compare and discuss the performance of all three adaptive algorithms, and we verify the derived expressions through simulation experiments.

1 Introduction

Acoustic feedback/echo problems occur in audio systems/devices with simultaneous recording and playback, where the microphones pick up part of the output signal from the loudspeakers. In applications such as public address systems and hearing aids, the acoustic feedback problem often degrades the system performance. In the worst-case, the systems become unstable and howling occurs as a consequence. Acoustic echo problems often occur in telephony and teleconferencing systems, where users hear their own voices as disturbing echoes.

During the past half-century, many approaches have been proposed to minimize the effects of the acoustic feedback/echo problems such as gain reduction, phase modification and frequency shifting/transposition, e.g. in [1–3]. A widely used solution for reducing the effect of this problem is the acoustic feedback cancellation (AFC) and acoustic

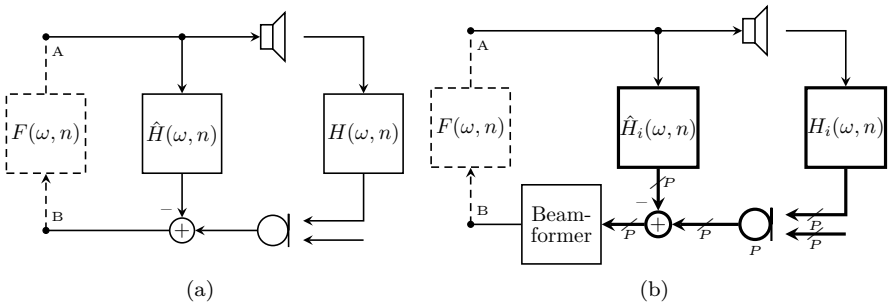


Fig. 1: Systems with acoustic feedback/echo and the cancellation. (a) A basic single-microphone and single-loudspeaker system. (b) A multiple-microphone and single-loudspeaker system with a beamformer, where $i = 1, \ldots, P$, and P denotes the number of microphone channels.

echo cancellation (AEC) algorithms which identify the acoustic feedback/echo paths by means of an adaptive filter in a system identification configuration, see e.g. [4–10].

Fig. 1(a) shows a single-microphone and single-loudspeaker (SMSL) audio processing system. An acoustic feedback/echo path is represented by the transfer function (TF) $H(\omega, n)$, where ω and n denote the discrete normalized frequency and the discrete-time index, respectively. An estimate $\hat{H}(\omega, n)$ of $H(\omega, n)$ is computed in the AFC/AEC system by means of an adaptive filter algorithm in order to cancel the effect of $H(\omega, n)$. The TF $F(\omega, n)$ denotes a forward path, which is found in closed-loop AFC applications, e.g. to implement a frequency dependent amplification in a hearing aid [11]; on the other hand, in the area of AEC, $F(\omega, n)$ represents a far-end TF and is usually ignored, resulting in an open-loop setup.

The adaptive filter approach was firstly developed in 1960s in the area of telecommunication [12]. Since then, a vast range of different adaptive algorithms have been proposed including the least mean square (LMS), normalized least mean square (NLMS), affine projection (AP), recursive least squares (RLS) and Kalman filter to mention a few [4, 5].

Many studies exist which analyze, characterize and improve adaptive algorithms in terms of e.g. robustness, stability bounds, convergence rate and steady-state behavior, see e.g. [13–22] and the references therein. Often, the analysis focuses on criteria such as mean-square error, mean-square deviation [4, 5] and variations of these. Although these criteria provide useful information about the behavior of the adaptive systems, they do not reveal frequency domain behavior, which could be more suitable in areas such as AFC and AEC, because the electro-acoustic properties of feedback/echo paths are easier described in the frequency domain in terms of the magnitude and phase spectra, and because in connection with sound quality assessment of the cancellation performance, a frequency domain measure is more directly linked to human auditory perception [23]. Some examples of frequency domain criteria can be found in e.g. [24].

In the following, we discuss a frequency domain criterion for characterizing both closed-loop AFC and open-loop AEC systems. In closed-loop systems such as hearing aids, the open-loop transfer function (OLTF) describes the system stability. In the example given in Fig. 1(a), the OLTF at a particular frequency ω and time instant n is determined by $\text{OLTF}(\omega, n) = F(\omega, n)(H(\omega, n) - \hat{H}(\omega, n))$. The stability in the system is determined by the OLTF according to the Nyquist stability criterion [25], which states that a linear and time-invariant closed-loop system becomes unstable whenever the following two criteria are both fulfilled:

$$1.\ |\text{OLTF}(\omega, n)| \geq 1;\ \ 2.\ \angle\text{OLTF}(\omega, n) = l2\pi,\ \ l = \mathbb{Z}. \tag{1}$$

In practice, the $\text{OLTF}(\omega, n)$ can not be calculated directly due to the unknown feedback/echo path $H(\omega, n)$. Instead, we express the expected magnitude-squared OLTF by

$$E\left[|\text{OLTF}(\omega, n)|^2\right] = |F(\omega, n)|^2 \xi(\omega, n), \tag{2}$$

where $\xi(\omega, n)$ denotes the expected magnitude-squared TF from point A to B in Fig. 1(a). We refer to $\xi(\omega, n)$ as the power transfer function (PTF).

As given in Eq. (1), there are two criteria for the system stability. However, we ignore the phase information in Eq. (2) because we consider a worst-case scenario for the system stability, namely, by assuming $\angle\text{OLTF}(\omega, n) = l2\pi$, where $l = \mathbb{Z}$, at all frequencies and times.

If the PTF $\xi(\omega, n)$ could be identified, then $E[|\text{OLTF}(\omega, n)|^2]$ would follow trivially, because in many closed-loop applications such as hearing aids, the forward path $F(\omega, n)$ is observable and can simply be added to $\xi(\omega, n)$ in order to determine $E[|\text{OLTF}(\omega, n)|^2]$. In the area of AEC, the influence of the far-end TF $F(\omega, n)$ is minimal, assuming that an acoustic echo cancellation system is applied at the far-end. Hence, the PTF $\xi(\omega, n)$ itself reveals the echo cancellation performance, over time and frequency, of the entire system. Therefore, in both the AFC and AEC systems, we are interested in the PTF $\xi(\omega, n)$. Ideally, with perfect cancellation $\hat{H}(\omega, n) = H(\omega, n)$, we would have $\xi(\omega, n) = 0$ for all frequencies. In practice, the PTF $\xi(\omega, n)$ is stochastic.

The goal of this work is to derive simple expressions for the PTF $\xi(\omega, n)$ in a unified framework of a multiple-microphone and single-loudspeaker (MMSL) system, whereas a conventional linear beamformer [26], performing spatial filtering of the incoming signals, processes the feedback canceled signals as illustrated in Fig. 1(b). We show that it is possible to derive a simple expression for the PTF $\xi(\omega, n)$ which allows prediction of the system behavior, without the knowledge of $H_i(\omega, n)$, as a function of system parameters, e.g. the estimation filter order, adaptive cancellation algorithm, assumptions of the feedback/echo path changes and the statistical properties of different signals. This work is inspired by the studies in [27, 28] of tracking characteristics of frequency domain mean-square errors $E[|\hat{H}(\omega, n) - H(\omega, n)|^2]$ for an SMSL system which can be viewed as a special case of the presented generalized framework.

The derivations in the following are based on the LMS, NLMS and the RLS algorithms. We chose the LMS algorithm for its simplicity and the NLMS algorithm for its popularity in practical applications, whereas we chose the RLS algorithm for its potentially much better convergence properties. With these choices, we derive and interpret analytic expressions for the convergence rate, system stability bound and the steady-state behavior in terms of the PTF. We show how to choose the step size parameter in the adaptive algorithms, for a given desired convergence rate and/or steady-state behavior. Furthermore, our derived expressions can be used to predict if an algorithm would meet given requirements at different frequencies and thereby would be suitable for a particular application.

Parts of this work were published in [29, 30], where the PTF in MMSL systems was introduced. In this paper, we present the in-depth mathematical derivation of the PTF and provide more detailed interpretation of the results, and their relations to existing work for SMSL systems. Finally, we verify the validity of the derived expressions through extensive simulation experiments and demonstrate the practical relevance in a hearing aid AFC system using real data.

In this paper, column vectors and matrices are emphasized using lower and upper letters in bold, respectively. Transposition, Hermitian transposition and complex conjugation are denoted by the superscripts T, H and $*$, respectively.

In Sec. 2, we introduce the system under analysis. We define the exact PTF $\xi(\omega, n)$ and its approximation in Sec. 3. In Sec. 4, we present the detailed derivations of the PTF. After that, we discuss and verify the derived expressions through simulations in Sec. 5 and Sec. 6, respectively. Finally, we give conclusions in Sec. 7.

2 System Description

A detailed overview of the MMSL system under analysis is given in Fig. 2. For convenience, we have expressed the feedback signal $v_i(n)$ and the incoming signal $x_i(n)$ as discrete-time signals, although in practice they are continuous-time signals.

A finite impulse response (FIR) $\mathbf{h}_i(n)$ of order $L-1$ is used to model the ith true feedback/echo path, as

$$\mathbf{h}_i(n) = [h_i(0,n), \ldots, h_i(L-1,n)]^T . \tag{3}$$

We derive results for sufficiently large filter orders, in principle, $L \to \infty$. This ensures that the error in representing the true underlying acoustic feedback/echo path by an FIR tends to zero, even if it has an infinite impulse response (IIR). Furthermore, a lower order FIR of the true acoustic feedback/echo path can be represented by zero-padding to the length L. Thus, knowledge of the exact length of the true acoustic feedback/echo paths is not needed in our analysis.

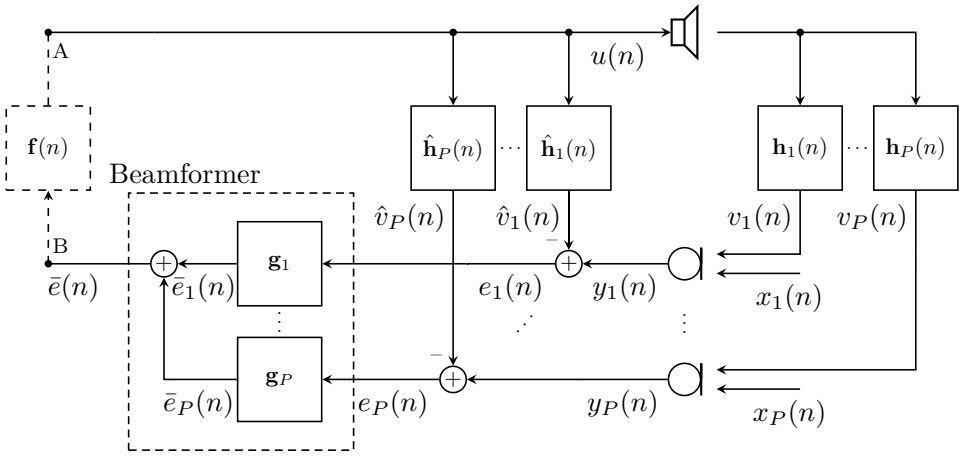


Fig. 2: A general multiple-microphone and single-loudspeaker system. In this work, we focus on the power transfer function from point A to B.

The frequency response determined as the discrete Fourier transform (DFT) of $\mathbf{h}_i(n)$ is expressed by

$$H_i(\omega, n) = \sum_{k=0}^{L-1} h_i(k, n) e^{-j\omega k}. \tag{4}$$

We allow feedback/echo path variations over time. There are different ways to model these variations, see e.g. [14]. In this work, we use a simple random walk model given by

$$H_i(\omega, n) = H_i(\omega, n-1) + \check{H}_i(\omega, n), \tag{5}$$

for the ith feedback/echo path, where $\check{H}_i(\omega, n) \in \mathbb{C}$ is a zero-mean Gaussian stochastic sequence with covariance

$$S_{\check{h}_{ij}}(\omega) = E\left[\check{H}_i(\omega, n)\check{H}_j^*(\omega, n)\right]. \tag{6}$$

In the time domain, the feedback/echo path variation vector is

$$\check{\mathbf{h}}_i(n) = \mathbf{h}_i(n) - \mathbf{h}_i(n-1). \tag{7}$$

The correlation matrix of the ith and jth feedback/echo path variation is defined as

$$\check{\mathbf{H}}_{ij} = E\left[\check{\mathbf{h}}_i(n)\check{\mathbf{h}}_j^T(n)\right]. \tag{8}$$

The adaptively estimated feedback/echo path $\hat{\mathbf{h}}_i(n)$ of order $L-1$ is expressed by

$$\hat{\mathbf{h}}_i(n) = \left[\hat{h}_i(0,n), \ldots, \hat{h}_i(L-1,n)\right]^T, \tag{9}$$

and the estimation error vector which expresses the difference between the true and estimated feedback/echo path is

$$\tilde{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n) - \mathbf{h}_i(n), \tag{10}$$

with a frequency response given by

$$\tilde{H}_i(\omega, n) = \sum_{k=0}^{L-1} \tilde{h}_i(k,n) e^{-j\omega k}. \tag{11}$$

In the analysis, we consider the loudspeaker signal $u(n)$ as a deterministic zero-mean signal, because it is measurable and thereby known with certainty. However, our results remain valid, even if the loudspeaker signal $u(n)$ is considered as a realization of a stochastic process; the same approach is applied and explained in details in [27]. This important point will be demonstrated by simulations in Sec. 6. The loudspeaker signal vector $\mathbf{u}(n)$ is defined as

$$\mathbf{u}(n) = \left[u(n), \ldots, u(n-L+1)\right]^T. \tag{12}$$

We assume the incoming signals $x_i(n)$ are zero-mean stationary stochastic signals with the correlation function

$$r_{x_{ij}}(k) = E\left[x_i(n) x_j(n-k)\right]. \tag{13}$$

The ith microphone signal is modeled as

$$y_i(n) = \mathbf{h}_i^T(n-1)\mathbf{u}(n) + x_i(n). \tag{14}$$

The ith feedback/echo compensated error signal is given by

$$e_i(n) = y_i(n) - \hat{\mathbf{h}}_i^T(n-1)\mathbf{u}(n). \tag{15}$$

In the MMSL system shown in Fig. 2, spatial filtering is carried out by applying beamformer filters on the error signals $e_i(n)$. Each beamformer filter $\mathbf{g}_i$ is an FIR of order $N-1$,

$$\mathbf{g}_i = \left[g_i(0), \ldots, g_i(N-1)\right]^T, \tag{16}$$

with frequency response

$$G_i(\omega) = \sum_{k=0}^{N-1} g_i(k) e^{-j\omega k}. \tag{17}$$

The output signal of the beamformer is therefore

$$\begin{aligned} \bar{e}(n) =& \sum_{i=1}^{P} \bar{e}_i(n) \\ =& \sum_{i=1}^{P} \sum_{k=0}^{N-1} g_i(k) e_i(n-k). \end{aligned} \tag{18}$$

In principle, the order of the beamformer and the acoustic feedback/echo cancellation system could be reversed. In that case, the beamformer would operate directly on the microphone signals, whereas a single-channel acoustic feedback/echo cancellation is carried out on the beamformer processed output signal. In this paper, we focus on the case where the cancellation is performed prior to the beamformer as given in Fig. 2. This setup requires more computational power due to multiple cancellation systems, but the beamformer would not affect the cancellation process negatively as demonstrated in [31].

3 Power Transfer Function

Consider the MMSL system shown in Fig. 2. The PTF $\xi(\omega, n)$ is defined as the expected magnitude-squared TF from point A to B. More specifically, the PTF $\xi(\omega, n)$ is given by

$$\begin{aligned} \xi(\omega, n) =& E\left[\left| \sum_{i=1}^{P} G_i(\omega) \tilde{H}_i(\omega, n) \right|^2 \right] \\ =& \sum_{i=1}^{P} \sum_{j=1}^{P} G_i(\omega) G_j^*(\omega) \xi_{ij}(\omega, n), \end{aligned} \tag{19}$$

where $\xi_{ij}(\omega, n) = E[\tilde{H}_i(\omega, n) \tilde{H}_j^*(\omega, n)]$, and the expectation is with respect to $H_i(\omega, n)$ which is considered as a stochastic variable.

Traditionally, time domain criteria such as mean-square error defined as $E[e^2(n)]$ and mean-square deviation defined as $E[\|\tilde{\mathbf{h}}(n)\|^2]$ have been used in adaptive filter design and performance evaluation to describe convergence rate, stability bound and steady-state behavior of a single adaptive filter, e.g. [4, 18–21]. These criteria are related to the

PTF, despite providing information in different ways. For instance, the mean-square deviation $E[\|\tilde{\mathbf{h}}(n)\|^2]$ can be seen as a time domain, fullband version of the PTF in an SMSL system. We study these relations further in Sec. 5.

The exact analytic expression for Eq. (19) is complicated and difficult to interpret. Thus, in this work, we derive a much simpler but accurate approximation $\hat{\xi}(\omega, n)$. We introduce the notation

$$\hat{\xi}_{ij}(\omega, n) \approx E\left[\tilde{H}_i(\omega, n)\tilde{H}_j^*(\omega, n)\right], \tag{20}$$

and a PTF approximation of $\xi(\omega, n)$ in Eq. (19) is defined as

$$\hat{\xi}(\omega, n) = \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)\hat{\xi}_{ij}(\omega, n). \tag{21}$$

Our derivations are based on an open-loop setup, i.e. $\mathbf{f}(n)$ is omitted in Fig. 2, as in a traditional AEC system. As we demonstrate in a simulation experiment in Sec. 6, the derived results also provide accurate approximations in a closed-loop hearing aid AFC system with a realistic delay in $\mathbf{f}(n)$.

4 System Analysis

4.1 PTF for LMS Algorithm

In this section, we derive the PTF approximation $\hat{\xi}(\omega, n)$ for an MMSL system where cancellation filters $\hat{\mathbf{h}}_i(n)$ are estimated using the LMS algorithm. The LMS update using a step size $\mu(n)$ of the ith channel is expressed by, see e.g. [4],

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{u}(n)e_i(n). \tag{22}$$

Using Eqs. (22), (15), (14) and (7), the estimation error vector defined in Eq. (10) can also be expressed by

$$\tilde{\mathbf{h}}_i(n) = \left(\mathbf{I} - \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)\right)\tilde{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{u}(n)x_i(n) - \check{\mathbf{h}}_i(n), \tag{23}$$

where $\mathbf{I}$ is the identity matrix. Assuming the feedback/echo path variation vector $\check{\mathbf{h}}_i(n)$ is uncorrelated with every other term in Eq. (23), we introduce the matrix $\mathbf{A}(n) = \mathbf{I} - \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)$ and compute the estimation error correlation matrix $\mathbf{H}_{ij}(n) =$

$E\left[\tilde{\mathbf{h}}_i(n)\tilde{\mathbf{h}}_j^T(n)\right]$ as

$$\begin{aligned}\mathbf{H}_{ij}(n) =& E\Big[\tilde{\mathbf{h}}_i(n-1)\tilde{\mathbf{h}}_j^T(n-1) - \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{h}}_i(n-1)\tilde{\mathbf{h}}_j^T(n-1)\mathbf{A}^T(n) \\ & -\tilde{\mathbf{h}}_i(n-1)\tilde{\mathbf{h}}_j^T(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mu(n) + \check{\mathbf{h}}_i(n)\check{\mathbf{h}}_j^T(n) \\ & +\mu^2(n)\mathbf{u}(n)x_i(n)x_j(n)\mathbf{u}^T(n) + \mu(n)\mathbf{u}(n)x_i(n)\tilde{\mathbf{h}}_j^T(n-1)\mathbf{A}^T(n) \\ & +\mathbf{A}(n)\tilde{\mathbf{h}}_i(n-1)x_j(n)\mathbf{u}^T(n)\mu(n)\Big]. \end{aligned} \tag{24}$$

Under the assumption of sufficiently small step size $\mu(n)$, in principle, $\mu(n) \to 0$, it follows that $\mathbf{A}(n) \approx \mathbf{I}$. Using this in Eq. (24) corresponds to neglecting all second order terms involving the matrix $\mu(n)\mathbf{u}(n)\mathbf{u}^T(n)$ due to the presence of the first-order terms. Eq. (24) can now be simplified as

$$\begin{aligned}\mathbf{H}_{ij}(n) \approx& \mathbf{H}_{ij}(n-1) - \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{H}_{ij}(n-1) - \mathbf{H}_{ij}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mu(n) \\ & + \check{\mathbf{H}}_{ij} + \mu^2(n)\mathbf{u}(n)\mathbf{u}^T(n)E\left[x_i(n)x_j(n)\right] \\ & + E\left[\mu(n)\mathbf{u}(n)x_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right] + E\left[\tilde{\mathbf{h}}_i(n-1)x_j(n)\mathbf{u}^T(n)\mu(n)\right]. \end{aligned} \tag{25}$$

Eq. (25) is a difference equation in $\mathbf{H}_{ij}(n)$. According to the direct-averaging method described in [32], and using again the small step size assumption $\mu(n) \to 0$, $\mathbf{u}(n)\mathbf{u}^T(n)$ can be replaced by its sample average $\mathbf{R}_u(0)$, where $\mathbf{R}_u(k)$ is defined as

$$\mathbf{R}_u(k) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbf{u}(n)\mathbf{u}^T(n-k). \tag{26}$$

Using Eqs. (25)-(26), the approximated estimation error correlation matrix $\hat{\mathbf{H}}_{ij}(n) \approx E[\tilde{\mathbf{h}}_i(n)\tilde{\mathbf{h}}_j^T(n)]$ is written as

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =& \hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\mathbf{R}_u(0)\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{R}_u(0) \\ & + \check{\mathbf{H}}_{ij} + \mu^2(n)\mathbf{R}_u(0)r_{x_{ij}}(0) + E\left[\mu(n)\mathbf{u}(n)x_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right] \\ & + E\left[\tilde{\mathbf{h}}_i(n-1)x_j(n)\mathbf{u}^T(n)\mu(n)\right]. \end{aligned} \tag{27}$$

Assuming $\mu(n) \to 0$ and that the incoming signals $x_i(n)$ have a finite correlation function, i.e.

$$r_{x_{ij}}(k) = 0 \;\forall\; |k| > k_0, \tag{28}$$

where k_0 is a finite integer number, it can be shown (see Appendix A for details) that Eq. (27) can be written as

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =& \hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\mathbf{R}_u(0)\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{R}_u(0) \\ & + \check{\mathbf{H}}_{ij} + \mu^2(n)\sum_{k=-k_0}^{k_0} \mathbf{R}_u(k)r_{x_{ij}}(k). \end{aligned} \tag{29}$$

Recall that $\hat{\mathbf{H}}_{ij}(n) \approx E[\tilde{\mathbf{h}}_i(n)\tilde{\mathbf{h}}_j^T(n)]$ and $\hat{\xi}_{ij}(\omega,n) \approx E[\tilde{H}_i(\omega,n)\tilde{H}_j^*(\omega,n)]$, where $\tilde{H}_i(\omega,n)$ is the frequency response of $\tilde{\mathbf{h}}_i(n)$. To find an expression for $\hat{\xi}_{ij}(\omega,n)$, we let $\mathbf{F} \in \mathbb{C}^{L\times L}$ be a DFT matrix. It is well-known that $\mathbf{F}$ diagonalizes a Toeplitz matrix asymptotically as $L \to \infty$ [33]. Thus, the matrix

$$\hat{\mathbf{\Xi}}_{ij}(n) = \mathbf{F}\hat{\mathbf{H}}_{ij}(n)\mathbf{F}^H \tag{30}$$

approaches a diagonal matrix, as $L \to \infty$, with the diagonal elements $\hat{\xi}_{ij}(\omega,n)$ as given in Eq. (20).

Similarly, both $\mathbf{F}\check{\mathbf{H}}_{ij}\mathbf{F}^H$ and $\frac{1}{L}\mathbf{F}\mathbf{R}_u(0)\mathbf{F}^H$ approach diagonal matrices as $L \to \infty$. The resulting diagonal elements $S_{\check{h}_{ij}}(\omega)$ and $S_u(\omega)$ are the covariances of the underlying feedback/echo path changes, and the power spectrum density (PSD) of the loudspeaker signal $u(n)$, respectively.

Inserting Eq. (29) in Eq. (30) and using that $\frac{1}{L}\mathbf{F}^H\mathbf{F} = \mathbf{I}$, the matrix $\hat{\mathbf{\Xi}}_{ij}(n)$ is expressed by

$$\begin{aligned}\hat{\mathbf{\Xi}}_{ij}(n) =& \mathbf{F}\hat{\mathbf{H}}_{ij}(n-1)\mathbf{F}^H + \mathbf{F}\check{\mathbf{H}}_{ij}\mathbf{F}^H - \mu(n)\frac{1}{L}\mathbf{F}\mathbf{R}_u(0)\mathbf{F}^H\mathbf{F}\hat{\mathbf{H}}_{ij}(n-1)\mathbf{F}^H \\ &- \mu(n)\frac{1}{L}\mathbf{F}\hat{\mathbf{H}}_{ij}(n-1)\mathbf{F}^H\mathbf{F}\mathbf{R}_u(0)\mathbf{F}^H + \mu^2(n)\sum_{k=-k_0}^{k_0}\mathbf{F}\mathbf{R}_u(k)\mathbf{F}^H r_{x_{ij}}(k). \end{aligned}\tag{31}$$

$\hat{\xi}_{ij}(\omega,n)$, defined in Eq. (20), follow as the diagonal elements of $\hat{\mathbf{\Xi}}_{ij}(n)$ which are given by

$$\hat{\xi}_{ij}(\omega,n) = (1-2\mu(n)S_u(\omega))\,\hat{\xi}_{ij}(\omega,n-1) + L\mu^2(n)S_u(\omega)S_{x_{ij}}(\omega) + S_{\check{h}_{ij}}(\omega), \tag{32}$$

where $S_{x_{ij}}(\omega)$ denotes the cross(auto) PSDs of the incoming signals $x_i(n)$ and $x_j(n)$.

Finally, inserting Eq. (32) in Eq. (21), we arrive at

$$\begin{aligned}\hat{\xi}(\omega,n) =& (1-2\mu(n)S_u(\omega))\,\hat{\xi}(\omega,n-1) \\ &+ L\mu^2(n)S_u(\omega)\sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega) \\ &+ \sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega). \end{aligned}\tag{33}$$

4.2 PTF for NLMS Algorithm

We can use the methodology from Sec. 4.1 to derive the PTF approximation $\hat{\xi}(\omega,n)$ for the NLMS algorithm. However, in this section, we show how to obtain the same result

more easily by adapting the results of the LMS algorithm. The NLMS update of the ith cancellation filter is, see e.g. [4],

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \bar{\mu}(n)\frac{\mathbf{u}(n)e_i(n)}{\mathbf{u}^T(n)\mathbf{u}(n)+\delta}, \tag{34}$$

where $\bar{\mu}(n)$ is the NLMS step size, and $\delta > 0$ is a scalar often referred to as the regularization term.

Note that $\mathbf{u}^T(n)\mathbf{u}(n) = L\hat{\sigma}_u^2$, where $\hat{\sigma}_u^2(n)$ is an estimate of the variance σ_u^2 of the loudspeaker signal $u(n)$. Using the fact that for small step sizes $\bar{\mu}(n) \to 0$, the NLMS algorithm has a low-pass effect on the loudspeaker signal $u(n)$, and this allows us to replace $\hat{\sigma}_u^2(n)$ by σ_u^2. Hence, Eq. (34) can be rewritten as

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \frac{\bar{\mu}(n)}{L\sigma_u^2+\delta}\mathbf{u}(n)e_i(n). \tag{35}$$

From Eqs. (35) and (22), it is seen that the relation between the LMS and NLMS algorithms is a normalized step size according to

$$\mu(n) = \frac{\bar{\mu}(n)}{L\sigma_u^2+\delta}. \tag{36}$$

Inserting Eq. (36) in Eq. (33), the PTF approximation $\hat{\xi}(\omega,n)$ of the MMSL system using the NLMS algorithm, under the same assumptions as for the LMS algorithm, is expressed by

$$\begin{aligned}\hat{\xi}(\omega,n) = &\left(1 - 2\frac{\bar{\mu}(n)}{L\sigma_u^2+\delta}S_u(\omega)\right)\hat{\xi}(\omega,n-1)\\ &+ L\frac{\bar{\mu}^2(n)}{(L\sigma_u^2+\delta)^2}S_u(\omega)\sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)\\ &+ \sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega).\end{aligned} \tag{37}$$

4.3 PTF for RLS Algorithm

The RLS update step is given by, see e.g. [4],

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mathbf{Z}(n)\mathbf{u}(n)e_i(n), \tag{38}$$

$$\mathbf{Z}(n) = \frac{\mathbf{P}(n-1)}{\lambda + \mathbf{u}^T(n)\mathbf{P}(n-1)\mathbf{u}(n)}, \tag{39}$$

$$\mathbf{P}(n) = \frac{1}{\lambda}\left(\mathbf{P}(n-1) - \mathbf{Z}(n)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{P}(n-1)\right), \tag{40}$$

where $0 < \lambda < 1$ denotes the forgetting factor, and $\mathbf{P}(0) = \delta\mathbf{I}$, δ is a regularization parameter.

The same methodology used in Sec. 4.1 to derive the PTF approximation $\hat{\xi}(\omega, n)$ for the LMS algorithm can be used for the RLS algorithm. The resulting PTF $\hat{\xi}(\omega, n)$ can be found to be

$$\begin{aligned}\hat{\xi}(\omega, n) = &\left(1 - 2z(\omega, n)S_u(\omega)\right)\hat{\xi}(\omega, n-1) \\ &+ Lz^2(\omega, n)S_u(\omega)\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega) \\ &+ \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega),\end{aligned} \tag{41}$$

where $z(\omega, n)$ is the element in the diagonal of $\frac{1}{L}\mathbf{F}\mathbf{Z}(n)\mathbf{F}^H$.

We are interested in finding an expression for $z(\omega, n)$ in Eq. (41). The matrix $\mathbf{P}(n)$ is a recursively updated inverse matrix in the RLS algorithm expressed by $\mathbf{P}(n) = \left(\sum_{m=1}^{n} \lambda^{n-m}\mathbf{u}(m)\mathbf{u}^T(m) + \delta\lambda^n\mathbf{I}\right)^{-1}$. Asymptotically, as $\lambda \to 1$ and for large values of n, the matrix $\sum_{m=1}^{n} \lambda^{n-m}\mathbf{u}(m)\mathbf{u}^T(m)$ contains large values, and therefore the matrix $\mathbf{P}(n)$ tends to have small entries. Hence, for large values of n, and asymptotically, as $\lambda \to 1$, the matrix $\mathbf{Z}(n)$ in Eq. (39) can be approximated by $\mathbf{Z}(n) \approx \mathbf{P}(n)$[1], and the matrix $\mathbf{P}(n)$ in Eq. (40) can therefore be expressed by

$$\mathbf{Z}(n) \approx \frac{1}{\lambda}\left(\mathbf{Z}(n) - \mathbf{Z}(n)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{Z}(n)\right). \tag{42}$$

The matrix $\mathbf{P}(n) \approx \mathbf{Z}(n)$ becomes Toeplitz structure when converged and can be diagonalized using the DFT matrix $\mathbf{F}$. Based on Eq. (42), we calculate $z(\omega, n)$ as the diagonal entries of the matrix $\frac{1}{L}\mathbf{F}\mathbf{Z}(n)\mathbf{F}^H$, which are given by

$$z(\omega, n) \approx \frac{1}{\lambda}\left(z(\omega, n) - z^2(\omega, n)S_u(\omega)\right). \tag{43}$$

Solving the second-order difference equation in $z(\omega, n)$ gives

$$z(\omega, n) = \frac{1 - \lambda}{S_u(\omega)}. \tag{44}$$

Inserting Eq. (44) in Eq. (41), the PTF approximation $\hat{\xi}(\omega, n)$ for the RLS algorithm is

[1]This requires that the matrix $\mathbf{P}(n)$ has converged, i.e. $\mathbf{P}(n) \approx \mathbf{P}(n-1)$. Convergence of $\mathbf{P}(n)$ does not necessarily mean that $n \to \infty$, we observed from the simulations that $\mathbf{P}(n)$ may already converge for $n < 1000$.

finally expressed by

$$\begin{aligned}\hat{\xi}(\omega,n) = (2\lambda-1)\,\hat{\xi}(\omega,n-1) \\ + L\frac{(1-\lambda)^2}{S_u(\omega)}\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega) \\ + \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega). \end{aligned} \tag{45}$$

5 Discussion

In this section, we use the derived expressions for the LMS, NLMS and the RLS algorithms to predict the system behavior. Specifically, we discuss the system behavior, in terms of convergence rate defined by the decay rate of $\hat{\xi}(\omega,n)$, system stability bound of the step size parameters to ensure algorithm convergence, and steady-state behavior describing $\hat{\xi}(\omega,n)$ when the adaptive algorithm has converged. Furthermore, we discuss how to use the derived PTF expressions to choose the step size parameters when given a desired system behavior for a specific frequency ω; this is especially useful for setting up the parameters in closed-loop applications such as a hearing aid because the system instability as consequence of the acoustic feedback often occurs at a single frequency at a time.

Eqs. (33), (37) and (45) are all first-order difference equations in $\hat{\xi}(\omega,n)$ expressed by a TF $T(z)=\frac{\beta}{1-\alpha z^{-1}}$, where $\alpha,\beta\in\mathbb{R}$. The coefficient α determines the pole location of $T(z)$ and thereby the convergence rate of the system. The convergence rate CR in dB per iteration (in this case, for each time instant n) can be calculated as the derivative of the logarithm of the envelope of the impulse response (IR) function, $t(n)=\beta\cdot\alpha^n$, of $T(z)$ as

$$\begin{aligned}\mathrm{CR[dB/iteration]} &= \frac{d}{dn}10\log_{10}(\beta\cdot|\alpha|^n) \\ &= 10\log_{10}(|\alpha|). \end{aligned} \tag{46}$$

Furthermore, stability of $T(z)$ is ensured whenever

$$|\alpha|<1. \tag{47}$$

The steady-state behavior is described through evaluation of

$$\hat{\xi}(\omega,\infty)=\lim_{n\to\infty}\hat{\xi}(\omega,n). \tag{48}$$

5.1 System Behavior for LMS Algorithm

In Eq. (33), the frequency dependent coefficient $\alpha(\omega)$ is expressed by

$$\alpha(\omega) = 1 - 2\mu(n)S_u(\omega). \tag{49}$$

Using Eqs. (49) and (47), the step size $\mu(n)$ to ensure system stability is given by

$$0 < \mu(n) < \frac{1}{\max_\omega S_u(\omega)}. \tag{50}$$

Inserting Eq. (33) in Eq. (48), the steady-state behavior is

$$\hat{\xi}(\omega,\infty) = \underbrace{\lim_{n\to\infty} L\frac{\mu(n)}{2}\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}_{\text{Steady-State Error}} + \underbrace{\lim_{n\to\infty}\frac{\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega)}{2\mu(n)S_u(\omega)}}_{\text{Tracking Error}}. \tag{51}$$

Eq. (51) consists of two parts. The first part indicates the lowest possible steady-state value of $\hat{\xi}(\omega,n)$; the second part gives the additional tracking error as a result of the variations $S_{\check{h}_{ij}}(\omega) > 0$ in the feedback/echo paths.

Using Eqs. (46) and (49), a desired convergence rate in dB/iteration is achieved by choosing the step size according to

$$\mu(n) = \frac{1 - 10^{\text{CR[dB/iteration]}/10}}{2S_u(\omega)}. \tag{52}$$

Using Eq. (51), a desired steady-state error $\hat{\xi}(\omega,\infty)$, ignoring the tracking error for simplicity, could be achieved by setting the step size $\mu(n)$ as

$$\mu(n) = \frac{2\hat{\xi}(\omega,\infty)}{L\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}. \tag{53}$$

We observe from Eq. (49) that the convergence rate only depends on the step size $\mu(n)$ and the PSD $S_u(\omega)$ of the loudspeaker signal $u(n)$. Similar results for the LMS algorithm in an SMSL system are derived in [27].

From the first part of Eq. (51), it is observed that the model order parameter L, the step size $\mu(n)$, and the PSDs $S_{x_{ij}}(\omega)$ weighted by the frequency responses $G_i(\omega)$ and $G_j^*(\omega)$ are linearly proportional to the steady-state error. However, the second part of

Eq. (51) shows that a larger step size $\mu(n)$ and higher PSD $S_u(\omega)$ lead to smaller tracking error when the system is undergoing variations, i.e. $S_{\check{h}_{ij}}(\omega) > 0$. The overall steady-state value $\hat{\xi}(\omega, \infty)$ is therefore a compromise between the steady-state behavior in situations with time invariant feedback/echo paths and tracking behavior in situations with time varying feedback/echo paths, this trade-off is well-known from existing fullband SMSL system analyses, e.g. in [4]. Furthermore, the frequency responses $G_i(\omega)$ and $G_j^*(\omega)$ act as weighting factors for $S_{x_{ij}}(\omega)$ and $S_{\check{h}_{ij}}(\omega)$, thus, the expected steady-state value $\hat{\xi}(\omega, \infty)$ according to Eq. (51) would change instantly followed by any changes in $G_i(\omega)$ and $G_j^*(\omega)$, even when the signals such as $u(n)$ and $x_i(n)$ were stationary, and the step size parameter $\mu(n)$ was unchanged.

5.2 System Behavior for NLMS Algorithm

From Eq. (37), the convergence rate for the NLMS algorithm is determined by the coefficient

$$\alpha(\omega) = 1 - 2\frac{\bar{\mu}(n)}{L\sigma_u^2 + \delta} S_u(\omega). \tag{54}$$

Inserting Eq. (54) in (47), the range of the step size $\bar{\mu}(n)$ to ensure system stability is determined as

$$0 < \bar{\mu}(n) < \frac{L\sigma_u^2 + \delta}{\max_\omega S_u(\omega)}. \tag{55}$$

Furthermore, the steady-state behavior is expressed by

$$\hat{\xi}(\omega, \infty) = \underbrace{\lim_{n\to\infty} L\frac{\bar{\mu}(n)}{2(L\sigma_u^2 + \delta)} \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}_{\text{Steady-State Error}} + \underbrace{\lim_{n\to\infty} (L\sigma_u^2 + \delta)\frac{\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega)}{2\bar{\mu}(n)S_u(\omega)}}_{\text{Tracking Error}}. \tag{56}$$

In order to achieve desired convergence rate and steady-state error, the NLMS step size $\bar{\mu}(n)$ should be chosen as:

$$\bar{\mu}(n) = (L\sigma_u^2 + \delta)\frac{1 - 10^{\text{CR[dB/iteration]}/10}}{2S_u(\omega)}, \tag{57}$$

$$\bar{\mu}(n) = \frac{2(L\sigma_u^2 + \delta)\hat{\xi}(\omega, \infty)}{L\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}. \tag{58}$$

Although the structure of Eqs. (54)-(56) for the NLMS algorithm is similar to Eqs. (49)-(51) for the LMS algorithm, there are some important differences in the system behavior due to the normalization of the step size.

The convergence rate and the tracking error in the NLMS adaptation are no more dependent on the absolute value of $S_u(\omega)$ due to the presence of the variance σ_u^2 in Eqs. (54) and (56), but rather the value of $S_u(\omega)$ relative to $L\sigma_u^2 + \delta$. Increasing the value of L results in decreased convergence rate and increased tracking error. The steady-state error is now also dependent on the variance σ_u^2. A higher variance leads to lower steady-state error.

We observe from Eq. (55) that the stability upper-bound of the NLMS step size $\bar{\mu}(n)$ is identical to the LMS stability upper-bound scaled by the factor of $L\sigma_u^2 + \delta$. As in the case when using the LMS algorithm, the expected steady-state error $\hat{\xi}(\omega, \infty)$ in Eq. (56) would instantly follow any variations in the frequency responses $G_i(\omega)$ and $G_j^*(\omega)$.

5.3 System Behavior for RLS Algorithm

For the RLS algorithm, the coefficient α, which expresses the convergence rate, is obtained from Eq. (45) as

$$\alpha = 2\lambda - 1, \tag{59}$$

and according to Eq. (47), stability is ensured for

$$0 < \lambda < 1. \tag{60}$$

The steady-state behavior is expressed by

$$\hat{\xi}(\omega, \infty) = \underbrace{L\frac{1-\lambda}{2S_u(\omega)} \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}_{\text{Steady-State Error}} + \underbrace{\frac{\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\tilde{h}_{ij}}(\omega)}{2(1-\lambda)}}_{\text{Tracking Error}}. \tag{61}$$

Furthermore, in order to achieve a desired convergence rate and steady-state error, the RLS forgetting factor λ should be:

$$\lambda = \frac{1 + 10^{\text{CR[dB/iteration]}/10}}{2}, \tag{62}$$

$$\lambda = 1 - \frac{2S_u(\omega)\hat{\xi}(\omega, \infty)}{L\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}. \tag{63}$$

Table 1: System behavior in terms of convergence rate (CR), steady-state error (SSE) and tracking error (TE) at frequency ω, when increasing the value of different system parameters. ($\uparrow$: increase, $\downarrow$: decrease, $-$: unchanged.)

System	Parameter	CR	SSE	TE
LMS	$\uparrow \mu(n)$	$\uparrow$	$\uparrow$	$\downarrow$
	$\uparrow L$	-	$\uparrow$	-
	$\uparrow S_u(\omega)$	$\uparrow$	-	$\downarrow$
NLMS	$\uparrow \bar{\mu}(n)$	$\uparrow$	$\uparrow$	$\downarrow$
	$\uparrow L/(L\sigma_u^2+\delta)$	-	$\uparrow$	-
	$\uparrow S_u(\omega)/(L\sigma_u^2+\delta)$	$\uparrow$	-	$\downarrow$
RLS	$\uparrow \lambda$	$\downarrow$	$\downarrow$	$\uparrow$
	$\uparrow L$	-	$\uparrow$	-
	$\uparrow S_u(\omega)$	-	$\downarrow$	-
Common	$\uparrow G_i(\omega)$	-	As weights	As weights
	$\uparrow S_{x_{ij}}(\omega)$	-	$\uparrow$	-
	$\uparrow S_{\check{h}_{ij}}(\omega)$	-	-	$\uparrow$

From Eqs. (59)-(61), we observe some major differences in the RLS algorithm compared to the LMS and NLMS algorithms. For the RLS algorithm, the convergence rate is not only frequency independent, but also signal independent; it only depends on the value of the forgetting factor λ. Increasing the value of λ gives slower convergence and tracking, but also lower steady-state error. In contrast to the LMS and NLMS algorithms, the PSD $S_u(\omega)$ is inversely proportional to the steady-state error when using the RLS algorithm, and it has no influences on the convergence rate and the tracking error. However, similarly to the LMS algorithm, increasing the model order parameter L leads to higher steady-state error. Once more, we observe that any variations in the frequency responses $G_i(\omega)$ and $G_j^*(\omega)$ would lead to an instant change in the expected steady-state error $\hat{\xi}(\omega, n)$ given by Eq. (61).

5.4 Summary of System Behavior

In Table 1, system behaviors for the LMS, NLMS and the RLS algorithms are summarized in terms of convergence rate, steady-state error and tracking error, when varying the values of most important system parameters.

The similarities and differences shown in Table 1 are well expected due to the underlying assumptions and procedures of the algorithms. Identical behaviors can simply

be obtained for the LMS and NLMS algorithms, because the NLMS algorithm can be seen as a step size adjusted LMS algorithm according to Eq. (36). It is also possible to achieve identical convergence rate and steady-state behavior for a specific frequency ω in the NLMS and RLS algorithms. To see this, equate Eqs. (54) and (59), and solve for the forgetting factor λ, to find

$$\lambda = 1 - \frac{\bar{\mu}(n)}{L\sigma_u^2 + \delta} S_u(\omega). \tag{64}$$

It is seen from Eq. (64) that the NLMS and RLS behavior is generally different across frequencies, unless the values of the PSD $S_u(\omega)$ are identical at different frequencies.

In principle, if the step size $\bar{\mu}(n)$ could be independently adjusted across frequencies in the NLMS algorithm, then it would be possible to obtain identical behavior for the NLMS and RLS algorithms for all frequencies. In practice, this could e.g. be achieved approximately using frequency-domain adaptive filters and/or a sub-band structure with NLMS adaptation carried out in each sub-band, see e.g. [34–36]. In this case, the NLMS algorithm might be preferred because the frequency dependent step size $\bar{\mu}(\omega, n)$ gives more choices to obtain desired properties which are not possible with an RLS algorithm with a single design parameter λ.

5.5 Relation to Existing Work

As mentioned, many studies exist already for adaptive system performance analysis using the LMS, NLMS and RLS algorithms in different contexts. The relationships observed in this work between the convergence rate, steady-state error in a time invariant system, and the additional tracking error in a time varying system are well in line with these studies. The PTF analysis in this work differs mainly from the existing studies in two ways. First, we analyze a multiple-microphone system, whereas existing works mainly focus on the single-microphone situation. Second, we evaluate the system performance for each frequency over time, whereas existing works tend to focus on fullband performance over time. In the following, we relate the PTF to the analysis results in [4] and [18], which both focus on the fullband performance analysis of adaptive algorithms in a single-microphone time-varying system.

Using the mean-square deviation criterion $\mathcal{D}(n) = ||\mathbf{h}(n) - \hat{\mathbf{h}}(n)||^2$, it was shown in [4] that the first-order difference equations describing the kth natural mode of the LMS filter contain the iteration coefficients $1 - \mu\lambda_{u,k}$, where μ is the step size and $\lambda_{u,k}$ is the kth eigenvalue of the correlation matrix of the loudspeaker signal $u(n)$. From this, the well-known system stability bound of the LMS algorithm [4] is given by

$$0 < \mu < \frac{2}{\max_k \lambda_{u,k}}. \tag{65}$$

Using the fact that $\max_k \lambda_{u,k} \leq \max_\omega S_u(\omega)$ [4], we observe that our derived stability upper-bound in Eq. (50) is smaller than the upper-bound defined in Eq. (65). This

difference is due to the approximation made from Eq. (24) to Eq. (25), where we, under the assumption of $\mu(n) \to 0$, ignored the second-order term $E[\mu^2(n)\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{h}}_i(n-1)\tilde{\mathbf{h}}_j^T(n-1)\mathbf{u}(n)\mathbf{u}^T(n)]$ in Eq. (24). Nevertheless, asymptotically, as $\mu(n) \to 0$, the derived stability bound is still valid. Although not explicitly treated in [4], the coefficients $1-\mu\lambda_{u,k}$ are linked to convergence rates $\alpha(\omega)$ in Eq. (49), as the eigenvalues $\lambda_{u,k}$ of the $L \times L$ dimensional correlation matrix of the loudspeaker signal $u(n)$ approach the PSD $S_u(\omega)$ of the signal, as $L \to \infty$ [4].

The derived steady-state behaviors in this work are also well related to the existing studies. In [4], the mean-square deviation $\mathcal{D}(n)$ for the LMS algorithm and large n is expressed by

$$\mathcal{D}(n) \approx \underbrace{\frac{\mu}{2}L\sigma_x^2}_{\text{Steady-State Error}} + \underbrace{\frac{1}{2\mu}\sum_{k=1}^{L}\frac{\lambda_{\check{h},k}}{\lambda_{u,k}}}_{\text{Tracking Error}}, \tag{66}$$

where σ_x^2 denotes the variance of the incoming signal $x(n)$, $\lambda_{\check{h},k}$ is the kth eigenvalue of the correlation matrix of the feedback/echo path variation vector $\check{\mathbf{h}}(n)$, and where we have adapted the notation from our analysis for convenience.

A similar result to Eq. (66) was obtained in [18], where the performance analysis is carried out for the NLMS(LMS), AP and RLS algorithms and their tap-selective partial updating versions in a single-microphone time-varying system. However, since a slightly different model is used for the time-varying acoustic feedback/echo paths, the mean-square deviation $\mathcal{D}(n)$ for the LMS algorithm was derived as[2],

$$\mathcal{D}(n) \approx \underbrace{\frac{\mu}{2}L\sigma_x^2}_{\text{Steady-State Error}} + \underbrace{\frac{(1-\zeta)L\sigma_{\check{h}}^2}{\mu\sigma_u^2}}_{\text{Tracking Error}}, \tag{67}$$

where σ_u^2 is the variance of the loudspeaker signal $u(n)$, and ζ is a parameter used in [18] to model the variations of acoustic feedback/echo paths $\mathbf{h}(n+1) = \zeta\mathbf{h}(n) + \sqrt{1-\zeta^2}\check{\mathbf{h}}(n)$, where elements of $\check{\mathbf{h}}(n)$ are drawn from the normal distribution $N(0, \sigma_{\check{h}}^2)$.

Comparing Eqs. (66)-(67) to the steady-state error of Eq. (51), it is clear that the equivalent term to σ_x^2 in $\mathcal{D}(n)$ is $\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)$ in $\hat{\xi}(\omega, n)$, which is a combined result from different microphone channels through the beamformer including the cross-channel effects. Choosing $P = 1$ and $G(\omega) = 1$ simplifies this term to $S_x(\omega)$, which is the spectral variance measured at frequency ω; summing over all frequencies, we get σ_x^2.

[2]Note that the LMS step size used in [18] is scaled by a factor of $\frac{1}{2}$ compared to the step size from this study.

In our analysis, we used a model for the variations of feedback/echo paths identical to [4], so a comparison between Eqs. (51) and (66) is therefore most meaningful. The eigenvalues $\lambda_{\check{h},k}$ and $\lambda_{u,k}$ in Eq. (66) approach $S_{\check{h}_{ij}}(\omega)$ and $S_u(\omega)$, respectively, as $L \to \infty$. As before, choosing $P = 1$, $G(\omega) = 1$ and summing over all frequencies in the expression for the tracking error in Eq. (51) lead to the expression for the tracking error in Eq. (66). Similar expression of the tracking error is observed in Eq. (67), however, since different models were used for the acoustic feedback/echo path variations $\check{\mathbf{h}}(n)$, the result is further scaled by the factor $2(1-\zeta)$.

Similar comparisons can be made between the derived PTF expressions and the analysis results from e.g. [4, 18], for the NLMS and RLS algorithms. In all cases, the well-known fullband behavior for SMSL systems can be obtained from the derived PTF expressions by considering a $P = 1$ microphone setup, ignoring the beamformer coefficients, and summing over all frequencies.

In addition to the open-loop analyses discussed above, a closed-loop steady-state analysis of an AFC system in hearing aids using the NLMS algorithm was studied in [37]. Let us focus on the first part of Eq. (56) and simplify it to a $P = 1$ microphone setup with the beamformer frequency response $G(\omega) = 1$. Adopting the assumptions from [37] that $x(n)$ is white noise, $\sigma_u^2 = S_x(\omega)$, and $\delta = 0$, we obtain $\hat{\xi}(\omega, \infty) = \lim_{n\to\infty} \bar{\mu}(n)/2$, which is the result presented in [37].

6 Simulation Experiments

In this section, we verify the derived expressions by simulation experiments. In the first experiment, we verify the derived expressions for the convergence rate, steady-state behavior and the step size parameters using synthetic signals in an open-loop MMSL system. In the second experiment, we demonstrate the practical relevance of the derived expressions in a closed-loop MMSL system, in particular, a hearing aid AFC system using real data including speech signals.

6.1 Simulation Experiment Using Synthetic Signals

In the first simulation experiment, we consider an open-loop MMSL system with $P = 3$ microphones. During the simulations, the true feedback/echo paths $\mathbf{h}_i(n)$ are known. Therefore, it is possible to compute the true PTF $\xi(\omega, n)$ given by Eq. (19) using Eqs. (10), (11) and (17) to verify the predicted convergence rate and steady-state behavior. In order to compute $\xi(\omega, n)$, the simulations consist of a number of R runs; an averaged result based on all R runs, at each frequency, is calculated as $\bar{\xi}(\omega, n) = \frac{1}{R}\sum_{r=1}^{R} \xi_r(\omega, n)$, where $\xi_r(\omega, n)$ denotes the result of the simulation run r.

The duration of each simulation run is denoted by D_s, where $n = 0, \ldots, D_s - 1$. The feedback/echo path estimates are initialized to zeros, i.e. $\hat{\mathbf{h}}_i(0) = \mathbf{0}$, in all sim-

ulation runs. In the first part of the experiment, the true feedback/echo paths are fixed during the first half of the simulation, whereas variations are added to these during the second half of the simulation according to the random walk model $h_i(k,n) = h_i(k,0) + \sum_{m=D_s/2}^{n} \epsilon_{h_i}(m)$, where $h_i(k,n)$ is the kth tap of the IR of the true feedback/echo paths at time index n, and $\epsilon_{h_i}(n)$ denotes the nth sample of a realization of a Gaussian stochastic sequence with mean value μ_{h_i} and variance $\sigma^2_{h_i}$. In the second part of the experiment, the feedback/echo paths $\mathbf{h}_i(n)$ remain fixed.

For each simulation run, the loudspeaker signal $u(n)$ is generated by filtering a realization of Gaussian stochastic sequence through an $L_u - 1$ order shaping filter, $\mathbf{h}_u = [h_u(0), \ldots, h_u(L_u-1)]^T$, i.e. $u(n) = \sum_{k=0}^{L_u-1} h_u(k)\epsilon_u(n-k)$, where $\epsilon_u(n)$ denotes the nth sample of a realization of the standard Gaussian stochastic sequence to generate $u(n)$. The incoming signal $x_1(n)$ is created as $x_1(n) = \sum_{k=0}^{L_{x_1}-1} h_{x_1}(k)\epsilon_{x_1}(n-k)$, where an $L_{x_1} - 1$ order shaping filter, $\mathbf{h}_{x_1} = [h_{x_1}(0), \ldots, h_{x_1}(L_{x_1}-1)]^T$ is applied to a realization of standard Gaussian stochastic sequence ϵ_{x_1} for generating $x_1(n)$. The remaining incoming signals $x_i(n)$ for $i = 2, \ldots, P$ are generated as $x_i(n) = \kappa_i x_1(n) + \sum_{k=0}^{L_{x_i}-1} h_{x_i}(k)\epsilon_{x_i}(n-k)$, where κ_i is a mixing factor.

All common simulation parameters for this experiment are given in Table 2. A minimum order of various filters are used to be able to provide the numerical values of all coefficients.

6.1.1 Simulation Experiment for LMS Algorithm

A constant step size $\mu = 2^{-9}$ is used in the first part of this experiment. The simulated and predicted values of convergence rate and steady-state values of $\hat{\xi}(\omega, n)$ obtained from Eqs. (49) and (51) at two representative frequencies $\omega = \frac{2\pi l}{L}, l = 7, 11$ are shown in Figs. 3(a)-(b). Clearly, the simulation results support the predicted convergence rate and steady-state behavior. As expected, we observe different convergence rates, steady-state errors and the additional tracking errors due to the spectral shaping of $\mathbf{h}_u$, $\mathbf{h}_{x_i}$ and $\mathbf{g}_i$. Differences between the simulation results and the predicted values at selected time indices are computed and shown in Table 3. In all cases, only very small deviations are observed.

Despite the underlying assumptions of $L \to \infty$ and $\mu(n) \to 0$ in the analysis, we observe from the simulation results that, in practice, the derived expressions can give accurate approximations already at values such as $L = 32$ and $\mu = 2^{-9}$. Another important point demonstrated by the simulations is that the derived results are valid, when the loudspeaker signal $u(n)$ is a realization of a stochastic process, although $u(n)$ was considered deterministic in the analysis.

In the second part of this experiment, using Eqs. (52)-(53), we calculated the step size $\mu \approx 0.0012$ and $\mu \approx 0.0008$ to achieve a desired convergence rate of -0.01 dB/iteration and a steady-state error of -12 dB, respectively. Only the step size and the variations of feedback/echo paths differ from the first part of the experiment. The simulated

Table 2: Common simulation parameters for all adaptive algorithms.

Symbol	Value	Description
D_s	10000	Duration of simulation.
R	100	Number of simulation runs.
P	3	Number of microphone channels.
L	32	Length of $\hat{\mathbf{h}}_i(n)$.
$\mathbf{g}_1$	$[1, 0.36]^T$	IR of beamformer filter 1.
$\mathbf{g}_2$	$[1, -0.32]^T$	IR of beamformer filter 2.
$\mathbf{g}_3$	$[1, 0.23]^T$	IR of beamformer filter 3.
$\mathbf{h}_1(0)$	$[1, 0.14]^T$	Initial IR of $\mathbf{h}_1(n)$.
$\mathbf{h}_2(0)$	$[1, -0.40]^T$	Initial IR of $\mathbf{h}_2(n)$.
$\mathbf{h}_3(0)$	$[1, 0.23]^T$	Initial IR of $\mathbf{h}_3(n)$.
$N(\mu_{h_1}, \sigma^2_{h_1})$	$N(0, 0.0384^2)$	Gaussian statistics for $\mathbf{h}_1(n)$.
$N(\mu_{h_2}, \sigma^2_{h_2})$	$N(0, 0.0332^2)$	Gaussian statistics for $\mathbf{h}_2(n)$.
$N(\mu_{h_3}, \sigma^2_{h_3})$	$N(0, 0.0024^2)$	Gaussian statistics for $\mathbf{h}_3(n)$.
L_u	2	Length of shaping filter for $u(n)$.
$\mathbf{h}_u$	$[1, -0.3]^T$	IR of shaping filter for $u(n)$.
L_{x_1}	2	Length of shaping filter for $x_1(n)$.
L_{x_2}	2	Length of shaping filter for $x_2(n)$.
L_{x_3}	2	Length of shaping filter for $x_3(n)$.
$\mathbf{h}_{x_1}$	$[1, 0.3]^T$	IR of shaping filter for $x_1(n)$.
$\mathbf{h}_{x_2}$	$[1, -0.2]^T$	IR of shaping filter for $x_2(n)$.
$\mathbf{h}_{x_3}$	$[1, 0.5]^T$	IR of shaping filter for $x_3(n)$.
κ_2	-0.3	Mixing factor for $x_2(n)$.
κ_3	0.45	Mixing factor for $x_3(n)$.

and predicted results at frequency bin $l = 7$ are shown in Figs. 3(c)-(d). Again, the simulation results support the theory.

6.1.2 Simulation Experiment for NLMS Algorithm

We now repeat the experiment from above for the NLMS algorithm. In the first part of the experiment, a fixed step size is chosen as $\bar{\mu} = 2^{-4}$ and the regularization term δ is

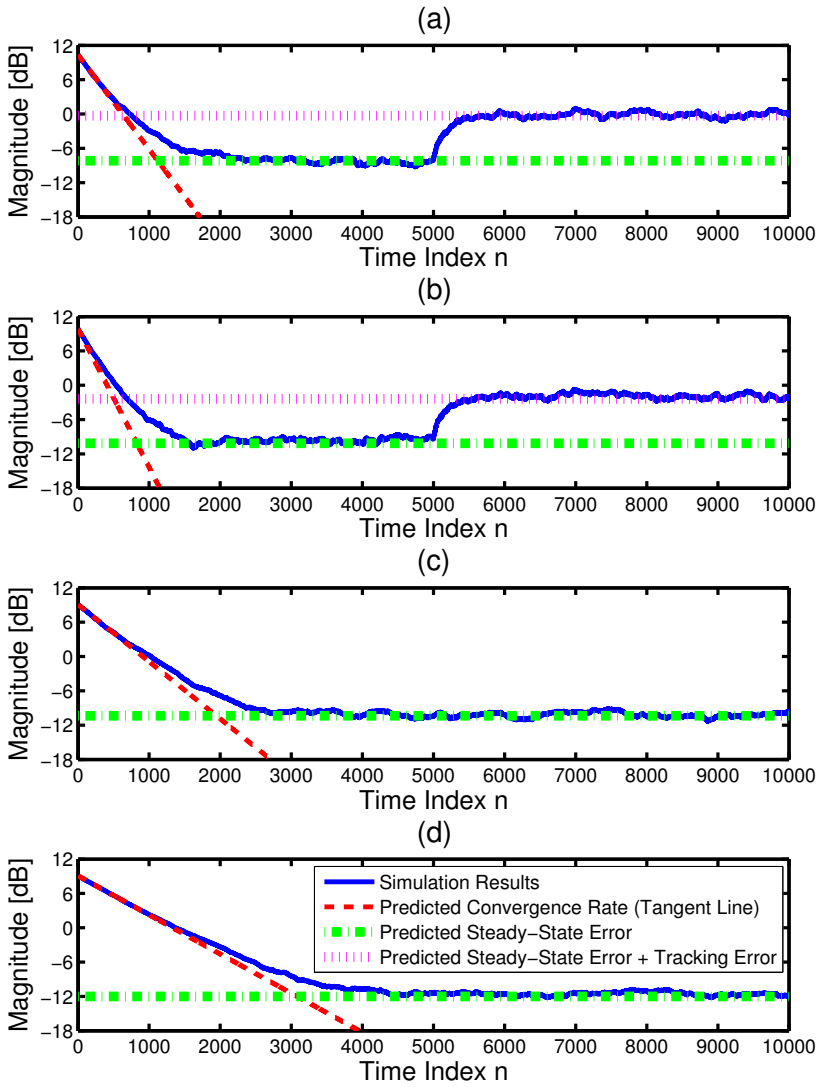


Fig. 3: LMS algorithm: the simulation results based on 100 simulation runs and the predicted values (a)-(b) Using Eqs. (49) and (51) at frequency bin $l = 7$ and $l = 11$. (c)-(d) Using Eqs. (52)-(53) at frequency bin $l = 7$. (a) PTF at frequency bin $l = 7$. (b) PTF at frequency bin $l = 11$. (c) Desired convergence rate $= -0.01$ dB/iteration. (d) Desired steady-state error $= -12$ dB.

Table 3: Difference between the simulation results and predicted values for the convergence rate (CR), steady-state error (SSE) and the tracking error (TE) in Fig. 3(a).

	Time Index	**Difference [dB]**				
CR	1:5	0.001	0.002	0.002	-0.002	-0.000
SSE	4996:5000	0.051	0.090	0.180	0.201	0.179
TE	9996:10000	0.126	0.095	0.040	0.062	0.130

set to zero[3]. Hence, the only difference to the LMS algorithm is the step size $\bar{\mu}$. Figs. 4(a)-(b) show the results for two representative frequencies $l = 7, 11$, to verify Eqs. (54) and (56).

In the second part of the experiment, using Eqs. (57)-(58), the step size values are calculated as $\bar{\mu} \approx 0.0412$ and $\bar{\mu} \approx 0.0283$ to achieve a desired convergence rate of -0.01 dB/iteration and a steady-state error of -12 dB, respectively. In Figs. 4(c)-(d), the simulation results are provided.

As for the LMS algorithm, the simulation results support the predicted results, and the derived expressions are accurate for practical values such as $L = 32$ and $\bar{\mu} = 2^{-4}$.

The NLMS algorithm can be considered an LMS algorithm with adjusted step size according to Eq. (36); thus, with an appropriate step size choice, it is possible to obtain identical system behavior for the two algorithms, at all frequencies. This is already implicitly done in the second part of this experiment, when we chose the step size parameters μ and $\bar{\mu}$ according to the desired convergence rate and steady-state value. It is seen from Figs. 3(c)-(d) and 4(c)-(d) that identical behavior at frequency bin $l = 7$ is obtained. Actually, making similar plots for other frequencies would demonstrate that this is the case at all frequencies.

6.1.3 Simulation Experiment for RLS Algorithm

In the first part of the experiment for the RLS algorithm, the forgetting factor $\lambda = 0.999$ was used to verify Eqs. (59) and (61). In the second part, the forgetting factor λ was set to $\lambda \approx 0.9989$ and $\lambda \approx 0.9992$, by using Eqs. (62)-(63), to obtain a desired convergence rate of -0.01 dB/iteration and a steady-state error of -12 dB, respectively.

It is seen in Fig. 5 that the simulation results again support the predicted results. Again, we observe that the derived expressions are already accurate for practical values such as $L = 32$ and $\lambda = 0.999$. Comparing Figs. 4(c)-(d) and 5(c)-(d) shows that identical behaviors for the NLMS and RLS algorithms are obtained at one specific frequency.

[3]The regularization term δ, besides the step size parameter $\bar{\mu}(n)$, might have significant influence on the performance of the NLMS algorithm, especially with a large step size $\bar{\mu}(n)$, as demonstrated in [38]. However, in the following simulations, we apply a small step size and focus on its influence on the system behavior in terms of the PTF by setting $\delta = 0$.

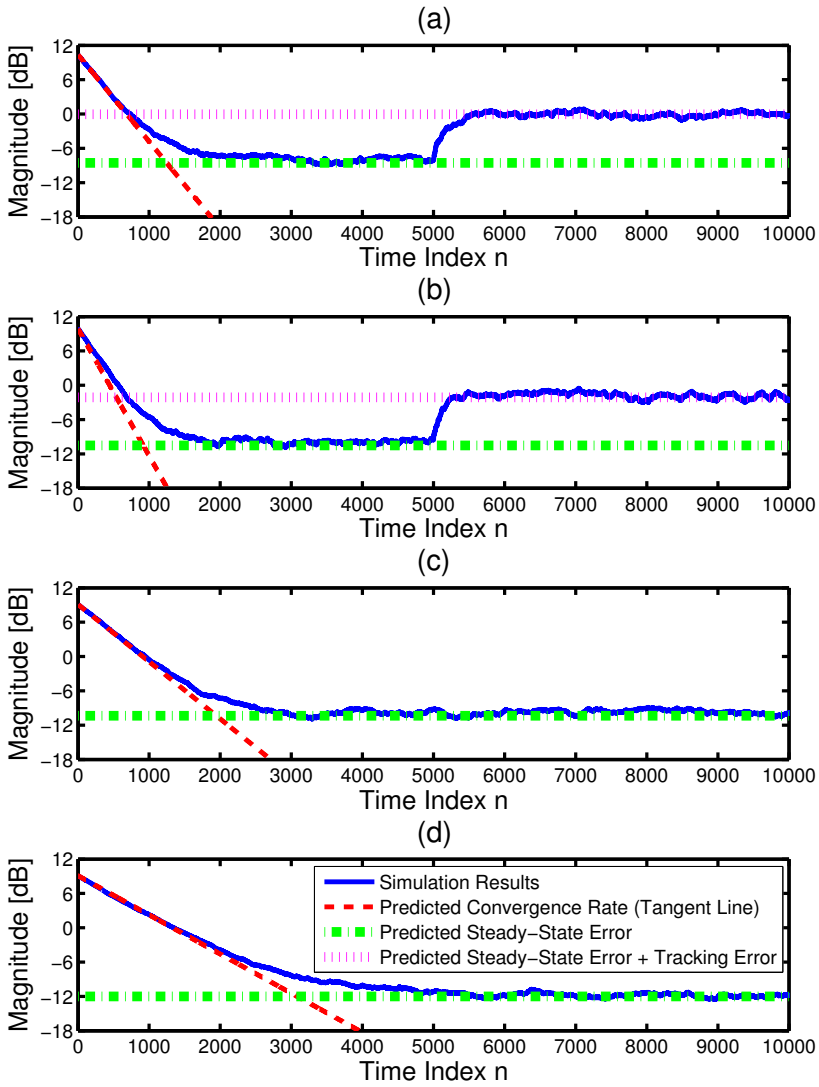


Fig. 4: NLMS algorithm: the simulation results based on 100 simulation runs and the predicted values (a)-(b) Using Eqs. (54) and (56) at frequency bin $l = 7$ and $l = 11$. (c)-(d) Using Eqs. (57)-(58) at frequency bin $l = 7$. (a) PTF at frequency bin $l = 7$. (b) PTF at frequency bin $l = 11$. (c) Desired convergence rate $= -0.01$ dB/iteration. (d) Desired steady-state error $= -12$ dB.

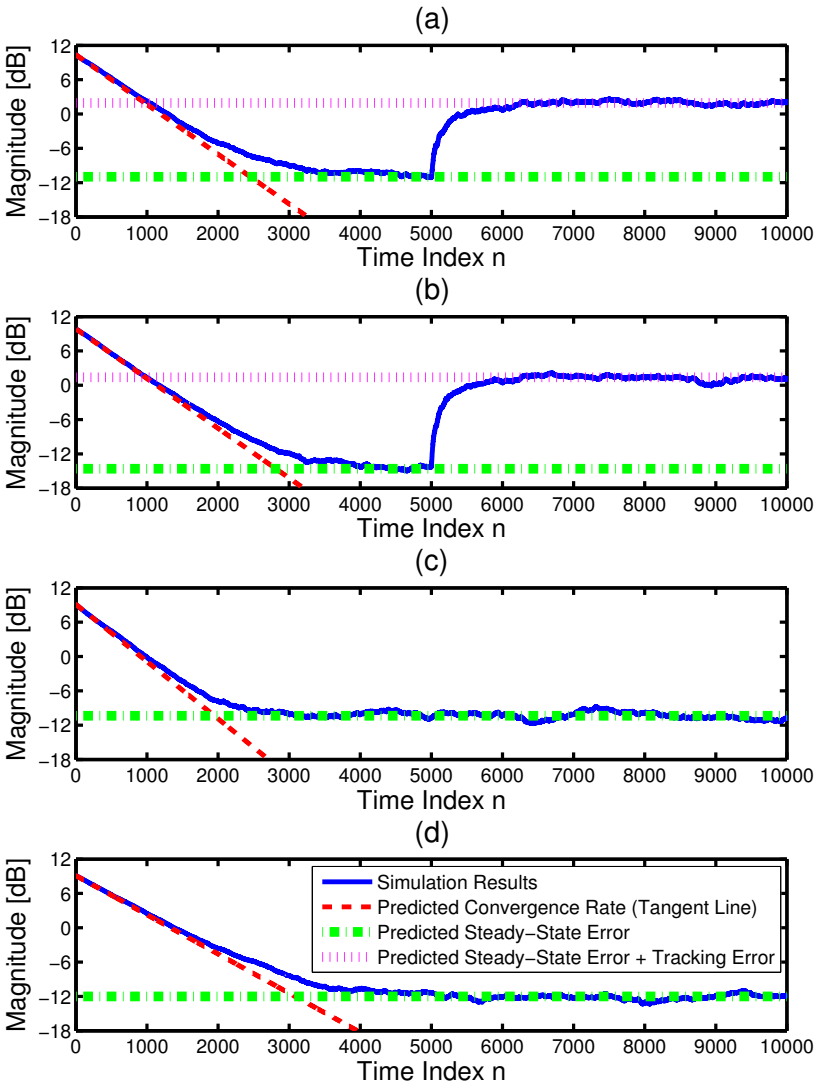


Fig. 5: RLS algorithm: the simulation results based on 100 simulation runs and the predicted values (a)-(b) Using Eqs. (59) and (61) at frequency bin $l = 7$ and $l = 11$. (c)-(d) Using Eqs. (62)-(63) at frequency bin $l = 7$. (a) PTF at frequency bin $l = 7$. (b) PTF at frequency bin $l = 11$. (c) Desired convergence rate $= -0.01$ dB/iteration. (d) Desired steady-state error $= -12$ dB.

However, in this given example, the behaviors of these two algorithms are different at all other frequencies because of the different values of $S_u(\omega)$ across frequencies.

6.2 Simulation Experiment for Acoustic Feedback Cancellation

As discussed in Sec. 3, the PTF derivation is performed under open-loop assumption, i.e. by omitting $\mathbf{f}(n)$ in Fig. 2. In closed-loop systems, the loudspeaker signal $u(n)$ is a processed and delayed version of the incoming signals $x_i(n)$. Thus, the correlation between $u(n)$ and $x_i(n)$ is nonzero for many natural signals, e.g. speech signals. This nonzero correlation is a general problem in closed-loop system identification as it leads to a biased solution in the adaptive estimation algorithms, see e.g. [6]. Besides, the assumption of uncorrelated $u(n)$ and $x_i(n)$ for the PTF derivation is violated.

Nevertheless, as demonstrated through this simulation experiment, the derived PTF expressions can be useful for predicting system behavior even in a closed-loop AFC application, e.g. in a hearing aid system.

6.2.1 Background of Feedback Problem in Hearing Aids

In hearing aids, acoustic feedback typically occurs in high frequency regions for two reasons. First, a larger amplification is usually implemented in the forward path $\mathbf{f}(n)$ for high frequencies above 3-5 kHz following the typical hearing loss patterns for the hearing aid users [11, 23]. Second, the peak magnitude response of the acoustic feedback paths in hearing aids represented by $\mathbf{h}_i(n)$ in Fig. 2 is typically located in the frequency region above 5 kHz due to the resonance frequency caused by the ventilation canals in the ear plugs [39].

Thus, feedback cancellation in hearing aids is particularly necessary in the high frequency region above 3 kHz. For many everyday signals including speech signals, the correlation functions in the critical frequency region above 3-5 kHz often decay rapidly for increasing correlation lags. For a typical hearing aid processing delay of 5-7 ms [24] in the forward path $\mathbf{f}(n)$, the correlation between the loudspeaker signal $u(n)$ and the incoming signal $x_i(n)$ at these frequencies is rather low. Thus, it is expected that the assumption of uncorrelated $u(n)$ and $x_i(n)$ is approximately valid in this high frequency region. However, this depends strongly on the incoming signals $x_i(n)$, e.g. for musical signals with many sustained high frequency components, the prediction of the system behavior using the derived PTF expressions would be less accurate.

6.2.2 Simulation Experiment

There exists a wide range of practical attempts for reducing the bias caused by the correlation problem in hearing aid AFC systems including band-limited estimation [40, 41], where the estimation of acoustic feedback paths is only carried out in a limited frequency band, typically above 1-2 kHz, selective step size algorithms [42] where the

step size parameter is adjusted according to an estimated correlation function between $u(n)$ and $x_i(n)$, and decorrelation methods [22] where e.g. frequency shifting is used to actively decorrelate $u(n)$ and $x_i(n)$. Applying these methods would reduce the bias problem further and allows an even better PTF prediction. However, our goal here is, to avoid these complications and to demonstrate that our analysis is approximately valid even without these methods.

We consider an MMSL system with $P = 2$ microphones, the NLMS algorithm with relatively small step size $\bar{\mu} = 2^{-9}$ and a regularization parameter $\delta = 2^{-30}$ is used for a fullband feedback path estimation and cancellation. The duration of the simulation is 90 s. The incoming signal $x_1(n)$ is a female speech, and $x_2(n)$ is a delayed version of $x_1(n)$ by one sample using a sampling frequency of $f_s = 20$ kHz. This models a situation where the sound signal is coming from the frontal direction of the hearing aid, and the distance between the front and rear microphones in the hearing aid is about 15 mm. The beamformer filters are simply set to $\mathbf{g}_1 = \mathbf{g}_2 = \frac{1}{2}$. The acoustic feedback paths $\mathbf{h}_i(n)$ are obtained from measurements of a behind-the-ear hearing aid. The impulse responses, magnitude responses and phase responses for both microphone channels are shown in Fig. 6.

The first 64 taps of the measured feedback paths $\mathbf{h}_i(n)$ are shown in Fig. 6. The true length of $\mathbf{h}_i(n)$ is unknown but considered to be about 50-55 taps, which corresponds to about 2.6 ms. In the simulation experiment, the length of adaptive filter is chosen to be $L = 64$. Furthermore, feedback path variations are added during the last 30 s of the simulation, similarly to the first experiment; the variances are $\sigma^2_{h_1} = 5.3 \cdot 10^{-5}$ and $\sigma^2_{h_2} = 3.8 \cdot 10^{-5}$, respectively. The forward path $\mathbf{f}(n)$ consists of a pure delay of 120 samples corresponding to 6 ms to model the input-to-output processing delay in a hearing aid, and a single-channel fullband compressor [11] to provide the amplification as a function of the energy of its input signal $\bar{e}(n)$ over time. In our simulation, the compressor provided a fullband amplification of 20 dB for the entire simulation. In this way, the most critical frequency is found at approximately 7.7 kHz, where the magnitude value of the OLTF is approximately -1 dB and the phase value is 0 rad.

To obtain the PTF prediction values, we computed the long-term PSDs $S_{x_{ij}}(\omega)$ and $S_u(\omega)$ of the incoming signals $x_i(n)$ and the loudspeaker signal $u(n)$ and inserted them in Eqs. (54) and (56). Using long-term PSDs is motivated by the fact that for small step size $\bar{\mu}$, the adaptive algorithm has a low-pass effect on the adaptive estimation and the PTF. In this simulation experiment, the PSD estimates are obtained from the entire signal sequences $x_i(n)$ and $u(n)$.

As in the first experiment, we compute the true PTF $\xi(\omega, n)$ using Eq. (19). The results from just one simulation run is shown in Fig. 7, where the simulation and prediction results for the frequency bin at 7.8 kHz are given. It is seen that especially the convergence rate and the steady-state value for the time-varying system are well predicted. The steady-state values for the time invariant system obtained from the simulation might be slightly biased, but they are still close to the predicted value. These

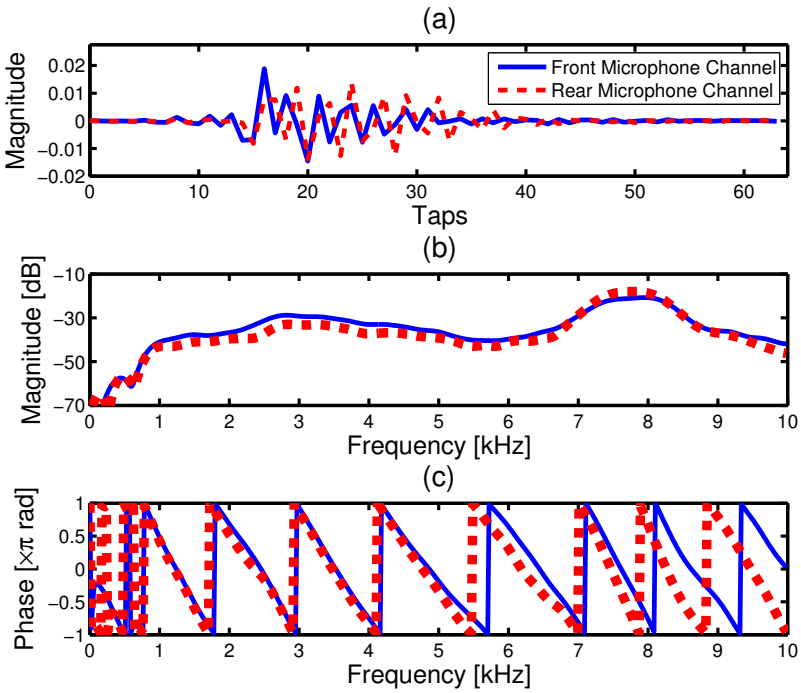


Fig. 6: Measured acoustic feedback paths from a behind-the-ear hearing aid with a sampling frequency $f_s = 20$ kHz. (a) Impulse response. (b) Magnitude response. (c) Phase response.

results are obtained despite the fact that the magnitude of the OLTF at the start of the simulation $n = 0$ was as high as -1 dB and a speech signal is used as the incoming signals $x_i(n)$. In this simulation experiment, the PTF predictions for frequencies below 2-3 kHz were less precise due to the correlation between $u(n)$ and $x_i(n)$.

The relatively rapid variations in the simulation results between 25-60 s in Fig. 7 are caused by the dynamics in the speech signal. Thus, there are mismatches between the true instantaneous PSDs and the long-terms estimates $S_{x_{ij}}(\omega)$ and $S_u(\omega)$ used in the PTF predictions. More precise and time-varying PSD estimates over shorter time duration could be used to improve the prediction. Ideally, the time duration should match the time constant of the averaging effect caused by the applied step size. It should also be noted that the simulation result in Fig. 7 is from one simulation run only, which corresponds to a practical situation. Much smoother simulation curves similar to the ones in Figs. 3-5 can be expected, if more simulation runs were carried out and an average result was computed.

In practice, the incoming signals $x_i(n)$ are unknown, and it is therefore not possible

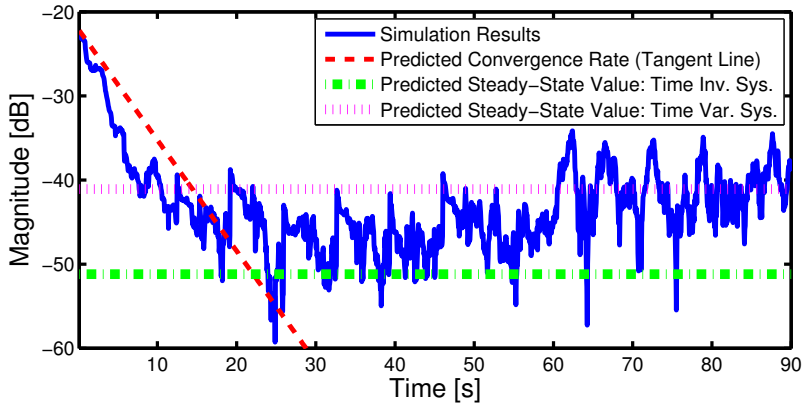


Fig. 7: Simulation results based on one simulation run in a closed-loop hearing aid acoustic feedback cancellation system with a female speech as the incoming signal. The results are given at the frequency bin 25 ($\sim$ 7.8 kHz), closest to the most critical frequency of 7.7 kHz.

to compute $S_{x_{ij}}(\omega)$ directly. However, from the derived PTF expressions Eqs. (49) and (51) for the LMS algorithm, Eqs. (54) and (56) for the NLMS algorithm, and Eqs. (59) and (61) for the RLS algorithm, it can be seen that $S_{x_{ij}}(\omega)$ only have influence on the steady-state error. This allows the use of the feedback/echo compensated signals $e_i(n)$ as estimates of $x_i(n)$ upon convergence of the adaptive filters $\hat{\mathbf{h}}_i(n)$. For this simulation experiment, estimating $S_{x_{ij}}(\omega)$ based on $e_i(n)$ led to prediction results of the steady-state values within 1 dB of those shown in Fig. 7.

7 Conclusion

In this work, we dealt with acoustic feedback/echo cancellation in a multiple-microphone and single-loudspeaker audio processing system. We derived analytic expressions for a frequency domain measure referred to as the power transfer function. These expressions are used to predict the feedback/echo cancellation performance in terms of the convergence rate, system stability bound and the steady-state behavior for the entire audio processing system at each frequency and time instant. The power transfer function is determined as a function of system parameters, e.g. the estimation filter order, and the statistical properties of different signals. We showed that the derived power transfer function approximations can be used to control system parameters, e.g. the step size parameter in the adaptive cancellation algorithms in order to achieve desired behaviors such as convergence rate and steady-state behavior at a specific frequency.

We considered three example adaptive algorithms, namely the least mean square,

normalized least mean square, and the recursive least squares, and we compared their system behaviors in terms of the power transfer function. Furthermore, we related the derived power transfer function expressions to other well-known fullband analysis results for single-microphone systems from existing works. Finally, the derived results are successfully verified by simulations using synthetic signals, and in a closed-loop hearing aid acoustic feedback cancellation system using real data including speech signals.

A Estimation Error Correlation Matrix

In this appendix, we derive Eq. (29) from Eq. (27). The estimation error vector $\tilde{\mathbf{h}}_i(n)$ in Eq. (23) can be expressed by

$$\begin{aligned}
\tilde{\mathbf{h}}_i(n) =& \mathbf{A}(n)\tilde{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{u}(n)x_i(n) - \check{\mathbf{h}}_i(n) \\
=& \mathbf{A}(n)\left(\mathbf{A}(n-1)\tilde{\mathbf{h}}_i(n-2) + \mu(n-1)\mathbf{u}(n-1)x_i(n-1) - \check{\mathbf{h}}_i(n-1)\right) \\
&+ \mu(n)\mathbf{u}(n)x_i(n) - \check{\mathbf{h}}_i(n) \\
=& \dots \\
=& \prod_{l=1}^{n}\mathbf{A}(l)\tilde{\mathbf{h}}_i(0) + \sum_{m=1}^{n}\left(\prod_{l=m+1}^{n}\mathbf{A}(l)\right)\left(\mu(m)\mathbf{u}(m)x_i(m) - \check{\mathbf{h}}_i(m)\right), \qquad (68)
\end{aligned}$$

where we use that $\prod_{l=n_0}^{n}\mathbf{A}(l) = \mathbf{I}$ for $n_0 > n$. Furthermore, we assume that the adaptation starts at $n = 0$.

Inserting Eq. (68) in the second last term of Eq. (27) and under the assumption that $\tilde{\mathbf{h}}_i(0) = \hat{\mathbf{h}}_i(0) - \mathbf{h}_i(0)$ is uncorrelated with $\mathbf{u}(n)$, we get

$$\begin{aligned}
E\left[\mu(n)\mathbf{u}(n)x_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right] =& \mu(n)E\Bigg[\mathbf{u}(n)x_i(n) \\
&\cdot\left(\prod_{l=1}^{n-1}\mathbf{A}(l)\tilde{\mathbf{h}}_j(0) + \sum_{m=1}^{n-1}\left(\prod_{l=m+1}^{n-1}\mathbf{A}(l)\right)\right. \\
&\left.\cdot\left(\mu(m)\mathbf{u}(m)x_j(m) - \check{\mathbf{h}}_j(m)\right)\right)^T\Bigg] \\
=& \mu(n)\sum_{k=-1}^{-(n-1)}\mu(n+k)\mathbf{u}(n)\mathbf{u}^T(n+k)r_{x_{ij}}(k) \\
&\cdot\left(\prod_{l=n+k+1}^{n-1}\mathbf{A}(l)\right)^T, \text{ where } k = m-n. \qquad (69)
\end{aligned}$$

It can be shown that Eq. (69) only influences the steady-state behavior of the PTF $\xi(\omega, n)$. Thus, it is sufficient to consider the situation where n is large, specifically, for the case $n-1 \geq k_0$. Using Eq. (28) and considering the case where $n-1 \geq k_0$, Eq. (69) can be further expressed by

$$E\left[\mu(n)\mathbf{u}(n)x_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right] = \mu(n) \cdot \sum_{k=-1}^{-k_0} \mu(n+k)\mathbf{u}(n)\mathbf{u}^T(n+k)r_{x_{ij}}(k)\left(\prod_{l=n+k+1}^{n-1}\mathbf{A}(l)\right)^T. \tag{70}$$

The matrix $\prod_{l=n+k+1}^{n-1}\mathbf{A}(l)$ in Eq. (70) can be simplified. Considering the case $k=-k_0$ where this matrix contains most factors expressed by

$$\begin{aligned}\prod_{l=n-k_0+1}^{n-1}\mathbf{A}(l) = &\prod_{l=n-k_0+1}^{n-1}\left(\mathbf{I}-\mu(l)\mathbf{u}(l)\mathbf{u}^T(l)\right)\\ =&\mathbf{I}-\mu(n-k_0+1)\mathbf{u}(n-k_0+1)\mathbf{u}^T(n-k_0+1)\\ &+\ldots+\prod_{l=n-k_0+1}^{n-1}\left(-\mu(l)\mathbf{u}(l)\mathbf{u}^T(l)\right).\end{aligned} \tag{71}$$

Inserting Eq. (71) in Eq. (70), it can be seen, that all terms besides $\mathbf{I}$ in Eq. (71) result in higher order terms involving $\mu(n)\mu(n+k)\mathbf{u}(n)\mathbf{u}^T(n+k)$ in Eq. (70) and are thereby neglected. Hence, Eq. (70) can now be expressed by

$$\begin{aligned}E\left[\mu(n)\mathbf{u}(n)x_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right] =&\mu(n)\sum_{k=-1}^{-k_0}\mu(n+k)\mathbf{u}(n)\mathbf{u}^T(n+k)r_{x_{ij}}(k)\\ =&\mu^2(n)\sum_{k=-1}^{-k_0}\mathbf{R}_u(k)r_{x_{ij}}(k).\end{aligned} \tag{72}$$

The last line in Eq. (72) is carried out under the assumption of slowly varying step size $\mu(n)$ over time, so that $\mu(n)\mu(n-k_0)=\mu^2(n)$. This requires that the variation in $\mu(n)$ must be slower than the decay of $r_{x_{ij}}(k)$. Furthermore, we replaced $\mathbf{u}(n)\mathbf{u}^T(n+k)$ by $\mathbf{R}_u(k)$ using the direct-averaging method.

Similarly, we can express the last term in Eq. (27) as

$$E\left[\tilde{\mathbf{h}}_i(n-1)x_j(n)\mathbf{u}^T(n)\mu(n)\right] = \mu^2(n)\sum_{k=1}^{k_0}\mathbf{R}_u(k)r_{x_{ij}}(k). \tag{73}$$

Finally, inserting Eqs. (72)-(73) in Eq. (27), we get Eq. (29).

Acknowledgment

The authors would like to thank the reviewers and the associate editor for their valuable suggestions and comments.

References

[1] E. T. Patronis, Jr., "Electronic detection of acoustic feedback and automatic sound system gain control," *J. Audio Eng. Soc.*, vol. 26, no. 5, pp. 323–326, May 1978.

[2] L. N. Mishin, "A method for increasing the stability of sound amplification systems," *Sov. Phys. Acoust.*, vol. 4, pp. 64–71, 1958.

[3] M. R. Schroeder, "Improvement of acoustic-feedback stability by frequency shifting," *J. Acoust. Soc. Am.*, vol. 36, no. 9, pp. 1718–1724, Sep. 1964.

[4] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, US: Prentice Hall, Sep. 2001.

[5] A. H. Sayed, *Fundamentals of Adaptive Filtering.* Hoboken, NJ, US: Wiley, Jun. 2003.

[6] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Elsevier J. Franklin Inst.*, vol. 343, no. 6, pp. 545–573, Sep. 2006.

[7] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio applications," *Elsevier Signal Process.*, vol. 89, no. 11, pp. 2185–2201, Nov. 2009.

[8] C. Paleologu, J. Benesty, and S. Ciochină, "A variable step-size affine projection algorithm designed for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 8, pp. 1466–1478, Nov. 2008.

[9] P. Loganathan, A. W. H. Khong, and P. A. Naylor, "A class of sparseness-controlled algorithms for echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 8, pp. 1591–1601, Nov. 2009.

[10] D.-Q. Nguyen, W.-S. Gan, and A. W. H. Khong, "Time-reversal approach to the stereophonic acoustic echo cancellation problem," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 2, pp. 385–395, Feb. 2011.

[11] H. Dillon, *Hearing Aids.* Stuttgart, Germany: Thieme, May 2001.

[12] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *IRE WESCON Conv. Record, Part 4*, pp. 96–104, 1960.

[13] G. Long, F. Ling, and J. G. Proakis, "The LMS algorithm with delayed coefficient adaptation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 9, pp. 1397–1405, Sep. 1989.

[14] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification – a survey," *Elsevier Automatica*, vol. 26, no. 1, pp. 7–21, Jan. 1990.

[15] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2937–2944, Nov. 1994.

[16] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Heidelberg, Germany: Springer, May 2001.

[17] S. Koike, "A class of adaptive step-size control algorithms for adaptive filters," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1315–1326, Jun. 2002.

[18] A. W. H. Khong and P. A. Naylor, "Selective-tap adaptive filtering with performance analysis for identification of time-varying systems," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 5, pp. 1681–1695, Jul. 2007.

[19] Y. Avargel and I. Cohen, "Adaptive system identification in the short-time Fourier transform domain using cross-multiplicative transfer function approximation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 1, pp. 162–173, Jan. 2008.

[20] P. Loganathan, E. A. P. Habets, and P. A. Naylor, "Performance analysis of IPNLMS for identification of time-varying systems," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 317–320.

[21] V. H. Nascimento, M. T. M. Silva, L. A. Azpicueta-Ruiz, and J. Arenas-García, "On the tracking performance of combinations of least mean squares and recursive least squares adaptive filters," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 3710–3713.

[22] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

[23] B. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. Bingley, UK: Emerald Group Publishing Limited, Apr. 2003.

[24] A. Spriet, M. Moonen, and J. Wouters, "Evaluation of feedback reduction techniques in hearing aids based on physical performance measures," *J. Acoust. Soc. Am.*, vol. 128, no. 3, pp. 1245–1261, Sep. 2010.

[25] H. Nyquist, "Regeneration theory," *Bell System Tech. J.*, vol. 11, pp. 126–147, 1932.

[26] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Heidelberg, Germany: Springer, Jun. 2001.

[27] S. Gunnarsson and L. Ljung, "Frequency domain tracking characteristics of adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 1072–1089, Jul. 1989.

[28] S. Gunnarsson, "On the quality of recursively identified FIR models," *IEEE Trans. Signal Process.*, vol. 40, no. 3, pp. 679–682, Mar. 1992.

[29] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of adaptive feedback and echo cancelation algorithms in a general multiple-microphone and single-loudspeaker system," in *Proc. 2011 IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 433–436.

[30] ——, "Comparison of multiple-microphone and single-loudspeaker adaptive feedback/echo cancellation systems," in *Proc. 19th European Signal Process. Conf.*, Aug. 2011, pp. 1279–1283.

[31] A. Lombard, K. Reindl, and W. Kellermann, "Combination of adaptive feedback cancellation and binaural adaptive filtering in hearing aids," *EURASIP J. Advances Signal Process.*, vol. 2009, pp. 1–15, Apr. 2009.

[32] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory*. Cambridge, MA, US: MIT Press, 1984.

[33] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. Hanover, MA, US: Now Publishers Inc., Jan. 2006.

[34] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

[35] D. R. Morgan and J. C. Thi, "A delayless subband adaptive filter architecture," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1819–1830, Aug. 1995.

[36] J. Huo, S. Nordholm, and Z. Zang, "New weight transform schemes for delayless subband adaptive filtering," in *Proc. IEEE Global Telecommunications Conf.*, vol. 1, Nov. 2001, pp. 197–201.

[37] A. Kaelin, A. Lindgren, and S. Wyrsch, "A digital frequency-domain implementation of a very high gain hearing aid with compensation for recruitment of loudness and acoustic echo cancellation," *Elsevier Signal Process.*, vol. 64, no. 1, pp. 71–85, Jan. 1998.

[38] J. Benesty, C. Paleologu, and S. Ciochină, "On regularization in adaptive filtering," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1734–1742, Aug. 2011.

[39] J. Hellgren, T. Lunner, and S. Arlinger, "Variations in the feedback of hearing aids," *J. Acoust. Soc. Am.*, vol. 106, no. 5, pp. 2821–2833, Nov. 1999.

[40] H.-F. Chi, S. X. Gao, S. D. Soli, and A. Alwan, "Band-limited feedback cancellation with a modified filtered-X LMS algorithm for hearing aids," *Elsevier Speech Commun.*, vol. 39, no. 1-2, pp. 147–161, Jan. 2003.

[41] J. M. Kates, "Feedback cancellation in a hearing aid with reduced sensitivity to low-frequency tonal inputs," United States Patent, US 6,831,986 B2, Dec. 2004.

[42] K. T. Klinkby, P. M. Norgaard, and H. P. Foeh, "Hearing aid, and a method for control of adaptation rate in anti-feedback systems for hearing aids," Int. Patent Application, WO 2007/113282 A1, Oct. 2007.

Paper B

Acoustic Feedback and Echo Cancellation Strategies for Multiple–Microphone and Single–Loudspeaker Systems

Meng Guo, Thomas Bo Elmedyb, Søren Holdt Jensen, and Jesper Jensen

Published in
Proc. 45th Asilomar Conf. Signals, Syst., Comput., Nov. 2011, pp. 556–560.

The layout has been revised.

Acoustic Feedback and Echo Cancellation Strategies for Multiple-Microphone and Single-Loudspeaker Systems

Meng Guo, Thomas Bo Elmedyb, Søren Holdt Jensen, and Jesper Jensen

Abstract

Acoustic feedback/echo cancellation in a multiple-microphone and single-loudspeaker system is often carried out using a cancellation filter for each microphone channel, and the filters are adaptively estimated, independently of each other. In this work, we consider another strategy by estimating all the cancellation filters jointly and in this way exploit information from all microphone channels. We determine the statistical system behavior for the joint estimation strategy in terms of the convergence rate and steady-state behavior across time and frequency. We assess if an improved cancellation performance is achievable compared to the independent estimation strategy. Furthermore, we relate the joint estimation strategy to a stereophonic acoustic echo cancellation system and provide analytic expressions for its system behavior.

1 Introduction

Acoustic feedback/echo problems occur when the microphone of a sound system picks up the acoustic output signal from the loudspeaker. In practical applications such as public address systems, teleconferencing systems and hearing aids, the acoustic feedback/echo problem often degrades the system performance. Feedback/echo cancellation using adaptive filters is one of the most applied methods to compensate for this problem, see e.g. [1, 2] and the references therein.

In this work, we focus on a multiple-microphone and single-loudspeaker (MMSL) system shown in Fig. 1, where the acoustic feedback/echo cancellation is performed using the adaptive filters $\hat{\mathbf{h}}_i(n)$ for estimating the true feedback/echo paths $\mathbf{h}_i(n)$, where n is the time index and P denotes the number of microphone channels. A beamformer performs spatial filtering of the feedback/echo compensated signals $e_i(n)$. This MMSL system could e.g. be a teleconferencing system or a hearing aid system. In acoustic feedback cancellation (AFC) applications, $\mathbf{f}(n)$ denotes a forward path e.g. to implement amplification in a hearing aid. In the area of acoustic echo cancellation (AEC), $\mathbf{f}(n)$

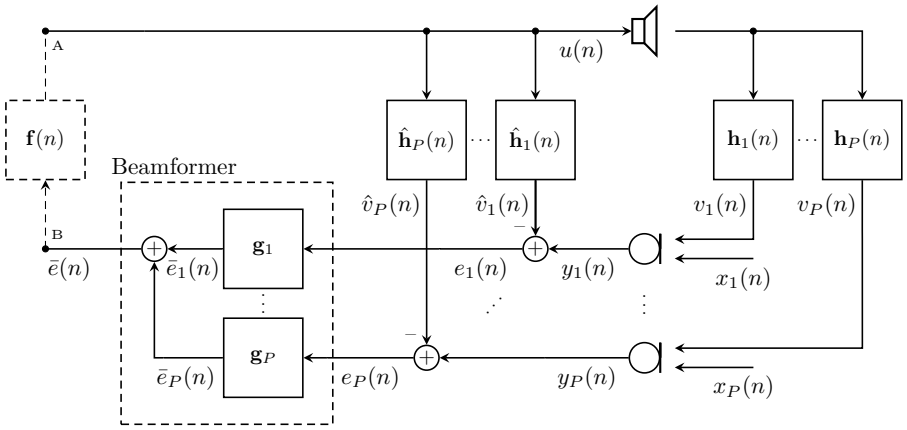


Fig. 1: A general multiple-microphone and single-loudspeaker system.

represents a far-end impulse response function and is usually ignored.

More specifically, we study estimation strategies for the adaptive filters $\hat{\mathbf{h}}_i(n)$ in this work. Very often, the estimation of these in MMSL systems is carried out independently for each microphone channel, e.g. by minimizing the error cost function $E[e_i^2(n)]$ using an adaptive algorithm such as the least mean square (LMS) algorithm [3]. We refer to this strategy as the independent estimation strategy.

Instead of focusing on each error signal $e_i(n)$ as in the independent estimation strategy, we consider minimizing the beamformer output signal as $E[(e(n) - \bar{e}(n))^2]$, where $e(n)$ is the ideal beamformer output signal at point B in Fig. 1, assuming no feedback signals, i.e. $\hat{v}_i(n) = v_i(n) = 0 \ \forall \ i$. In this way, we make use of all available information from different microphone channels in the adaptive system at once, instead of considering each of them separately. We refer to this strategy as the joint estimation strategy.

In contrast to the traditional independent estimation strategy, the joint estimation strategy studied here allows for a mismatch in each pair of $\hat{\mathbf{h}}_i(n)$ and $\mathbf{h}_i(n)$, as long as the sum of such mismatches lead to a smaller value (or faster convergence) of $E[(e(n) - \bar{e}(n))^2]$. It is clear that these two strategies lead to the same statistical behaviors, if all error signals $e_i(n)$ are mutually independent. However, this is obviously not the case when strong correlations can be expected between different microphone signals $y_i(n)$ due to the closely located microphones.

In this paper, we first derive analytical expressions for the convergence and steady-state behavior in the MMSL system using an example LMS algorithm[1] in the cancellation systems with the joint estimation strategy. This derivation is based on a recently

[1]We study the LMS algorithm for its simplicity, extending the results to other adaptive algorithms e.g. the normalized least mean square is possible.

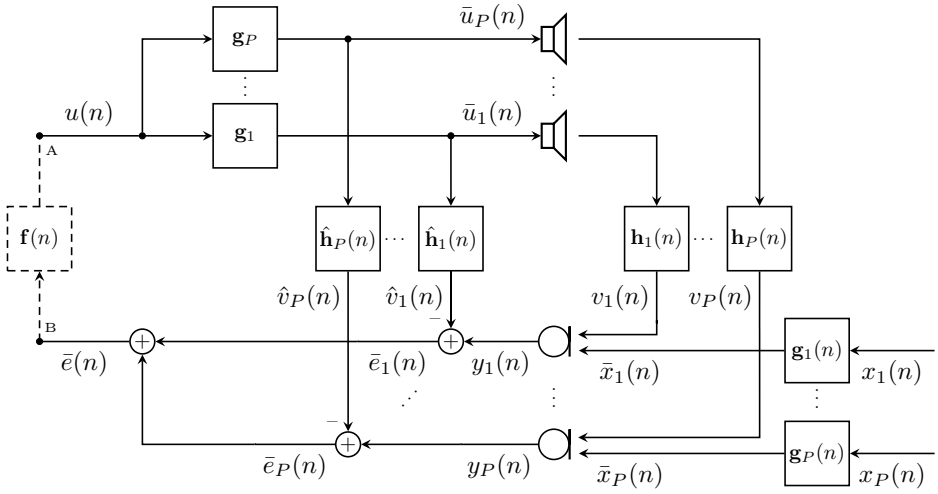


Fig. 2: The redrawn system with auxiliary loudspeakers and signals $\bar{u}_i(n)$ and $\bar{x}_i(n)$.

introduced frequency domain design and evaluation criterion referred to as the power transfer function (PTF) [4], inspired by the work in [5]. After that, we compare the derived results to the obtained results using the independent estimation strategy from an identical MMSL system [4], in terms of the convergence rate and steady-state behavior. Furthermore, we relate the results to a stereophonic acoustic echo cancellation system.

2 System Description

In order to ease the analysis, we redraw the MMSL system in Fig. 1 as the system shown in Fig. 2. The beamformer filters $\mathbf{g}_i$ are repositioned, and different auxiliary loudspeakers and signals $\bar{u}(n)$ and $\bar{x}(n)$ are created. It can be shown that both systems are identical, when minimizing $E[(e(n) - \bar{e}(n))^2]$ in the adaptive cancellation system, by assuming linear and time-invariant filters $\mathbf{g}_i$, $\mathbf{h}_i(n)$ and $\hat{\mathbf{h}}_i(n)$ at a given time index n. The following analysis is carried out on the redrawn system in Fig. 2.

The true but unknown feedback/echo path from the ith (auxiliary) loudspeaker to the ith microphone is modeled by a finite impulse response (FIR) of order $L-1$,

$$\mathbf{h}_i(n) = [h_i(0,n), \ldots, h_i(L-1,n)]^T . \tag{1}$$

The frequency response as the discrete Fourier transform (DFT) of $\mathbf{h}_i(n)$ is expressed

by

$$H_i(\omega, n) = \sum_{k=0}^{L-1} h_i(k, n) e^{-j\omega k}. \tag{2}$$

As in [4], we model time variations of the true feedback/echo paths $\mathbf{h}_i(n)$ using the random walk model

$$H_i(\omega, n) = H_i(\omega, n-1) + \check{H}_i(\omega, n), \tag{3}$$

where $\check{H}_i(\omega, n) \in \mathbb{C}$ is a sample from a zero-mean Gaussian stochastic process with variance

$$S_{\check{h}_{ii}}(\omega) = E\left[\check{H}_i(\omega, n)\check{H}_i^*(\omega, n)\right]. \tag{4}$$

In the time domain, the feedback/echo path variation vector is

$$\check{\mathbf{h}}_i(n) = \mathbf{h}_i(n) - \mathbf{h}_i(n-1). \tag{5}$$

The estimate $\hat{\mathbf{h}}_i(n)$ of the ith true feedback/echo path is given by

$$\hat{\mathbf{h}}_i(n) = \left[\hat{h}_i(0, n), \ldots, \hat{h}_i(L-1, n)\right]^T, \tag{6}$$

and the corresponding estimation error vector is defined as

$$\tilde{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n) - \mathbf{h}_i(n), \tag{7}$$

with a frequency response of

$$\tilde{H}_i(\omega, n) = \sum_{k=0}^{L-1} \tilde{h}_i(k, n) e^{-j\omega k}. \tag{8}$$

Each beamformer filter $\mathbf{g}_i$ is represented by an FIR filter

$$\mathbf{g}_i = [g_i(0), \ldots, g_i(N-1)]^T, \tag{9}$$

and its frequency response is

$$G_i(\omega) = \sum_{k=0}^{N-1} g_i(k) e^{-j\omega k}. \tag{10}$$

We consider the original loudspeaker signal $u(n)$ as a deterministic signal because it is measurable and thereby known. However, as argued in [5], our results remain valid,

even if the loudspeaker signal $u(n)$ is considered as a realization of a stochastic process, which is statistically independent of the incoming signals $x_i(n)$; this important point will be demonstrated by simulations in Sec. 6.

In the redrawn system, the auxiliary loudspeaker signal $\bar{u}_i(n)$ is expressed by

$$\bar{u}_i(n) = \sum_{k=0}^{N-1} g_i(k)u(n-k), \tag{11}$$

and the signal vector $\bar{\mathbf{u}}_i(n)$ is

$$\bar{\mathbf{u}}_i(n) = [\bar{u}_i(n), \ldots, \bar{u}_i(n-L+1)]^T. \tag{12}$$

We assume that the incoming signals $x_i(n)$ are zero-mean, stationary stochastic signals, and the auxiliary incoming signals are expressed by

$$\bar{x}_i(n) = \sum_{k=0}^{N-1} g_i(k)x_i(n-k). \tag{13}$$

Furthermore, the ith microphone signal $y_i(n)$ is modeled as

$$y_i(n) = \mathbf{h}_i^T(n-1)\bar{\mathbf{u}}_i(n) + \bar{x}_i(n). \tag{14}$$

The ith error signal $\bar{e}_i(n)$ is expressed by

$$\bar{e}_i(n) = y_i(n) - \hat{\mathbf{h}}_i^T(n-1)\bar{\mathbf{u}}_i(n). \tag{15}$$

The beamformer output signal $\bar{e}(n)$ is given by

$$\bar{e}(n) = \sum_{i=1}^{P} \bar{e}_i(n). \tag{16}$$

The impulse response $\mathbf{f}(n) = [f(0,n), \ldots, f(0,L_f-1)]^T$ has a frequency response $F(\omega,n) = \sum_{k=0}^{L_f-1} f(k,n)e^{-j\omega k}$.

3 Review of Power Transfer Function

In our analysis, we define the PTF as the expected magnitude-squared transfer function from point A to B in Fig. 2. The frequency responses $H_i(\omega,n)$ of the true feedback/echo

paths $\mathbf{h}_i(n)$ are unknown and considered as stochastic. Hence, as in [4], we define the exact PTF of the MMSL system as

$$\begin{aligned}\xi(\omega,n) &= E\left[\left|\sum_{i=1}^{P} G_i(\omega)\tilde{H}_i(\omega,n)\right|^2\right] \\ &= \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)\xi_{ij}(\omega,n),\end{aligned} \tag{17}$$

where $\xi_{ij}(\omega,n) = E[\tilde{H}_i(\omega,n)\tilde{H}_j^*(\omega,n)]$ and $*$ denotes complex conjugation.

In AFC systems, the PTF $\xi(\omega,n)$ is the unknown part of the expected magnitude-squared open-loop transfer function given by $E[|\mathrm{OLTF}(\omega,n)|^2] = |F(\omega,n)|^2\xi(\omega,n)$. If $|\mathrm{OLTF}(\omega,n)| < 1$, system stability is guaranteed [6]. For AEC applications, the PTF $\xi(\omega,n)$ is similar to the mean-square deviation $E[||\tilde{\mathbf{h}}(n)||^2]$ [3], but more importantly it describes the cancellation behavior, across frequency in addition to time, for the entire system rather than a single cancellation filter. Thus, for both AFC and AEC applications, the PTF is a useful criterion for designing and evaluating the cancellation system.

In general, however, we can not calculate the PTF $\xi(\omega,n)$ directly because $H_i(\omega,n)$ is unknown. In this work, we derive a simple but accurate approximation of the PTF as

$$\hat{\xi}(\omega,n) = \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)\hat{\xi}_{ij}(\omega,n), \tag{18}$$

where $\hat{\xi}_{ij}(\omega,n) \approx E[\tilde{H}_i(\omega,n)\tilde{H}_j^*(\omega,n)]$.

4 System Analysis

4.1 Review of MMSL System With Independent Estimation

In [4], we analyzed an MMSL system in terms of the PTF where feedback paths $\mathbf{h}_i(n)$ were estimated independently using the LMS algorithm. We showed the convergence rate describing the decay of $\hat{\xi}(\omega,n)$ is given by the coefficient

$$\alpha(\omega) = 1 - 2\mu_0(n)S_u(\omega), \tag{19}$$

where $\mu_0(n)$ is the LMS step size and $S_u(\omega)$ denotes the power spectrum density (PSD) of the loudspeaker signal $u(n)$. Furthermore, the steady-state behavior describing

$\hat{\xi}(\omega, n)$ upon convergence, is

$$\begin{aligned}\hat{\xi}(\omega,\infty) &= \lim_{n\to\infty} \hat{\xi}(\omega, n) \\ &= \underbrace{\lim_{n\to\infty} L\frac{\mu_0(n)}{2} \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}_{\text{Steady-State Error}} + \underbrace{\lim_{n\to\infty} \frac{\sum_{i=1}^{P}|G_i(\omega)|^2 S_{\check{h}_{ii}}(\omega)}{2\mu_0(n)S_u(\omega)}}_{\text{Tracking Error}}, \quad (20)\end{aligned}$$

where $S_{x_{ij}}(\omega)$ denotes the cross(auto) PSDs of the incoming signals $x_i(n)$ and $x_j(n)$. For details of the derivation of Eqs. (19) and (20), we refer to [4].

4.2 Analysis of MMSL System With Joint Estimation

Using the joint estimation strategy and considering $u(n)$ as a deterministic signal, it can be shown that the cost function $E[(e(n) - \bar{e}(n))^2]$ to be minimized is identical to

$$J(n) = E\left[\bar{e}^2(n)\right] = E\left[\left(\sum_{i=1}^{P}\bar{e}_i(n)\right)^2\right]. \quad (21)$$

The partial derivative of Eq. (21) is found as

$$\frac{\partial J(n)}{\partial \hat{\mathbf{h}}_i(n-1)} = -2E\left[\bar{\mathbf{u}}_i(n)\sum_{p=1}^{P}\bar{e}_p(n)\right]. \quad (22)$$

The LMS adaptation [3], using a step size $\mu(n)$, is given by

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\bar{\mathbf{u}}_i(n)\sum_{p=1}^{P}\bar{e}_p(n). \quad (23)$$

Using Eqs. (23), (15), (14) and (5), the ith feedback/echo path estimation error vector given by Eq. (7) is also described as

$$\tilde{\mathbf{h}}_i(n) = \tilde{\mathbf{h}}_i(n-1) + \mu(n)\bar{\mathbf{u}}_i(n)\sum_{p=1}^{P}\left(\bar{x}_p(n) - \bar{\mathbf{u}}_p^T\tilde{\mathbf{h}}_p(n-1)\right) - \check{\mathbf{h}}_i(n). \quad (24)$$

We assume that $\mu(n)$ varies slowly over time in addition to $\mu(n)$ is sufficiently small, in principle $\mu(n) \to 0$, and the correlation function $r_{x_{ij}}(k) = E[x_i(n)x_j(n-k)]$ of the incoming signals $x_i(n)$ and $x_j(n)$ fulfills $r_{x_{ij}}(k) = 0 \;\forall\; |k| > k_0 \in \mathbb{N}$. Furthermore, by

ignoring higher order terms, an approximation of the estimation error correlation matrix $\mathbf{H}_{ij}(n) = E[\tilde{\mathbf{h}}_i(n)\tilde{\mathbf{h}}_j^T(n)]$ can be written as

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =&\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\left(\sum_{p=1}^{P}\mathbf{R}_{\bar{u}_{ip}}(0)\hat{\mathbf{H}}_{pj}(n-1) + \hat{\mathbf{H}}_{ip}(n-1)\mathbf{R}_{\bar{u}_{pj}}(0)\right)\\ &+\mu^2(n)\sum_{p=1}^{P}\sum_{q=1}^{P}\sum_{k=-k_0}^{k_0}\mathbf{R}_{\bar{u}_{ij}}(k)r_{\bar{x}_{pq}}(k) + \check{\mathbf{H}}_{ij},\end{aligned} \tag{25}$$

under the assumption of $\mu(n) \to 0$ and using the direct-averaging method [7] to replace $\bar{\mathbf{u}}_i(n)\bar{\mathbf{u}}_j^T(n)$ by its sample average $\mathbf{R}_{\bar{u}_{ij}}(k) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\bar{\mathbf{u}}_i(n)\bar{\mathbf{u}}_j^T(n-k)$. In Eq. (25), $r_{\bar{x}_{ij}}(k) = E[\bar{x}_i(n)\bar{x}_j(n-k)]$ denotes the correlation function of the signals $\bar{x}_i(n)$ and $\bar{x}_j(n)$, and $\check{\mathbf{H}}_{ij} = E[\check{\mathbf{h}}_i(n)\check{\mathbf{h}}_j^T(n)]$ is the correlation matrix of the ith and jth feedback/echo path variations. Furthermore, we assume, for simplicity, that $\check{\mathbf{H}}_{ij} = \mathbf{0}_{L\times L}\ \forall\ i \neq j$.

Assuming a sufficiently large L, asymptotically as $L \to \infty$, we can use the DFT matrix $\mathbf{F} \in \mathbb{C}^{L\times L}$ to diagonalize the Toeplitz matrix $\hat{\mathbf{H}}_{ij}(n)$ [8]. Introduce the notation $G_{ij} = G_i^*(\omega)G_j(\omega)$. The values of $\hat{\xi}_{ij}(\omega, n)$ are obtained as the diagonal elements of the matrix $\mathbf{F}\hat{\mathbf{H}}_{ij}(n)\mathbf{F}^H$ and can be expressed by

$$\begin{aligned}\hat{\xi}_{ij}(\omega,n) =&\hat{\xi}_{ij}(\omega,n-1) - \mu(n)S_u(\omega)\sum_{p=1}^{P}\Big(G_{ip}\hat{\xi}_{pj}(\omega,n-1) + G_{pj}\hat{\xi}_{ip}(\omega,n-1)\Big)\\ &+L\mu^2(n)S_u(\omega)G_{ij}\sum_{p=1}^{P}\sum_{q=1}^{P}G_{qp}S_{x_{pq}}(\omega) + S_{\check{h}_{ij}}(\omega).\end{aligned} \tag{26}$$

Inserting Eq. (26) in Eq. (18), the PTF approximation $\hat{\xi}(\omega, n)$ is given by

$$\begin{aligned}\hat{\xi}(\omega,n) =& \left(1 - 2\mu(n)S_u(\omega)\sum_{i=1}^{P}|G_i(\omega)|^2\right)\hat{\xi}(\omega,n-1)\\ &+L\mu^2(n)S_u(\omega)\left(\sum_{i=1}^{P}|G_i(\omega)|^2\right)^2\sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)\\ &+\sum_{i=1}^{P}|G_i(\omega)|^2S_{\check{h}_{ii}}(\omega).\end{aligned} \tag{27}$$

5 Interpretation

5.1 System Behavior

Considering Eq. (27) as a first-order difference equation in $\hat{\xi}(\omega, n)$, the convergence rate is determined by the coefficient

$$\alpha(\omega) = 1 - 2\mu(n)S_u(\omega)\sum_{i=1}^{P}|G_i(\omega)|^2, \tag{28}$$

which determines the pole location of the first-order system. The convergence rate in dB/iteration is computed as $10\log_{10}|\alpha(\omega)|$. Therefore, increasing the values of $\mu(n)$, $S_u(\omega)$ and $\sum_{i=1}^{P}|G_i(\omega)|^2$ lead to higher convergence rate.

Algorithm stability is ensured if $|\alpha(\omega)| < 1$. Hence, using Eq. (28), the step size range to ensure stability is expressed by

$$0 < \mu(n) < \frac{1}{\max_\omega S_u(\omega)\sum_{i=1}^{P}|G_i(\omega)|^2}. \tag{29}$$

Furthermore, the steady-state behavior $\hat{\xi}(\omega,\infty) = \lim_{n\to\infty}\hat{\xi}(\omega, n)$ is given by

$$\begin{aligned}\hat{\xi}(\omega,\infty) = &\underbrace{\lim_{n\to\infty} L\frac{\mu(n)}{2}\sum_{i=1}^{P}|G_i(\omega)|^2\sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}_{\text{Steady-State Error}}\\ &+\underbrace{\lim_{n\to\infty}\frac{\sum_{i=1}^{P}|G_i(\omega)|^2 S_{\check{h}_{ii}}(\omega)}{2\mu(n)S_u(\omega)\sum_{i=1}^{P}|G_i(\omega)|^2}}_{\text{Tracking Error}}.\end{aligned} \tag{30}$$

The first term in Eq. (30) is the steady-state value of $\hat{\xi}(\omega, n)$ with time invariant feedback/echo paths, i.e. $S_{\check{h}_{ii}}(\omega) = 0$; the second term is the extra error contribution as a result of variations in the feedback/echo paths. In a time invariant system, $\hat{\xi}(\omega,\infty)$ is linearly proportional to L, $\mu(n)$, $\sum_{i=1}^{P}|G_i(\omega)|^2$, and the PSDs $S_{x_{ij}}(\omega)$ weighted by the frequency responses $G_i(\omega)$ and $G_j^*(\omega)$. On the other hand, it is seen from the second term that a larger values of $\mu(n)$, $S_u(\omega)$ and $\sum_{i=1}^{P}|G_i(\omega)|^2$ lead to smaller additional error when the feedback paths are undergoing variations. Therefore, the overall steady-state value $\hat{\xi}(\omega,\infty)$ is a compromise between the steady-state error and the tracking error, as in the single adaptive filter case [3].

More interestingly, by comparing Eq. (28) to Eq. (19) and Eq. (30) to Eq. (20), we observe that $\sum_{i=1}^{P}|G_i(\omega)|^2$ is the only difference between the independent and joint estimation strategies. Thus, by choosing the step size according to

$$\mu_0(\omega, n) = \mu(n)\sum_{i=1}^{P}|G_i(\omega)|^2, \tag{31}$$

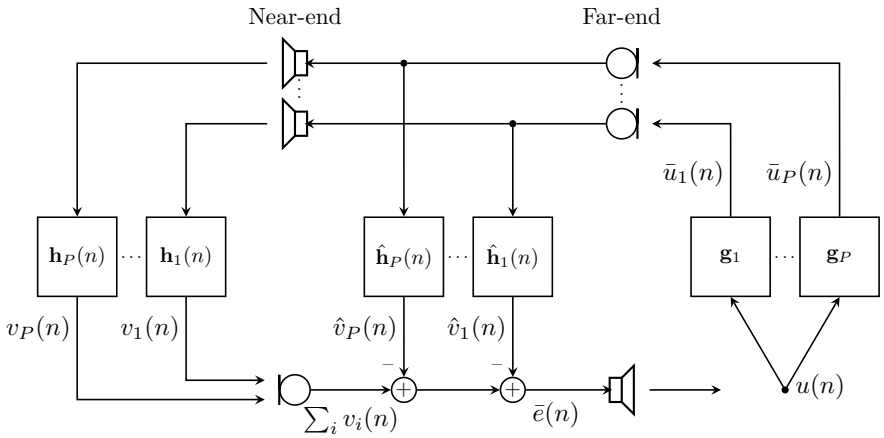


Fig. 3: A stereophonic acoustic echo cancellation system.

statistically identical behavior is obtained in both strategies. The LMS step size now becomes frequency dependent as $\mu_0(\omega, n)$, this can be achieved using e.g. frequency-domain adaptive filters [9]. Hence, perhaps somewhat surprisingly, there is no performance advantage by using the joint estimation strategy over the computationally simpler independent estimation strategy. Note that this conclusion and the relation in Eq. (31) are valid under the assumption of fixed or at least slowly varying beamformer filters $\mathbf{g}_i$.

5.2 Relation to Stereophonic Acoustic Echo Cancellation

From Fig. 2, we observe that the redrawn MMSL system is very similar in structure to a stereophonic acoustic echo cancellation (SAEC) system [1] given in Fig. 3. Consider $u(n)$ as a far-end source signal which is modified by the room impulse responses (RIRs) $\mathbf{g}_i$ into the far-end microphone signals $\bar{u}_i(n)$ which are transmitted to the near-end and played through loudspeakers. These loudspeaker signals are modified by the near-end RIRs $\mathbf{h}_i(n)$ before recorded by a near-end microphone. This corresponds to a summation $\sum_i v_i(n)$. The estimates $\hat{\mathbf{h}}_i(n)$ are used to cancel $\mathbf{h}_i(n)$, resulting in the echo canceled signals $\bar{e}(n)$.

The slight difference between these two systems, shown in Figs. 2 and 3, is that the summation of the signals $\bar{e}_i(n)$ in Fig. 2 is moved in front of the near-end microphone in the SAEC system, and only the mixture signal $\sum_i v_i(n)$, not the individual signal $v_i(n)$, is presented at the near-end microphone.

Hence, independent estimation strategy is not an option in the SAEC situation, and a joint estimation strategy minimizing the error signal $\bar{e}(n)$ must be carried out.

Thus, Eqs. (28) and (30) describe the cancellation behavior in terms of the convergence rate and steady-state behavior, over frequency and time, in an SAEC system with a corresponding near-end microphone signal $\sum_{i=1}^{P} \bar{x}_i(n)$.

6 Experiment

Simulations are carried out in an MMSL system with $P = 3$ microphones, to verify Eqs. (28) and (30).

In the simulations, the beamformer filters $\mathbf{g}_i$ and true feedback/echo paths $\mathbf{h}_i(n)$ are modeled by first-order FIR filters and thereby known. Thus, it is possible to compute the true PTF $\xi(\omega, n)$ given by Eq. (17) to verify the prediction values using Eqs. (28) and (30). The duration of each simulation run is 10^4 iterations, and 100 simulation runs are performed to obtain an averaged $\xi(\omega, n)$. In each simulation run, new realizations of standard Gaussian stochastic sequences are drawn; they are then filtered by various fixed first-order FIR shaping filters $\mathbf{h}_u$ and $\mathbf{h}_{x_i}$ to generate the loudspeaker signal $u(n)$ and the incoming signals $x_i(n)$, respectively.

The feedback/echo path estimates start from zeros, i.e. $\hat{\mathbf{h}}_i(0) = \mathbf{0}_{L\times 1}$, and $L = 32$. The true feedback/echo paths $\mathbf{h}_i(n)$, modeled by first-order FIR filters for simplicity, are fixed during the first half of the simulations, whereas random walk variations with Gaussian statistics $N(\mu_{h_i}, \sigma^2_{h_i})$ are added during the second half. Furthermore, a fixed step size $\mu = 2^{-9}$ is used. The numerical values of all simulation parameters are given in Table 1.

The simulated and predicted results, at a representative example frequency $\omega = 2\pi l/L$, $l = 7$, are shown in Fig. 4. The simulation results agree with the predicted convergence rate and steady-state values. The fact that a new realization of $u(n)$ is drawn for each simulation run demonstrates, as expected, that the derived results are valid, when the loudspeaker signal $u(n)$ is a realization of a stochastic process, even though we considered $u(n)$ as a deterministic signal in the analysis. For a detailed discussion of this, we refer to [5]. Furthermore, we also observed that the derived expressions are accurate for practical values as $L = 32$ and $\mu(n) = 2^{-9}$.

7 Conclusions

This work deals with acoustic feedback and echo cancellation strategies for multiple-microphone and single-loudspeaker systems. We analyzed a new strategy where the filters are estimated jointly in order to minimize the error signal after the beamformer in the system. We derived analytic expressions for the convergence rate and steady-state behavior, and these expressions are verified by simulations. Our analysis showed that with appropriately chosen parameter values, and under the assumption of fixed or slowly varying beamformer filters, the independent estimation strategy is statistically

Table 1: The simulation parameters.

Symbol	Value	Description
D_s	10^4	Duration of simulation.
R	100	Number of simulation runs.
P	3	Number of microphone channels.
L	32	Length of $\hat{\mathbf{h}}_i(n)$.
$\mu(n)$	2^{-9}	A fixed step size.
$\mathbf{g}_1$	$[1, 0.36]^T$	FIR of beamformer filter 1.
$\mathbf{g}_2$	$[1, -0.32]^T$	FIR of beamformer filter 2.
$\mathbf{g}_3$	$[1, 0.23]^T$	FIR of beamformer filter 3.
$\mathbf{h}_1(0)$	$[1, 0.14]^T$	Initial FIR of $\mathbf{h}_1(n)$.
$\mathbf{h}_2(0)$	$[1, -0.40]^T$	Initial FIR of $\mathbf{h}_2(n)$.
$\mathbf{h}_3(0)$	$[1, 0.23]^T$	Initial FIR of $\mathbf{h}_3(n)$.
$N(\mu_{h_1}, \sigma^2_{h_1})$	$N(0, 0.0384^2)$	Gaussian statistics for $\mathbf{h}_1(n)$.
$N(\mu_{h_2}, \sigma^2_{h_2})$	$N(0, 0.0332^2)$	Gaussian statistics for $\mathbf{h}_2(n)$.
$N(\mu_{h_3}, \sigma^2_{h_3})$	$N(0, 0.0024^2)$	Gaussian statistics for $\mathbf{h}_3(n)$.
$\mathbf{h}_u$	$[1, -0.3]^T$	FIR of shaping filter for $u(n)$.
$\mathbf{h}_{x_1}$	$[1, 0.3]^T$	FIR of shaping filter for $x_1(n)$.
$\mathbf{h}_{x_2}$	$[1, -0.2]^T$	FIR of shaping filter for $x_2(n)$.
$\mathbf{h}_{x_3}$	$[1, 0.5]^T$	FIR of shaping filter for $x_3(n)$.

identical to the joint estimation strategy. Hence, there is no performance advantages by using the computationally more complex joint estimation strategy in the considered system. Finally, we showed that the derived results of the joint estimation strategy also describe the behavior of a stereophonic acoustic echo cancellation system.

References

[1] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Heidelberg, Germany: Springer, May 2001.

[2] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

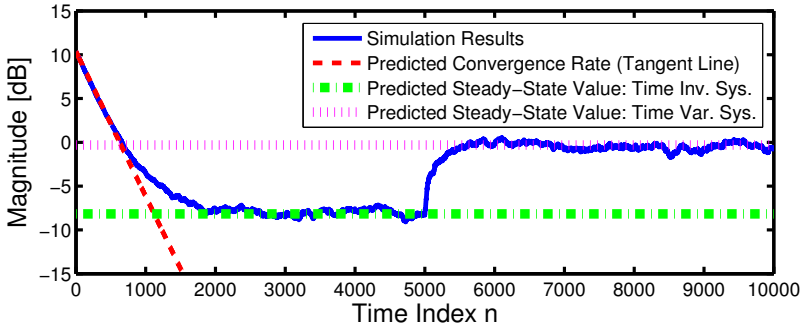


Fig. 4: PTF verification results of Eqs. (28) and (30), at the frequency $\omega = 2\pi l/L$, where $L = 32$ and $l = 7$.

[3] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, US: Prentice Hall, Sep. 2001.

[4] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of adaptive feedback and echo cancelation algorithms in a general multiple-microphone and single-loudspeaker system," in *Proc. 2011 IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 433–436.

[5] S. Gunnarsson and L. Ljung, "Frequency domain tracking characteristics of adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 1072–1089, Jul. 1989.

[6] H. Nyquist, "Regeneration theory," *Bell System Tech. J.*, vol. 11, pp. 126–147, 1932.

[7] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory.* Cambridge, MA, US: MIT Press, 1984.

[8] R. M. Gray, *Toeplitz and Circulant Matrices: A Review.* Hanover, MA, US: Now Publishers Inc., Jan. 2006.

[9] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

Paper C

On Acoustic Feedback Cancellation Using Probe Noise in Multiple–Microphone and Single–Loudspeaker Systems

Meng Guo, Thomas Bo Elmedyb, Søren Holdt Jensen, and Jesper Jensen

Published in
IEEE Signal Process. Lett., vol. 19, no. 5, pp. 283–286, May 2012.

The layout has been revised.

On Acoustic Feedback Cancellation Using Probe Noise in Multiple-Microphone and Single-Loudspeaker Systems

Meng Guo, Thomas Bo Elmedyb, Søren Holdt Jensen, and Jesper Jensen

Abstract

A probe noise signal can be used in an acoustic feedback cancellation system to prevent biased adaptive estimation of acoustic feedback paths. However, practical experiences and simulation results indicate that whenever a low-level and inaudible probe noise signal is used, the convergence rate of the adaptive estimation is significantly decreased when keeping the steady-state error unchanged. The goal of this work is to derive analytic expressions for the system behavior such as convergence rate and steady-state error for a multiple-microphone and single-loudspeaker audio system, where the acoustic feedback cancellation is carried out using a probe noise signal. The derived results show how different system parameters and signal properties affect the cancellation performance, and the results explain theoretically the decreased convergence rate. Understanding this is important for making further improvements in the existing probe noise approach.

1 Introduction

Acoustic feedback problems occur in audio systems when the microphone picks up part of the loudspeaker output signal. This problem often causes performance degradation in applications such as public address systems and hearing aids.

Acoustic feedback cancellation (AFC) using adaptive filters is one of the most applied methods to compensate for the feedback problem. The main problem of using adaptive filters for AFC is that the filter estimates become biased whenever the incoming signal and loudspeaker signal are correlated [1]. The bias generally leads to poor cancellation performance and in worst-case causes the cancellation system to fail. Different techniques have been proposed to prevent biased estimation including phase modification, frequency shifting, nonlinear processing, decorrelating prefilters and probe noise injection, see e.g. [2] and the references therein.

In this work, we focus on a probe noise approach in a multiple-microphone and single-loudspeaker (MMSL) audio system shown in Fig. 1. AFC is carried out using the

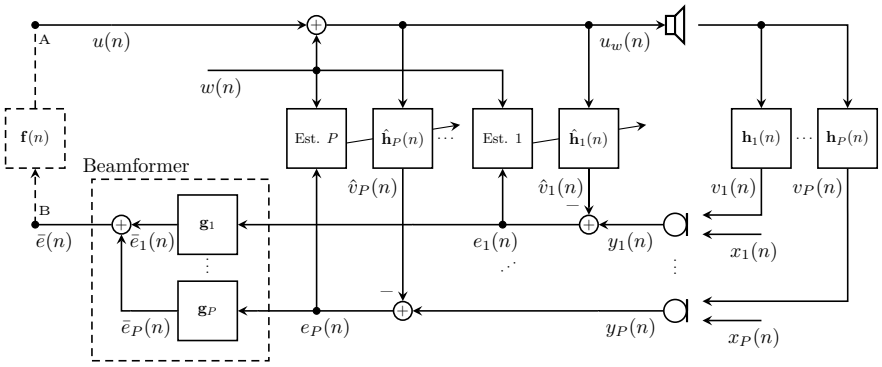


Fig. 1: A multiple-microphone and single-loudspeaker system using a probe noise signal $w(n)$ for unbiased estimation of $\mathbf{h}_i(n)$.

adaptive filters $\hat{\mathbf{h}}_i(n)$ for estimating the true feedback paths $\mathbf{h}_i(n)$, where n is the time index, $i = 1, \ldots, P$, and P is the number of microphones. The probe noise signal $w(n)$ is added to the original loudspeaker signal $u(n)$ to facilitate unbiased estimation of the true feedback paths $\mathbf{h}_i(n)$, and the estimation of adaptive filters $\hat{\mathbf{h}}_i(n)$ is based on the signals $w(n)$ and $e_i(n)$. More details on Fig. 1 are given in Sec. 2.

Different studies have reported that the probe noise approach is feasible in practice, only if the power of the probe noise $w(n)$ is high enough compared to the original loudspeaker signal $u(n)$, see e.g. [3–6]. Unfortunately, the required probe noise power is generally so high that the probe noise becomes clearly audible and annoying for the users of the audio system. When the probe noise is adjusted to be inaudible, the convergence rate of the adaptive algorithm is highly decreased while maintaining the steady-state error, which limits the practical use of the probe noise approach in an AFC system.

In this work, we derive theoretically the behavior of an AFC system using probe noise in an MMSL system as shown in Fig. 1. In [7], we introduced a frequency domain evaluation criterion referred to as the power transfer function (PTF), which describes AFC system behavior for a general MMSL system without probe noise. The PTF criterion was a further development of the work on frequency domain tracking characteristics of adaptive algorithms in [8]. In this paper, we show theoretically that existing probe noise approaches give a significantly decreased convergence rate compared to systems without probe noise, assuming that the probe noise must be inaudible and the steady-state error in AFC systems unchanged. The derived expressions provide a theoretical description of the existing probe noise approach, which we believe is important for further improvements.

2 System Description

Fig. 1 shows a general MMSL system. The ith unknown feedback path $\mathbf{h}_i(n)$ is represented by a finite impulse response (FIR) of order $L-1$, $\mathbf{h}_i(n) = [h_i(0,n), \ldots, h_i(L-1,n)]^T$ and its frequency response is given by the discrete Fourier transform (DFT), $H_i(\omega,n) = \sum_{k=0}^{L-1} h_i(k,n)e^{-j\omega k}$.

For simplicity, we model time variations of the feedback paths $\mathbf{h}_i(n)$ using the random walk model $H_i(\omega,n) = H_i(\omega,n-1) + \check{H}_i(\omega,n)$. The complex scalar $\check{H}_i(\omega,n) \in \mathbb{C}$ is a sample from a zero-mean Gaussian[1] stochastic process with variance $S_{\check{h}_{ii}}(\omega) = E[\check{H}_i(\omega,n)\check{H}_i^*(\omega,n)]$, where $*$ denotes complex conjugation. For simplicity, we assume $S_{\check{h}_{ij}}(\omega) = 0 \;\forall\; i \neq j$. Alternatively, different feedback path variation models [9] can be used, but the methodology of analysis would remain the same.

In the time domain, the feedback path variation vector is

$$\check{\mathbf{h}}_i(n) = \mathbf{h}_i(n) - \mathbf{h}_i(n-1). \tag{1}$$

The estimate of the ith feedback path is given by $\hat{\mathbf{h}}_i(n) = [\hat{h}_i(0,n), \ldots, \hat{h}_i(L-1,n)]^T$ and the corresponding estimation error vector $\tilde{\mathbf{h}}_i(n) = [\tilde{h}_i(0,n), \ldots, \tilde{h}_i(L-1,n)]^T$ is

$$\tilde{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n) - \mathbf{h}_i(n), \tag{2}$$

with frequency response of $\tilde{H}_i(\omega,n) = \sum_{k=0}^{L-1} \tilde{h}_i(k,n)e^{-j\omega k}$.

In the analysis, we consider the original loudspeaker signal $u(n)$ and probe noise signal $w(n)$ as deterministic signals because they are measurable and thereby known. The original loudspeaker signal vector is defined as $\mathbf{u}(n) = [u(n), \ldots, u(n-L+1)]^T$, and the probe noise signal vector is defined as $\mathbf{w}(n) = [w(n), \ldots, w(n-L+1)]^T$. The probe noise signal is generated as $w(n) = \sum_{k=0}^{L_w-1} h_w(k)\epsilon(n-k)$, where $\epsilon(n)$ is a zero-mean Gaussian stochastic sequence with unit variance, and $\mathbf{h}_w = [h_w(0), \ldots, h_w(L_w-1)]^T$ is a known spectral shaping filter. The resulting loudspeaker signal is $u_w(n) = u(n)+w(n)$ with a signal vector $\mathbf{u}_w(n) = [u_w(n), \ldots, u_w(n-L+1)]^T$, and

$$\mathbf{u}_w(n) = \mathbf{u}(n) + \mathbf{w}(n). \tag{3}$$

The ith microphone signal $y_i(n)$ is expressed by

$$y_i(n) = x_i(n) + \mathbf{u}_w^T(n)\mathbf{h}_i(n-1), \tag{4}$$

and the ith error signal $e_i(n)$ is

$$e_i(n) = y_i(n) - \mathbf{u}_w^T(n)\hat{\mathbf{h}}_i(n-1). \tag{5}$$

[1] For convenience, we use the Gaussian distribution for different signals. However, we do not rely on the Gaussian distribution in our analysis.

In this work, the beamformer filters $\mathbf{g}_i$ are considered fixed because they are in general slowly varying compared to AFC systems; they are represented by FIR filters $\mathbf{g}_i = [g_i(0), \ldots, g_i(N-1)]^T$ with frequency responses of $G_i(\omega) = \sum_{k=0}^{N-1} g_i(k) e^{-j\omega k}$. The output $\bar{e}_i(n)$ of the ith beamformer filter is $\bar{e}_i(n) = \sum_{k=0}^{N-1} g_i(k) e_i(n-k)$, and the beamformer output signal is given by $\bar{e}(n) = \sum_{i=1}^{P} \bar{e}_i(n)$.

There is a slightly different but regularly used probe noise approach which uses $u_w(n)$ to estimate the adaptive filters $\hat{\mathbf{h}}_i(n)$, see e.g. [2]. However, analysis of this approach is outside the scope of this work.

3 Review of Power Transfer Function

In our analysis, we use the PTF introduced in [7] to describe the system behavior. The PTF $\xi(\omega, n)$ is defined as the expected magnitude-squared transfer function from point A to B in Fig. 1,

$$\begin{aligned}\xi(\omega, n) &= E\left[\left|\sum_{i=1}^{P} G_i(\omega)\tilde{H}_i(\omega, n)\right|^2\right] \\ &= \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)\xi_{ij}(\omega, n),\end{aligned} \tag{6}$$

where $\xi_{ij}(\omega, n) = E[\tilde{H}_i(\omega, n)\tilde{H}_j^*(\omega, n)]$.

The PTF $\xi(\omega, n)$ is the unknown part of the expected magnitude-squared open-loop transfer function given by $E[|\mathrm{OLTF}(\omega, n)|^2] = |F(\omega, n)|^2\xi(\omega, n)$, where $F(\omega, n)$ is the frequency response of the forward path impulse response $\mathbf{f}(n)$. If $|\mathrm{OLTF}(\omega, n)| < 1$, system stability is guaranteed [10].

In general, the frequency responses $H_i(\omega, n)$ of the true feedback paths $\mathbf{h}_i(n)$, and thereby $\tilde{H}_i(\omega, n) = \hat{H}_i(\omega, n) - H_i(\omega, n)$, are unknown and considered as stochastic. Hence, in practice, we can not compute the PTF $\xi(\omega, n)$ directly. Instead, we let $\hat{\xi}_{ij}(\omega, n) \approx E[\tilde{H}_i(\omega, n)\tilde{H}_j^*(\omega, n)]$ and derive a simple and easier-to-interpret approximation of the PTF,

$$\hat{\xi}(\omega, n) = \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)\hat{\xi}_{ij}(\omega, n). \tag{7}$$

4 System Analysis

In this section, we derive the PTF approximation $\hat{\xi}_{ij}(\omega, n)$ based on an example adaptive algorithm. It is possible to conduct derivations for other adaptive algorithms using the

same methodology.

With traditional adaptive algorithms such as the least mean square (LMS) algorithm [11], nonzero signal correlation $r_{xu}(k) = E[x(n)u(n-k)]$ between the incoming signal $x(n)$ and the original loudspeaker signal $u(n)$ leads to biased feedback path estimates $\hat{\mathbf{h}}_i(n)$, see e.g. [1]. The bias can be avoided by basing the estimation of $\hat{\mathbf{h}}_i(n)$ on a probe noise signal $w(n)$, which is uncorrelated with both $x(n)$ and $u(n)$; in this case, the actual signal correlation $r_{xu}(k)$ is no more important for the estimation [2]. With the probe noise approach in Fig. 1, an expression for the error signal $e_i(n)$ can be found by inserting Eqs. (3)-(4) in Eq. (5),

$$e_i(n) = x_i(n) - \mathbf{w}^T(n)\tilde{\mathbf{h}}_i(n-1) - \mathbf{u}^T(n)\tilde{\mathbf{h}}_i(n-1). \tag{8}$$

The update of $\hat{\mathbf{h}}_i(n)$ with a step size $\mu(n)$ is performed as

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{w}(n)e_i(n). \tag{9}$$

The PTF for the MMSL system with probe noise can be derived based on the update rule in Eq. (9) as follows. Using Eqs. (9), (8) and (1), the estimation error vector defined in Eq. (2) can be expressed by

$$\begin{aligned}\tilde{\mathbf{h}}_i(n) = &\left(\mathbf{I} - \mu(n)\mathbf{w}(n)\mathbf{w}^T(n) - \mu(n)\mathbf{w}(n)\mathbf{u}^T(n)\right)\tilde{\mathbf{h}}_i(n-1)\\ &+ \mu(n)\mathbf{w}(n)x_i(n) - \check{\mathbf{h}}_i(n),\end{aligned} \tag{10}$$

where $\mathbf{I}$ is the identity matrix. Furthermore, it can be shown that an approximation of the estimation error correlation matrix, $\hat{\mathbf{H}}_{ij}(n) \approx E\left[\tilde{\mathbf{h}}_i(n)\tilde{\mathbf{h}}_j^T(n)\right]$, is given by

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =&\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\mathbf{R}_w(0)\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{R}_w(0) + \check{\mathbf{H}}_{ij} + \mu^2(n)\sum_{k=-k_0}^{k_0}\mathbf{R}_w(k)r_{x_ij}(k).\end{aligned} \tag{11}$$

Eq. (11) is derived under the assumption of a sufficiently small step size, i.e. $\mu(n) \to 0$, and we used the direct-averaging method [12] to replace $\mathbf{w}(n)\mathbf{w}^T(n-k)$ and $\mathbf{w}(n)\mathbf{u}^T(n-k)$ by their sample average $\mathbf{R}_w(k) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathbf{w}(n)\mathbf{w}^T(n-k)$ and $\mathbf{R}_{wu}(k) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathbf{w}(n)\mathbf{u}^T(n-k) = \mathbf{0}$, respectively. Furthermore, $\check{\mathbf{H}}_{ij} = E[\check{\mathbf{h}}_i(n)\check{\mathbf{h}}_j^T(n)]$ is the correlation matrix of the ith and jth feedback path variations, and we assumed that the correlation function $r_{x_ij}(k) = E[x_i(n)x_j(n-k)]$ of the incoming signals fulfills $r_{x_ij}(k) = 0 \;\forall\; |k| > k_0 \in \mathbb{N}$.

To find an expression for $\hat{\xi}_{ij}(\omega, n)$, we let $\mathbf{F} \in \mathbb{C}^{L\times L}$ be a DFT matrix. It is well-known that $\mathbf{F}$ diagonalizes a Toeplitz matrix for sufficiently large L [13]. The matrix

$\hat{\mathbf{H}}_{ij}(n)$ is asymptotically a Toeplitz matrix; therefore, $\mathbf{F}\hat{\mathbf{H}}_{ij}(n)\mathbf{F}^H$ approaches a diagonal matrix asymptotically as $L \to \infty$. It can be shown that the diagonal elements are

$$\hat{\xi}_{ij}(\omega, n) = (1 - 2\mu(n)S_w(\omega))\,\hat{\xi}_{ij}(\omega, n-1) + L\mu^2(n)S_w(\omega)S_{x_{ij}}(\omega, n) + S_{\check{h}_{ij}}(\omega), \quad (12)$$

where $S_w(\omega)$ denotes the power spectrum density (PSD) of the probe noise signal $w(n)$, and $S_{x_{ij}}(\omega)$ denotes the cross(auto) PSDs of the incoming signals $x_i(n)$ and $x_j(n)$.

Finally, inserting Eq. (12) in Eq. (7), the PTF approximation $\hat{\xi}(\omega, n)$ can be expressed by

$$\begin{aligned}\hat{\xi}(\omega, n) = &(1 - 2\mu(n)S_w(\omega))\,\hat{\xi}(\omega, n-1) \\ &+ L\mu^2(n)S_w(\omega)\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega) + \sum_{i=1}^{P}|G_i(\omega)|^2 S_{\check{h}_{ii}}(\omega). \quad (13)\end{aligned}$$

5 Discussion

We can now determine the convergence rate, algorithm stability bound for the step size, and steady-state behavior of the probe noise based AFC algorithm in the MMSL system.

The PTF expression in Eq. (13) can be viewed as a first-order difference equation in $\hat{\xi}(\omega, n)$ described by the transfer function $H(z) = \frac{\beta}{1-\alpha z^{-1}}$. The coefficient α determines the pole location in $H(z)$ and thus the decay rate of $\hat{\xi}(\omega, n)$, it is given by

$$\alpha = 1 - 2\mu(n)S_w(\omega). \quad (14)$$

The convergence rate in dB/iteration is given by $10\log_{10}(|\alpha|)$. Algorithm stability is guaranteed if $|\alpha| < 1$, so the stability bound for $\mu(n)$ is

$$0 < \mu(n) < \frac{1}{\max_\omega S_w(\omega)}. \quad (15)$$

The steady-state behavior, $\hat{\xi}(\omega, \infty) = \lim_{n\to\infty}\hat{\xi}(\omega, n)$, is

$$\hat{\xi}(\omega, \infty) = \underbrace{\lim_{n\to\infty} L\frac{\mu(n)}{2}\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)}_{\text{Steady-State Error}} + \underbrace{\lim_{n\to\infty}\frac{\sum_{i=1}^{P}|G_i(\omega)|^2 S_{\check{h}_{ii}}(\omega)}{2\mu(n)S_w(\omega)}}_{\text{Tracking Error}}. \quad (16)$$

The tracking error is caused by the time-varying feedback paths $\mathbf{h}_i(n)$, and the steady-state error is due to the incoming signals $x_i(n)$ and $x_j(n)$. Eqs. (14) and (16) provide a theoretical description of the system behavior for a probe noise based AFC system.

They describe how the system parameters $\mu(n)$, L and $G_p(\omega)$, the signal properties $S_w(\omega)$ and $S_{x_{ij}}(\omega)$, and the feedback path variations $S_{\check{h}_{ii}}(\omega)$ affect the convergence rate and steady-state behavior. It is interesting to note that system behavior could be considered to be independent of the original loudspeaker signal $u(n)$.

Eqs. (14)-(16) are similar in structure to the corresponding expressions for the same MMSL system without probe noise injection [7]. The important difference is that in Eqs. (14)-(16), the PSD $S_u(\omega)$ of the original loudspeaker signal $u(n)$ is now replaced by the PSD $S_w(\omega)$ of the probe noise signal $w(n)$. It means that using the same step size $\mu(n)$ in both systems would lead to an identical steady-state error, but a *decrease* in convergence rate and an *increase* in tracking error both by a factor of $S_u(\omega)/S_w(\omega)$ in the probe noise system. In many practical systems, one wishes the probe noise $w(n)$ to be inaudible in the presence of the original loudspeaker signal $u(n)$, such that the resulting loudspeaker signal $u_w(n)$ is perceived as the original loudspeaker signal $u(n)$. This may be achieved by exploiting the masking effects of the human auditory system [14], to generate a probe noise signal $w(n)$ masked by $u(n)$. In this case, a time-varying $S_w(\omega, n)$ is chosen as a function of the short-time PSD $S_u(\omega, n)$, and is generally as much as 15-25 dB below the level of $S_u(\omega, n)$, according to e.g. perceptual audio coding techniques [15]. Hence, we would expect a reduction in convergence rate and an increase in tracking error by a factor of more than 30.

6 Simulation Verification

The goal of the simulations is to verify the derived theoretical expressions in Eqs. (14) and (16), and to compare these with the behavior of an identical MMSL system without probe noise [7]. The simulations are performed in an MMSL system with $P = 3$ microphones.

In the simulations, the beamformer filters $\mathbf{g}_i$ and the true feedback paths $\mathbf{h}_i(n)$ are modeled by first-order FIR filters[2] $\mathbf{g}_1 = [1, 0.36]^T$, $\mathbf{g}_2 = [1, -0.32]^T$, $\mathbf{g}_3 = [1, 0.23]^T$, $\mathbf{h}_1(n) = [1, 0.14]^T$, $\mathbf{h}_2(n) = [1, -0.4]^T$ and $\mathbf{h}_3(n) = [1, 0.21]^T$. Thus, it is possible to compute the true PTF $\xi(\omega, n)$ in Eq. (6). The duration of each simulation run is $2 \cdot 10^4$ iterations, and 100 simulation runs are performed to obtain an averaged $\xi(\omega, n)$.

In each simulation run, new realizations of standard Gaussian stochastic sequences are drawn; they are then filtered by various fixed first-order FIR shaping filters $\mathbf{h}_u(n) = [1, -0.3]^T$, $\mathbf{h}_w(n) = \frac{1}{2} \cdot [1, -0.3]^T$, $\mathbf{h}_{x_1}(n) = [1, 0.3]^T$, $\mathbf{h}_{x_2}(n) = [1, -0.2]^T$ and $\mathbf{h}_{x_3}(n) = [1, 0.5]^T$ to generate the original loudspeaker signal $u(n)$, probe noise signal $w(n)$ and the incoming signals $x_i(n)$, respectively. In this way, $u(n)$ and $x_i(n)$ are uncorrelated, although this is not necessary to verify the derived PTF expressions for the AFC system

[2] Using first-order FIR filters to model the true feedback paths $\mathbf{h}_i(n)$ is unrealistic in practice. However, the derivations show that the PTF expressions and thereby the AFC performance are actually independent of $\mathbf{h}_i(n)$ which we have verified by simulations. First-order filters are chosen for reproducibility in this experiment.

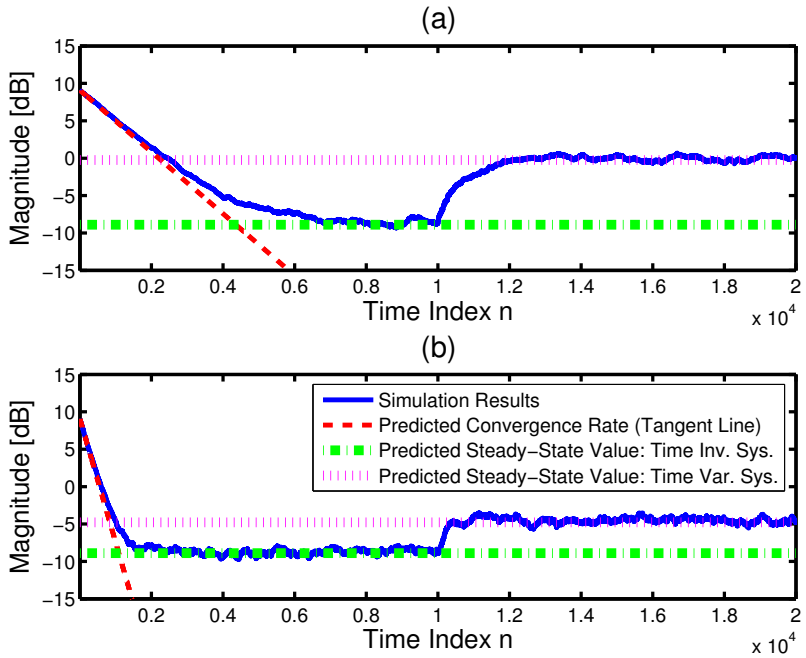


Fig. 2: Verification results at frequency $\omega = 2\pi l/L$, where $L = 32$ and $l = 7$. (a) An MMSL system using a probe noise signal. (b) The same MMSL system without probe noise.

using probe noise, which is immune to these correlations. However, using uncorrelated $u(n)$ and $x_i(n)$ signals, an unbiased estimation and PTF prediction in an AFC system *without* probe noise, carried out in the next simulation experiment, can be obtained; it makes a direct comparison between the convergence rate and steady-state error in both systems straightforward. Furthermore, for ease of demonstration, the probe noise signal level is simply chosen as half of the original loudspeaker signal.

The feedback path estimates are initialized to zero, i.e. $\hat{\mathbf{h}}_i(0) = \mathbf{0}_{L\times 1}$, and $L = 32$. The true feedback paths are fixed during the first 10^4 iterations of the simulations, whereas random walk variations with variances $\sigma^2_{h_1} = 0.0184$, $\sigma^2_{h_2} = 0.0132$ and $\sigma^2_{h_3} = 0.0124$ are added in the remaining period. A fixed step size $\mu = 2^{-9}$ is used. The simulated and predicted results, at a representative example frequency $\omega = 2\pi l/L$, where $l = 7$, are shown in Fig. 2(a). The simulation results agree with the predicted convergence rate and steady-state values in Eqs. (14) and (16).

A similar simulation experiment is carried out with the same MMSL system but without probe noise injection [7]. The estimation of $\hat{\mathbf{h}}_i(n)$ is based on the original loudspeaker signal $u(n)$ using a standard LMS algorithm $\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{u}(n)e_i(n)$,

where $e_i(n) = x_i(n) - \mathbf{u}^T(n)\tilde{\mathbf{h}}_i(n-1)$. As in the previous experiment, the step size is $\mu = 2^{-9}$. The prediction values are computed by replacing $S_w(\omega)$ with $S_u(\omega)$ in Eqs. (14) and (16). The simulation results are given in Fig. 2(b). The results confirm that when the steady-state error is equal in both systems, the convergence rate of the system with probe noise injection is decreased by the ratio of $S_u(\omega)/S_w(\omega)$ compared to the identical system without probe noise. In this example experiment, since $S_u(\omega) = 4S_w(\omega)$, the convergence rate is decreased by a factor of four. For the same reason, the tracking error as given in the second part of Eq. (16) becomes four times larger.

7 Conclusions

This work dealt with acoustic feedback cancellation for multiple-microphone and single-loudspeaker audio systems. We derived expressions for the cancellation behavior in terms of the convergence rate and steady-state behavior for the entire system, when using a probe noise signal in the adaptive feedback cancellation algorithm. The derived expressions describe the impact of different system parameters and signal properties on the convergence rate and steady-state behavior, and thereby provide a theoretical description of the probe noise approach, which we believe is important for further improvements. In particular, the analysis explain the decreased convergence rate in the cancellation system using probe noise compared to an identical system without it. Specifically, the price to pay for unbiased feedback cancellation using probe noise is a decrease in convergence rate determined by the ratio of $S_u(\omega)/S_w(\omega)$.

References

[1] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Elsevier J. Franklin Inst.*, vol. 343, no. 6, pp. 545–573, Sep. 2006.

[2] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

[3] A. Gilloire and V. Turbin, "Using auditory properties to improve the behaviour of stereophonic acoustic echo cancellers," in *Proc. 1998 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, May 1998, pp. 3681–3684.

[4] J. Hellgren and U. Forssell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 906–913, Nov. 2001.

[5] N. A. Shusina and B. Rafaely, "Unbiased adaptive feedback cancellation in hearing aids by closed-loop identification," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 658–665, Mar. 2006.

[6] T. van Waterschoot and M. Moonen, "Assessing the acoustic feedback control performance of adaptive feedback cancellation in sound reinforcement systems," in *Proc. 17th European Signal Process. Conf.*, Aug. 2009, pp. 1997–2001.

[7] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of acoustic feedback/echo cancellation in multiple-microphone and single-loudspeaker systems using a power transfer function method," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5774–5788, Dec. 2011.

[8] S. Gunnarsson and L. Ljung, "Frequency domain tracking characteristics of adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 1072–1089, Jul. 1989.

[9] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification – a survey," *Elsevier Automatica*, vol. 26, no. 1, pp. 7–21, Jan. 1990.

[10] H. Nyquist, "Regeneration theory," *Bell System Tech. J.*, vol. 11, pp. 126–147, 1932.

[11] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, US: Prentice Hall, Sep. 2001.

[12] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory.* Cambridge, MA, US: MIT Press, 1984.

[13] R. M. Gray, *Toeplitz and Circulant Matrices: A Review.* Hanover, MA, US: Now Publishers Inc., Jan. 2006.

[14] B. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. Bingley, UK: Emerald Group Publishing Limited, Apr. 2003.

[15] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. IEEE*, vol. 88, no. 4, pp. 451–513, Apr. 2000.

Paper D

Novel Acoustic Feedback Cancellation Approaches in Hearing Aid Applications Using Probe Noise and Probe Noise Enhancement

Meng Guo, Søren Holdt Jensen, and Jesper Jensen

Published in
IEEE Trans. Audio, Speech, Lang. Process., vol. 20, no. 9, pp. 2549–2563, Nov. 2012.

The layout has been revised.

Novel Acoustic Feedback Cancellation Approaches in Hearing Aid Applications Using Probe Noise and Probe Noise Enhancement

Meng Guo, Søren Holdt Jensen, and Jesper Jensen

Abstract

Adaptive filters are widely used in acoustic feedback cancellation systems and have evolved to be state-of-the-art. One major challenge remaining is that the adaptive filter estimates are biased due to the nonzero correlation between the loudspeaker signals and the signals entering the audio system. In many cases, this bias problem causes the cancellation system to fail. The traditional probe noise approach, where a noise signal is added to the loudspeaker signal can, in theory, prevent the bias. However, in practice, the probe noise level must often be so high that the noise is clearly audible and annoying; this makes the traditional probe noise approach less useful in practical applications. In this work, we explain theoretically the decreased convergence rate when using low-level probe noise in the traditional approach, before we propose and study analytically two new probe noise approaches utilizing a combination of *specifically designed probe noise signals and probe noise enhancement. Despite using low-level and inaudible probe noise signals, both approaches significantly improve the convergence behavior of the cancellation system compared to the traditional probe noise approach. This makes the proposed approaches much more attractive in practical applications. We demonstrate this through a simulation experiment with audio signals in a hearing aid acoustic feedback cancellation system, where the convergence rate is improved by as much as a factor of 10.*

1 Introduction

Acoustic feedback problems may occur in audio systems when the microphone picks up part of the acoustic output signal from the loudspeaker. This problem often causes significant performance degradations in applications such as public address systems and hearing aids. In the worst-case, the audio system becomes unstable and howling occurs. Many solutions have been proposed for reducing the effect of acoustic feedback, see e.g. [1, 2] and the references therein. A widely used and probably the best solution to date is to use adaptive filters in a system identification configuration [3].

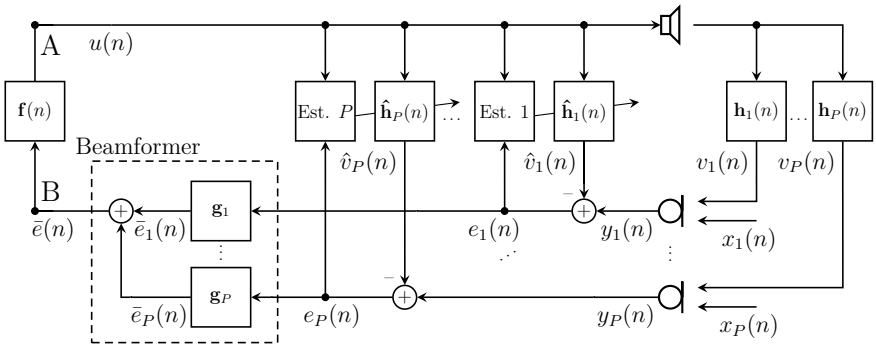


Fig. 1: A traditional acoustic feedback cancellation approach in a multiple-microphone and single-loudspeaker system.

Fig. 1 shows a general acoustic feedback cancellation (AFC) approach using adaptive filters in a multiple-microphone and single-loudspeaker (MMSL) audio system, where AFC is carried out using the adaptive filters $\hat{\mathbf{h}}_i(n)$ to compensate for the true acoustic feedback paths $\mathbf{h}_i(n)$, where n is the discrete-time index, $i = 1, \ldots, P$, and P is the number of microphones. The estimation of the true feedback paths $\mathbf{h}_i(n)$ by means of adaptive filters $\hat{\mathbf{h}}_i(n)$ is based on the loudspeaker signal $u(n)$ and the error signals $e_i(n)$ and can be performed using e.g. least mean square (LMS), normalized least mean square (NLMS), and recursive least squares (RLS) algorithms [3, 4]. The incoming signals to the microphones of the MMSL system are denoted by $x_i(n)$. Often, multiple-microphone audio systems are equipped with a beamforming algorithm to perform spatial filtering of the incoming signals. The beamformer filters $\mathbf{g}_i$ process the error signals $e_i(n)$ to form a spatially filtered beamformer output signal $\bar{e}(n)$, which is further modified by the forward path $\mathbf{f}(n)$ to produce the loudspeaker signal $u(n)$. More details on Fig. 1 are given in Sec. 2.1.

The adaptive filter $\hat{\mathbf{h}}_i(n)$ approximates the acoustic feedback path $\mathbf{h}_i(n)$. Although AFC using adaptive filters is one of the most applied methods to compensate for the feedback problem, one of the major problems remaining is that the estimates $\hat{\mathbf{h}}_i(n)$ become biased, i.e. $E[\hat{\mathbf{h}}_i(n)] \neq \mathbf{h}_i(n)$, where $E[\cdot]$ denotes the statistical expectation operator, whenever the loudspeaker signal $u(n)$ and the incoming signals $x_i(n)$ are correlated [5]. This is generally unavoidable in closed-loop systems as e.g. shown in Fig. 1, because the loudspeaker signal $u(n)$ is a processed and delayed version of the incoming signals $x_i(n)$. The biased estimation of $\mathbf{h}_i(n)$ may lead to a poor feedback cancellation and in the worst-case causes the cancellation system to fail.

Different techniques have been proposed to prevent or reduce the biased estimation problem. Nonlinear processing methods [1] add an ideally inaudible, nonlinearly distorted version of loudspeaker signal $u(n)$ to $u(n)$ to decorrelate it from the incoming

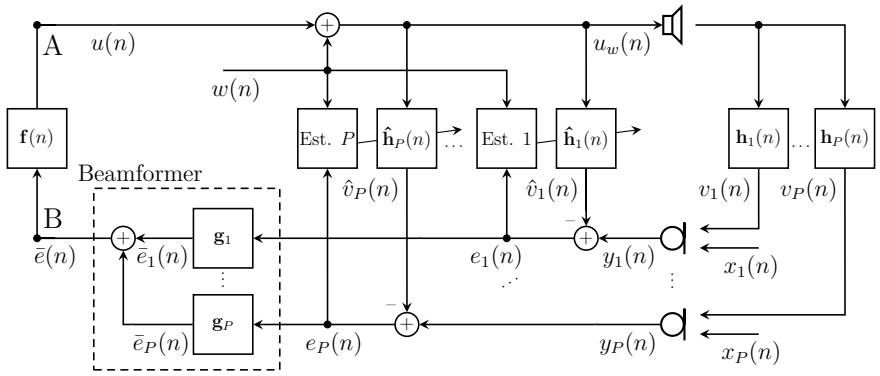


Fig. 2: A traditional probe noise based acoustic feedback cancellation approach in a multiple-microphone and single-loudspeaker system.

signals $x_i(n)$. Typically, a half-wave rectifier is used to introduce the distortion. The frequency transposition methods [6, 7] introduce a modification in the forward path $\mathbf{f}(n)$, by e.g. shifting the frequency components from the incoming signals $x_i(n)$ to other frequencies. Thus, it decorrelates the loudspeaker signal $u(n)$ and the incoming signals $x_i(n)$ and is thereby capable of reducing the bias problem. The prediction error method [8, 9] utilizes prefilters applied to the signals entering the adaptive filter estimation; the prefilters are used to approximately whiten the incoming signal components in these signals and thereby compensate for the biased estimation.

In this work, we focus on the probe noise approach. Fig. 2 shows a traditional probe noise approach in an MMSL system. The probe noise signal $w(n)$ is added to the original loudspeaker signal $u(n)$ to facilitate unbiased estimation of the true feedback paths $\mathbf{h}_i(n)$. In contrast to the traditional AFC approach shown in Fig. 1, the adaptive filters $\hat{\mathbf{h}}_i(n)$ are estimated based on the probe noise signal $w(n)$ and the error signals $e_i(n)$, and unbiased estimation is guaranteed since $w(n)$ is uncorrelated with $u(n)$ and $x_i(n)$ by construction. More details on Fig. 2 are given in Sec. 2.2.

Adaptive filter estimation based on probe noise can be carried out in different ways. In [10], the adaptive filter estimates are only updated when the system is detected to be close to instability; in this case, the original loudspeaker signal $u(n)$ is muted, and only the probe noise signal $w(n)$ is presented as the loudspeaker signal to perform the estimation. In [11], an attempt is made to reduce the audible artifacts introduced by a high-level probe noise signal; specifically, probe noise insertion and adaptive filter estimation is only performed during quiet intervals. In both cases, a non-continuous adaptation is carried out, and the cancellation performance is highly dependent on the decisions made by the stability and quiet-interval detectors, respectively. For input signals with few quiet passages, e.g. musical signals, these systems can not update their feedback path estimate and are therefore sensitive to feedback path changes. In

other probe noise approaches [12, 13], an estimate $\hat{u}(n)$ of the loudspeaker signal $u(n)$ is created by using a probe noise signal in an open-loop system identification configuration, and the adaptive filter estimation relies on this estimated signal $\hat{u}(n)$ instead of $u(n)$; $\hat{u}(n)$ is ideally uncorrelated with $x_i(n)$, and an unbiased estimation can thereby be obtained. However, the drawback is that a loud and clearly audible probe noise signal is required.

In principle, all these probe noise approaches can prevent the bias problem and improve the cancellation performance. In [14], it was shown that the traditional probe noise approach is capable of providing similar or even better performance than other state-of-the-art AFC approaches, but only if the level of probe noise $w(n)$ is powerful enough compared to the original loudspeaker signal $u(n)$, see also [13, 15]. On the other hand, when the probe noise level is adjusted to be inaudible, the convergence rate of the adaptive algorithm is often highly decreased (while maintaining the steady-state error), which limits the practical use of the probe noise approach in an AFC system. In [16], it was shown theoretically that when using the traditional probe noise approach with inaudible probe noise signals, the convergence rate of the adaptive system is decreased, by as much as a factor 30 in practice. Based on [16], a theoretical framework was proposed in [17] for an improved probe noise approach, which is capable of significantly increasing the convergence rate without compromising the steady-state error at a given probe noise level. In this paper, we present a comprehensive theoretical analysis of the improved approach in [17] and discuss some important practical aspects of its application in real situations. Within the same theoretical framework, we present a further improved probe noise approach, where the convergence rate is increased by up to a factor 2 compared to [17] with only minimal additional calculations.

The improvements by the proposed probe noise approaches are obtained by processing the signals entering the adaptive algorithms, such that the disturbance from the incoming signals $x_i(n)$ is reduced. Additionally, both improved approaches utilize a simple spectral masking model to generate a probe noise signal $w(n)$, which is inaudible in the presence of the original loudspeaker signal $u(n)$. This provides a resulting loudspeaker signal $u_w(n)$ that is perceived essentially identically to the original loudspeaker signal $u(n)$. This probe noise generation method was introduced for AFC applications in [18].

For both proposed approaches, we derive analytical expressions for their system behavior; we compare them to a traditional AFC system without probe noise [19], and a traditional probe noise based AFC system [16]. Furthermore, we demonstrate the improvements in simulation experiments using audio signals and practical parameter settings in a realistic hearing aid AFC system.

In this work, column vectors and matrices are emphasized using lower and upper letters in bold, respectively. Transposition, Hermitian transposition and complex conjugation are denoted by the superscripts T, H and $*$, respectively.

The rest of this paper is organized as follows. In Sec. 2, we introduce different

MMSL systems using the traditional AFC approach, traditional probe noise approach and the two proposed probe noise approaches. In Sec. 3, we derive analytic expressions for the system behavior in terms of convergence rate and steady-state behavior to explain analytically the improvements obtained using the proposed approaches. In Sec. 4, we perform simulation experiments, using audio signals, to compare the proposed probe noise approaches to the traditional probe noise approach and the traditional AFC approach. Finally, we conclude this work in Sec. 5.

2 System Overview

In this section, we introduce MMSL systems using the four different AFC approaches, which are considered in this work: 1) The traditional AFC approach (T-AFC). 2) A traditional probe noise approach (T-PN). 3) The proposed probe noise approach I (PN-I) in [17]. 4) The proposed probe noise approach II (PN-II).

For convenience, we express all signals as discrete-time signals, although in practice the signals entering the microphones and leaving the loudspeaker are continuous-time signals.

2.1 Traditional AFC Approach (T-AFC)

Fig. 1 shows the MMSL system using the T-AFC approach. The ith true acoustic feedback path $\mathbf{h}_i(n) = [h_i(0,n),\ldots,h_i(L-1,n)]^T$ is assumed to be a finite impulse response (FIR) of order $L-1$. The frequency response of $\mathbf{h}_i(n)$ is expressed by the discrete Fourier transform (DFT) $H_i(\omega,n) = \sum_{k=0}^{L-1} h_i(k,n)e^{-j\omega k}$, where ω is the discrete normalized frequency.

There are different ways to model feedback path variations over time, see e.g. [20]. In this work, we use a simple random walk model given by $H_i(\omega,n) = H_i(\omega,n-1) + \check{H}_i(\omega,n)$ for the ith feedback path, where $\check{H}_i(\omega,n) \in \mathbb{C}$ is a sample from an independent zero-mean Gaussian stochastic sequence with cross-covariance $S_{\check{h}_{ij}}(\omega) = E[\check{H}_i(\omega,n)\check{H}_j^*(\omega,n)]$. Thus, in the time domain, the feedback path variation vector is

$$\check{\mathbf{h}}_i(n) = \mathbf{h}_i(n) - \mathbf{h}_i(n-1). \tag{1}$$

The adaptively estimated feedback path $\hat{\mathbf{h}}_i(n)$ of order $L-1$ is expressed by $\hat{\mathbf{h}}_i(n) = [\hat{h}_i(0,n),\ldots,\hat{h}_i(L-1,n)]^T$, and the estimation error vector is

$$\tilde{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n) - \mathbf{h}_i(n), \tag{2}$$

with a frequency response $\tilde{H}_i(\omega,n) = \sum_{k=0}^{L-1} \tilde{h}_i(k,n)e^{-j\omega k}$.

In this work, we denote the lengths of both $\mathbf{h}_i(n)$ and $\hat{\mathbf{h}}_i(n)$ with L. We assume that $\hat{\mathbf{h}}_i(n)$ has a sufficient length L, in principle $L \to \infty$. Thus, the effective length of $\mathbf{h}_i(n)$ could be shorter than L, e.g. in the case when $\mathbf{h}_i(n)$ is zero-padded to the length L.

The signal vector $\mathbf{u}(n)$ for the loudspeaker signal $u(n)$ is defined as $\mathbf{u}(n) = [u(n), \ldots, u(n-L+1)]^T$, whereas the ith microphone signal $y_i(n)$ is modeled as[1]

$$y_i(n) = x_i(n) + \mathbf{u}^T(n-1)\mathbf{h}_i(n-1), \tag{3}$$

and the ith feedback compensated error signal is given by

$$e_i(n) = y_i(n) - \mathbf{u}^T(n-1)\hat{\mathbf{h}}_i(n-1). \tag{4}$$

The adaptive estimation of $\mathbf{h}_i(n)$ can e.g. be performed using the LMS algorithm [3] with the step size $\mu(n)$ and the update rule

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{u}(n-1)e_i(n), \tag{5}$$

although many more options exist, see e.g. [3, 4].

In the MMSL system shown in Fig. 1, spatial filtering is carried out using a simple linear beamformer [22] applied to the error signals $e_i(n)$. In this work, the beamformer filters $\mathbf{g}_i$ are considered fixed because they are often slowly varying compared to AFC systems; they are represented by FIR filters $\mathbf{g}_i$ of order $L_g - 1$, $\mathbf{g}_i = [g_i(0), \ldots, g_i(L_g-1)]^T$ with a frequency response $G_i(\omega) = \sum_{k=0}^{L_g-1} g_i(k)e^{-j\omega k}$. The output signal of the beamformer is therefore $\bar{e}(n) = \sum_{i=1}^{P} \bar{e}_i(n) = \sum_{i=1}^{P}\sum_{k=0}^{L_g-1} g_i(k)e_i(n-k)$.

Although it is possible to reverse the order of the beamformer and the acoustic feedback cancellation system, we only focus on the case where the cancellation is performed prior to the beamformer as given in Fig. 1. This setup requires more computational power due to multiple cancellation systems, but the beamformer would not affect the cancellation process negatively as demonstrated in [23].

The forward path $\mathbf{f}(n)$ represents the process of converting $\bar{e}(n)$ to the loudspeaker signal $u(n)$. Generally, the forward path $\mathbf{f}(n)$ consists of an amplification and a processing delay for closed-loop audio systems. The impulse response of the forward path is denoted by $\mathbf{f}(n) = [f(0,n), \ldots, f(L_f-1,n)]^T$ with a frequency response $F(\omega,n) = \sum_{k=0}^{L_f-1} f(k,n)e^{-j\omega k}$, and the loudspeaker signal is obtained as $u(n) = \sum_{k=0}^{L_f-1} f(k,n)\bar{e}(n-k)$.

[1] At least one delay element is needed in closed-loop systems to avoid an algebraic loop. As in [21], we chose to model this delay in $\mathbf{h}_i$ by using the time index $n-1$ for notational convenience, since it then would appear to have the same time index as its parallel-structured acoustic feedback path estimate $\hat{\mathbf{h}}_i$. This notation of time indexing does not affect the result.

2.2 Traditional Probe Noise Approach (T-PN)

Fig. 2 shows the MMSL system using the T-PN approach. The significant difference compared to the T-AFC system in Fig. 1 is that a probe noise signal $w(n)$ is added to the original loudspeaker signal $u(n)$, and $w(n)$ is used directly for updating $\hat{\mathbf{h}}_i(n)$. The probe noise signal vector is defined as $\mathbf{w}(n) = [w(n), \ldots, w(n-L+1)]^T$. The resulting loudspeaker signal is $u_w(n) = u(n) + w(n)$ with a signal vector $\mathbf{u}_w(n) = [u_w(n), \ldots, u_w(n-L+1)]^T$, where

$$\mathbf{u}_w(n) = \mathbf{u}(n) + \mathbf{w}(n). \tag{6}$$

The ith microphone signal $y_i(n)$ is given by

$$y_i(n) = x_i(n) + \mathbf{u}_w^T(n-1)\mathbf{h}_i(n-1), \tag{7}$$

and the ith error signal $e_i(n)$ is expressed by

$$e_i(n) = y_i(n) - \mathbf{u}_w^T(n-1)\hat{\mathbf{h}}_i(n-1). \tag{8}$$

The goal of the probe noise $w(n)$ is to ensure an unbiased estimation of $\mathbf{h}_i(n)$, because the probe noise signal $w(n)$ is constructed to be uncorrelated with both the incoming signals $x_i(n)$ and the original loudspeaker signal $u(n)$, see e.g. [2, 5] for details. The probe noise is generated, using a known spectral shaping filter $\mathbf{h}_w(n) = [h_w(0,n), \ldots, h_w(L_w-1,n)]^T$, as $w(n) = \sum_{k=0}^{L_w-1} h_w(k,n)\epsilon(n-k)$, where $\epsilon(n)$ is a zero-mean Gaussian stochastic sequence with unit variance. In this work, we generate a probe noise signal $w(n)$ which is ideally inaudible in the presence of $u(n)$ by adaptively updating $\mathbf{h}_w(n)$ based on the spectral properties of $u(n)$; details on this are given in Sec. 4.2.3. Generally speaking, the goal of this is to maximize the power of the probe noise such that it is just not audible.

The unbiased estimation of $\mathbf{h}_i(n)$ is driven by the probe noise signal $w(n)$ and can e.g. be obtained using an update rule similar to Eq. (5), that is

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{w}(n-1)e_i(n). \tag{9}$$

2.3 Proposed Probe Noise Approach I (PN-I)

Fig. 3 shows the PN-I approach presented in [17]. The difference from the T-PN approach in Fig. 2 is the introduction of the so-called enhancement filters $\mathbf{a}_i(n)$ applied to the error signals $e_i(n)$.

Ideally, in the adaptive filter estimation in a system identification configuration, the error signal entering the estimation block of $\hat{\mathbf{h}}_i(n)$ is $-\mathbf{w}^T(n-1)\tilde{\mathbf{h}}_i(n-1)$. In practice, however, the error signal $e_i(n)$ contains also signal components such as $x_i(n)$

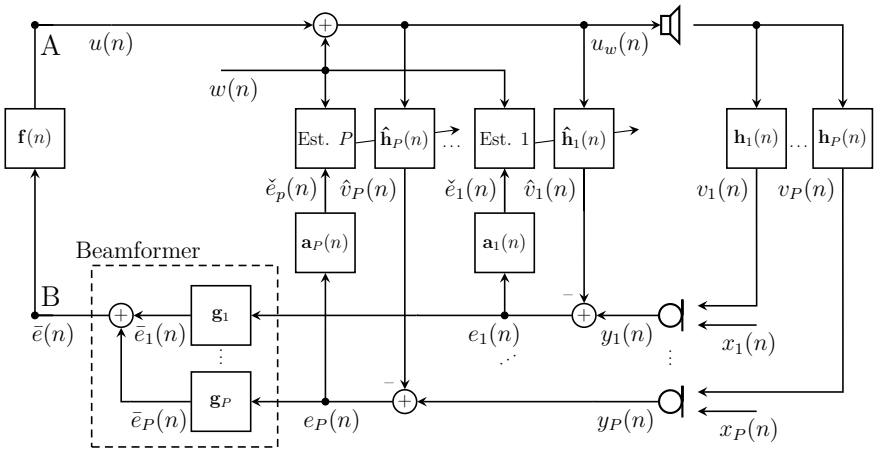


Fig. 3: The improved probe noise acoustic feedback cancellation approach in a multiple-microphone and single-loudspeaker system. The traditional probe noise approach is obtained by setting the filters $\mathbf{a}_i(n) = 1$.

and $-\mathbf{u}^T(n-1)\tilde{\mathbf{h}}_i(n-1)$, which are disturbing the estimation of $\mathbf{h}_i(n)$. The goal of the enhancement filters $\mathbf{a}_i(n)$ is to reduce the disturbing signal power, without changing the probe noise power for the estimation of $\mathbf{h}_i(n)$ at the same time [17]. As we will explain in more details in Sec. 3, the higher the power ratio between the probe noise and the disturbing signals, the faster convergence can be achieved given a fixed steady-state error in the adaptive cancellation system. As $\mathbf{a}_i(n)$ improves the probe noise to disturbing signal ratio, an increased convergence rate can be obtained compared to the T-PN approach without compromising the steady-state behavior in the cancellation system.

The increased probe noise to disturbing signal ratio is obtained by a specific design procedure of the enhancement filter $\mathbf{a}_i(n)$, which is closely related to the probe noise shaping filter length L_w and the feedback path length L. In this work, we assume that the same enhancement filter is applied across microphone channels, i.e. $\mathbf{a}_i(n) = \mathbf{a}(n)$. This is not strictly necessary, but gives a simple result. Furthermore, for audio systems with closely placed microphones such as hearing aids, this is a reasonable simplification. Furthermore, the enhancement filter is presented by an $L_a - 1$ order FIR $\mathbf{a}(n) = [a(0,n), \ldots, a(L_a-1,n)]^T$. Very importantly, the design of $\mathbf{a}(n)$ is constrained such that its frequency response is expressed by

$$A(\omega, n) = 1 + \sum_{k=D}^{L_a-1} a(k,n) e^{-j\omega k}, \tag{10}$$

and the value D is chosen as

$$D \geq L + L_w - 1. \tag{11}$$

Thus, the structure of the enhancement filter is $\mathbf{a}(n) = [1, 0, \ldots, 0, a(D, n), \ldots, a(L_a - 1, n)]^T$, and it is estimated as

$$\hat{\mathbf{a}}(n) = \underset{\mathbf{a}(n)}{\arg\min}\ E\left[\check{e}_i^2(n)\right]. \tag{12}$$

Thus, it is clear that $\hat{\mathbf{a}}(n)$ is simply the minimum mean square error (MMSE) prediction error filter [3]. Furthermore, for a large value of D in Eq. (11), it becomes a *long-term* prediction error filter [24]. We will explain the reason for these choices in Sec. 3.

The filtered error signal $\check{e}_i(n)$ is expressed by

$$\check{e}_i(n) = \sum_{k=0}^{L_a-1} a(k, n) e_i(n - k), \tag{13}$$

and the unbiased feedback path estimation is carried out by basing the estimation of $\mathbf{h}_i(n)$ on the probe noise signal $w(n)$ and filtered error signal $\check{e}_i(n)$, e.g. using the update rule

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{w}(n-1)\check{e}_i(n), \tag{14}$$

which is similar in structure to the LMS update rule used in Eq. (5).

2.4 Proposed Probe Noise Approach II (PN-II)

It is possible to further improve the PN-I approach. This is done by applying copies of the enhancement filter $\mathbf{a}(n)$ to the probe noise signal $w(n)$ to form $\check{w}_i(n)$ as shown in Fig. 4, in which the general terminology $\mathbf{a}_i(n)$ is used, although we assume $\mathbf{a}_i(n) = \mathbf{a}(n)$ for simplicity. As we will show through the theoretical analysis in Sec. 3, these copies of $\mathbf{a}(n)$ on the probe noise signal $w(n)$ lead to even higher probe noise to disturbing signals ratio than the PN-I approach shown in Fig. 3, by increasing the effective probe noise power for the estimation of $\mathbf{h}_i(n)$. Thus, a further increment of the convergence can be obtained in this cancellation system.

Due to the assumption of $\mathbf{a}_i(n) = \mathbf{a}(n)$, the filtered probe noise $\check{w}_i(n) = \check{w}(n)$ is obtained as

$$\check{w}(n) = \sum_{k=0}^{L_a-1} a(k, n) w(n - k). \tag{15}$$

The probe noise signal vector $\check{\mathbf{w}}(n)$ is defined as $\check{\mathbf{w}}(n) = [\check{w}(n), \ldots, \check{w}(n - L + 1)]^T$, and an unbiased feedback path estimation can be carried out e.g. using the update rule

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \mu(n)\check{\mathbf{w}}(n-1)\check{e}_i(n). \tag{16}$$

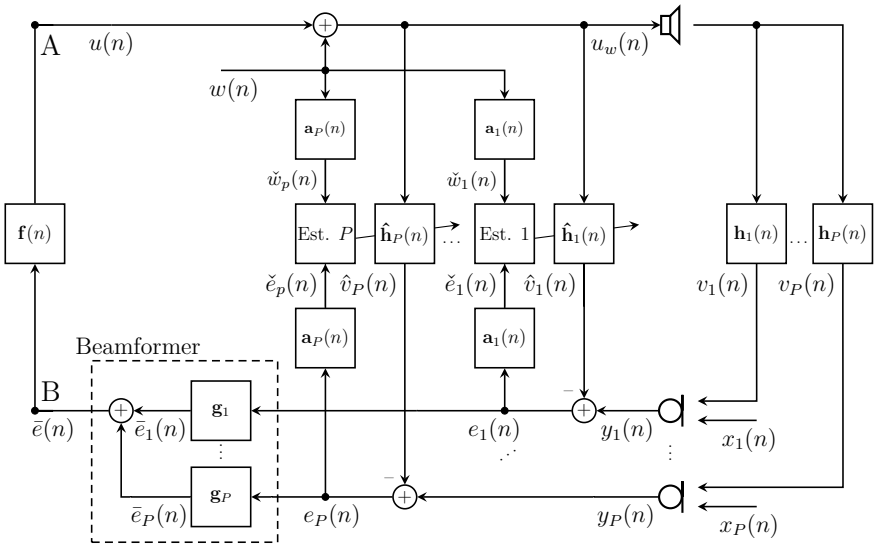


Fig. 4: The further improved probe noise acoustic feedback cancellation approaches in a multiple-microphone and single-loudspeaker system. The difference is that the copies of $\mathbf{a}_i(n)$ are applied on the probe noise signal $w(n)$ to form $\check{w}_i(n)$ which are used in the estimation of $\mathbf{h}_i(n)$.

3 Theoretical Analysis

In this section, we derive analytic expressions to describe system behavior in terms of convergence rate and steady-state error, as a function of time and frequency, based on the example update rules in Eqs. (5), (9), (14) and (16). The derived expressions explain analytically the differences between all four considered AFC approaches. Later in this section, simple simulations are performed to verify the derived expressions.

3.1 Review of Power Transfer Function

The theoretical analysis of the system behavior is based on a recently introduced frequency domain design and evaluation criterion for adaptive systems, the power transfer function (PTF) [19], which describes the expected magnitude-squared transfer function from point A to B in Figs. 1-4. More specifically, the PTF is expressed by

$$\xi(\omega, n) = E\left[\left|\sum_{i=1}^{P} G_i(\omega)\tilde{H}_i(\omega, n)\right|^2\right], \tag{17}$$

and it represents the unknown part of the expected magnitude-squared open-loop transfer function, $E[|\Theta(\omega, n)|^2] = |F(\omega, n)|^2\xi(\omega, n)$. If $|\Theta(\omega, n)| < 1 \ \forall \ \omega$, system stability

is guaranteed [25]. Hence, $\xi(\omega, n)$ provides important information of system behavior. The PTF can generally not be computed directly because the true acoustic feedback paths $\mathbf{h}_i(n)$ and thereby $\tilde{H}(\omega, n) = \hat{H}(\omega, n) - H(\omega, n)$ are unknown. However, as shown in [19], it is possible to obtain an accurate approximation $\hat{\xi}(\omega, n)$ of $\xi(\omega, n)$. This approximation is expressed by a first-order difference equation in $\hat{\xi}(\omega, n)$. Based on this, it is possible to determine the convergence rate and steady-state behavior for the system under concern.

As in [19], we let $\hat{\xi}_{ij}(\omega, n) \approx E[\tilde{H}_i(\omega, n)\tilde{H}_j^*(\omega, n)]$ and via Eq. (17) the PTF approximation can be shown to be

$$\hat{\xi}(\omega, n) = \sum_{i=1}^{P} \sum_{j=1}^{P} G_i(\omega) G_j^*(\omega) \hat{\xi}_{ij}(\omega, n). \tag{18}$$

In the following, we briefly review the PTF approximation $\hat{\xi}(\omega, n)$ for the MMSL system using the T-AFC approach [19], and the T-PN approach [16]. Then, we derive the PTF approximation $\hat{\xi}(\omega, n)$ for the PN-I and PN-II approaches. The derivations and comparisons provide a theoretical explanation of the motivation and improvements by the proposed approaches. For simplicity, the derivation is carried out in an open-loop configuration by omitting $\mathbf{f}(n)$ in the MMSL systems. It can be shown that this has only minor effects on the practical use of the derived results for closed-loop AFC approaches in general [26], and it has no influences on the technical explanations provided in this section. Finally, we assume for simplicity the incoming signals $x_i(n)$ are zero-mean stationary stochastic signals in the analysis.

3.2 Analytic Expressions for System Behavior

3.2.1 Some Definitions

To ease the derivation, we assume $\mathbf{a}(n) = \mathbf{a}$ and divide it further into the parts $\mathbf{a}_1 = [a_1(0), \ldots, a_1(L_a - 1)]^T = [1, 0, \ldots, 0]^T$ and $\mathbf{a}_0 = [a_0(0), \ldots, a_0(L_a - 1)]^T = [0, \ldots, 0, a(D), \ldots, a(L_a - 1)]^T$, such that

$$\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_0. \tag{19}$$

The frequency responses of $\mathbf{a}_1$ and $\mathbf{a}_0$ are $A_1(\omega) = 1$ and $A_0(\omega) = \sum_{k=D}^{L_a-1} a(k) e^{-j\omega k}$, respectively.

Furthermore, we define the Toeplitz-structured filtering matrix $\mathbf{A}_q$, with the dimen-

sion $[L + L_a - 1, L]$, as

$$\mathbf{A}_q = \begin{bmatrix} a_q(0) & 0 & \ddots & 0 \\ \vdots & a_q(0) & \ddots & 0 \\ a_q(L_a-1) & \vdots & \ddots & 0 \\ 0 & a_q(L_a-1) & \ddots & a_q(0) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & a_q(L_a-1) \end{bmatrix}, \tag{20}$$

where $q = 0, 1$, so that we get the matrices $\mathbf{A}_0$ and $\mathbf{A}_1$. Furthermore, we define

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1. \tag{21}$$

Finally, we define the vectors $\bar{\mathbf{w}}(n) = [w(n), \ldots, w(n - L - L_a + 2)]^T$ and $\bar{\mathbf{u}}(n) = [u(n), \ldots, u(n - L - L_a + 2)]^T$.

3.2.2 Traditional AFC Approach (T-AFC)

In [19], the PTF approximation for the MMSL system shown in Fig. 1, using the update rule in Eq. (5), was derived as

$$\begin{aligned} \hat{\xi}(\omega, n) = &(1 - 2\mu(n)S_u(\omega))\,\hat{\xi}(\omega, n-1) \\ &+ L\mu^2(n)S_u(\omega)\sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega) \\ &+ \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega), \end{aligned} \tag{22}$$

where $S_u(\omega)$ denotes the power spectrum density (PSD) of the loudspeaker signal $u(n)$, and $S_{x_{ij}}(\omega)$ denotes the auto/cross PSDs of the incoming signals $x_i(n)$ and $x_j(n)$. Eq. (22) was derived under the assumptions of sufficiently small step size $\mu(n)$ and large model order parameter L, in principle, $\mu(n) \to 0$ and $L \to \infty$. In Eq. (22), the last term is slightly modified compared to the result in [19], since the additional simplifying assumption of $S_{\check{h}_{ij}}(\omega) = 0 \;\forall\; i \neq j$ was applied in [19].

3.2.3 Traditional Probe Noise Approach (T-PN)

In [16], the PTF approximation for the MMSL system using the T-PN approach shown in Fig. 2 and the update rule in Eq. (9) was derived. Under the same assumptions of

$\mu(n)$ and L as for the T-AFC approach, it can be shown that

$$\begin{aligned}\hat{\xi}(\omega,n) =& (1-2\mu(n)S_w(\omega))\,\hat{\xi}(\omega,n-1)\\ &+ L\mu^2(n)S_w(\omega)\sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)\\ &+ \sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega),\end{aligned} \tag{23}$$

where $S_w(\omega)$ denotes the PSD of the probe noise signal $w(n)$. In Eq. (23), the last term is again slightly modified compared to the result in [16] with the additional simplifying assumption of $S_{\check{h}_{ij}}(\omega) = 0\ \forall\ i \neq j$.

3.2.4 Proposed Probe Noise Approach I (PN-I)

In [17], we provided the final PTF expression $\hat{\xi}(\omega,n)$ for the PN-I approach in Fig. 3 without detailed derivations. This section provides more details towards this result. The methodology used for the derivation is similar to the one presented in [26]. However, in contrast to [26], we consider the original loudspeaker signal $u(n)$ as a disturbing signal for the estimation of $\mathbf{h}_i(n)$. Additionally, we need to deal with the effects of enhancement filter $\mathbf{a}(n)$ on different signals and ensure that the estimation of $\mathbf{h}_i(n)$ is still unbiased. In the following, we derive $\hat{\xi}(\omega,n)$ for the PN-I approach with emphasis on this consideration.

Define the matrices $\mathbf{W}(n) = [\mathbf{w}(n),\ldots,\mathbf{w}(n-L_a+1)]^T$ and $\mathbf{U}(n) = [\mathbf{u}(n),\ldots,\mathbf{u}(n-L_a+1)]^T$. Then, using Eqs. (6)-(8) and (13), the example update rule for $\hat{\mathbf{h}}_i(n)$ given by Eq. (14), for the PN-I approach shown in Fig. 3, can be expressed as

$$\begin{aligned}\hat{\mathbf{h}}_i(n) =& \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{w}(n-1)\left(\mathbf{a}^T\mathbf{x}_i(n)\right.\\ &\left.- \mathbf{a}^T\mathbf{U}(n-1)\tilde{\mathbf{h}}_i(n-1) - \mathbf{a}^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)\right).\end{aligned} \tag{24}$$

It can be shown (see Appendix A) when the enhancement filter $\mathbf{a}$ fulfills the important constraint $D \geq L + L_w - 1$ in Eq. (11), then unbiased estimation of $\mathbf{h}_i(n)$ is ensured, i.e. $E[\hat{\mathbf{h}}_i(n)] = \mathbf{h}_i(n)$.

In order to derive the PTF expression $\hat{\xi}(\omega,n)$, we use Eqs. (24), (19) and (1) to express the estimation error vector defined in Eq. (2) as

$$\begin{aligned}\tilde{\mathbf{h}}_i(n) =& \left(\mathbf{I} - \mu(n)\mathbf{w}(n-1)\left(\mathbf{w}^T(n-1) + \mathbf{a}_0^T\mathbf{W}(n-1)\right)\right.\\ &\left.-\mu(n)\mathbf{w}(n-1)\mathbf{a}^T\mathbf{U}(n-1)\right)\tilde{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{w}(n-1)\mathbf{a}^T\mathbf{x}_i(n) - \check{\mathbf{h}}_i(n).\end{aligned} \tag{25}$$

The approximation of the estimation error (auto-) covariance matrix $\mathbf{H}_{ij}(n) = E[\tilde{\mathbf{h}}_i(n)\tilde{\mathbf{h}}_j^T(n)]$ is computed using Eq. (25), under the assumption of sufficiently small

$\mu(n)$, in principle $\mu(n) \to 0$, and by neglecting the second-order terms involving $\mu(n)$ due to the presence of their first-order versions. In addition, we consider $w(n)$ and $u(n)$ as deterministic signals in deriving Eq. (26). As argued in [21] and demonstrated in our simulation experiments, the resulting expression is valid even for the case where $w(n)$ and $u(n)$ are in fact realizations of stochastic processes. The approximation of $\mathbf{H}_{ij}(n)$ becomes

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =&\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\mathbf{w}(n-1)\mathbf{w}^T(n-1)\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\mathbf{w}(n-1)\mathbf{a}^T\mathbf{U}(n-1)\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\mathbf{w}(n-1)\mathbf{a}_0^T(n)\mathbf{W}(n-1)\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{w}(n-1)\mathbf{w}^T(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{U}^T(n-1)\mathbf{a}\mathbf{w}^T(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{W}^T(n-1)\mathbf{a}_0\mathbf{w}^T(n-1) + \check{\mathbf{H}}_{ij}\\ &+ E\left[\mu^2(n)\mathbf{w}(n-1)\mathbf{a}^T\mathbf{x}_i(n)\mathbf{x}_j^T(n)\mathbf{a}\mathbf{w}^T(n-1)\right]\\ &+ E\left[\mu(n)\mathbf{w}(n-1)\mathbf{a}^T\mathbf{x}_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right]\\ &+ E\left[\mu(n)\tilde{\mathbf{h}}_i(n-1)\mathbf{x}_j^T(n)\mathbf{a}\mathbf{w}^T(n-1)\right], \end{aligned} \tag{26}$$

where the correlation matrix of the ith and jth feedback path variations is defined as $\check{\mathbf{H}}_{ij} = E[\check{\mathbf{h}}_i(n)\check{\mathbf{h}}_j^T(n)]$.

Eq. (26) can be simplified. Recall that $w(n)$ is uncorrelated with $u(n)$, thereby $E[\mathbf{w}(n-1)\mathbf{a}^T\mathbf{U}(n-1)] = \mathbf{0}$. Furthermore, since $D \geq L + L_w - 1$ by construction, see Eq. (11), it can be shown that $E[\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)] = \mathbf{0}$ (see Appendix B). Using the direct-averaging method [27] to replace the matrix $\mathbf{w}(n-1)\mathbf{w}^T(n-1)$ with its sample average $\mathbf{R}_w(0) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathbf{w}(n-1)\mathbf{w}^T(n-1)$, the matrix $\mathbf{w}(n-1)\mathbf{a}^T\mathbf{U}(n-1)$ with its sample average $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathbf{w}(n-1)\mathbf{a}^T\mathbf{U}(n-1) = \mathbf{0}$, and the matrix $\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)$ with its sample average $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1) = \mathbf{0}$, the approximation $\hat{\mathbf{H}}_{ij}(n)$ in Eq. (26) can be simplified to

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =&\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\mathbf{R}_w(0)\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{R}_w(0) + \check{\mathbf{H}}_{ij}\\ &+ \mu^2(n)\mathbf{R}_w(0)E\left[\mathbf{a}^T\mathbf{x}_i(n)\mathbf{x}_j^T(n)\mathbf{a}\right]\\ &+ E\left[\mu(n)\mathbf{w}(n-1)\mathbf{a}^T\mathbf{x}_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right]\\ &+ E\left[\mu(n)\tilde{\mathbf{h}}_i(n-1)\mathbf{x}_j^T(n)\mathbf{a}\mathbf{w}^T(n-1)\right]. \end{aligned} \tag{27}$$

We now bring the time domain expression in Eq. (27) to the frequency domain to simplify it further. Recall that, asymptotically as $L \to \infty$, the DFT matrix $\mathbf{F} \in \mathbb{C}^{L\times L}$ diagonalizes any Toeplitz matrix [28]. Using this, we can show that $\hat{\xi}_{ij}(\omega, n)$ are obtained

as the diagonal values of the matrix $\mathbf{F}\hat{\mathbf{H}}_{ij}(n)\mathbf{F}^H$ expressed by

$$\begin{aligned}\hat{\xi}_{ij}(\omega,n) =& (1-2\mu(n)S_w(\omega))\,\hat{\xi}_{ij}(\omega,n-1)\\ &+ L\mu^2(n)S_w(\omega)|A(\omega)|^2 S_{x_{ij}}(\omega) + S_{\check{h}_{ij}}(\omega).\end{aligned} \tag{28}$$

Details on this derivation can be found in [26]. Inserting Eq. (28) in Eq. (18), the PTF approximation $\hat{\xi}(\omega,n)$ is finally obtained as

$$\begin{aligned}\hat{\xi}(\omega,n) =& (1-2\mu(n)S_w(\omega))\,\hat{\xi}(\omega,n-1)\\ &+ L\mu^2(n)S_w(\omega)|A(\omega)|^2 \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)\\ &+ \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega).\end{aligned} \tag{29}$$

3.2.5 Proposed Probe Noise Approach II (PN-II)

In the derivation of the PN-II approach, extra attention must be paid to the copies of the enhancement filter $\mathbf{a}$ filtering the probe noise signal $w(n)$; otherwise, the same procedure is applied as for the PN-I approach.

Using Eqs. (6)-(8), (13) and (15), the estimate of $\mathbf{h}_i(n)$ given by Eq. (16) can be written as

$$\begin{aligned}\hat{\mathbf{h}}_i(n) =& \hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{W}^T(n-1)\mathbf{a}\left(\mathbf{a}^T\mathbf{x}_i(n)\right.\\ &\left.- \mathbf{a}^T\mathbf{U}(n-1)\tilde{\mathbf{h}}_i(n-1) - \mathbf{a}^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)\right).\end{aligned} \tag{30}$$

Similarly to the PN-I approach, it can be shown that an unbiased estimation of $\mathbf{h}_i(n)$ can be obtained as long as the constraint on the enhancement filter $\mathbf{a}$ in Eq. (11) is obeyed.

Using Eqs. (30), (19) and (1), the estimation error vector defined in Eq. (2) can also be expressed by

$$\begin{aligned}\tilde{\mathbf{h}}_i(n) =& \left(\mathbf{I} - \mu(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1)\right.\\ &\left.- \mu(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{U}(n-1)\right)\tilde{\mathbf{h}}_i(n-1)\\ &+ \mu(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{x}_i(n) - \check{\mathbf{h}}_i(n).\end{aligned} \tag{31}$$

The approximation of the estimation error (auto-) covariance matrix $\mathbf{H}_{ij}(n)$ is again computed, under the assumption of sufficiently small $\mu(n)$, and by neglecting the second-

order terms involving $\mu(n)$ in the presence of their first-order versions, as

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =&\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1)\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{U}(n-1)\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{U}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1) + \check{\mathbf{H}}_{ij}\\ &+ E\left[\mu^2(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{x}_i(n)\mathbf{x}_j^T(n)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1)\right]\\ &+ E\left[\mu(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{x}_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right]\\ &+ E\left[\mu(n)\tilde{\mathbf{h}}_i(n-1)\mathbf{x}_j^T(n)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1)\right]. \end{aligned} \tag{32}$$

Considering that $\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1) = \mathbf{A}^T\bar{\mathbf{w}}(n-1)\bar{\mathbf{w}}^T(n-1)\mathbf{A}$, and using the direct-averaging method to further rewrite $\mathbf{A}^T\bar{\mathbf{w}}(n-1)\bar{\mathbf{w}}^T(n-1)\mathbf{A}$ as $\mathbf{A}^T\mathbf{R}_{\bar{w}}(0)\mathbf{A}$, then the matrix $\mathbf{R}_{\bar{w}}(0) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\bar{\mathbf{w}}(n-1)\bar{\mathbf{w}}^T(n-1)$ is identical to $\mathbf{R}_w(0)$ as $L\to\infty$, because $\mathbf{w}(n)$ and $\bar{\mathbf{w}}(n)$ are both signal vectors containing $w(n)$, but with different dimensions. Similarly, we rewrite $\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{U}(n-1)$ as $\mathbf{A}^T\mathbf{R}_{\bar{w}\bar{u}}(0)\mathbf{A} = \mathbf{0}$, where $\mathbf{R}_{\bar{w}\bar{u}}(0) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\bar{\mathbf{w}}(n-1)\bar{\mathbf{u}}^T(n-1)$. The approximation $\hat{\mathbf{H}}_{ij}(n)$ can therefore be simplified to

$$\begin{aligned}\hat{\mathbf{H}}_{ij}(n) =&\hat{\mathbf{H}}_{ij}(n-1) - \mu(n)\mathbf{A}^T\mathbf{R}_w(0)\mathbf{A}\hat{\mathbf{H}}_{ij}(n-1)\\ &- \mu(n)\hat{\mathbf{H}}_{ij}(n-1)\mathbf{A}^T\mathbf{R}_w(0)\mathbf{A} + \check{\mathbf{H}}_{ij}\\ &+ \mu^2(n)\mathbf{A}^T\mathbf{R}_w(0)\mathbf{A}E\left[\mathbf{a}^T\mathbf{x}_i(n)\mathbf{x}_j^T(n)\mathbf{a}\right]\\ &+ E\left[\mu(n)\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{x}_i(n)\tilde{\mathbf{h}}_j^T(n-1)\right]\\ &+ E\left[\mu(n)\tilde{\mathbf{h}}_i(n-1)\mathbf{x}_j^T(n)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1)\right]. \end{aligned} \tag{33}$$

Using similar considerations as in Appendix B, the matrix $\mathbf{A}^T\mathbf{R}_w(0)\mathbf{A}$ can be expressed by

$$\begin{aligned}\mathbf{A}^T\mathbf{R}_w(0)\mathbf{A} =& \left(\mathbf{A}_1 + \mathbf{A}_0\right)^T\mathbf{R}_w(0)\left(\mathbf{A}_1 + \mathbf{A}_0\right)\\ =&\mathbf{A}_1^T\mathbf{R}_w(0)\mathbf{A}_1 + \mathbf{A}_0^T\mathbf{R}_w(0)\mathbf{A}_0\\ =&\mathbf{R}_w(0) + \mathbf{A}_0^T\mathbf{R}_w(0)\mathbf{A}_0. \end{aligned} \tag{34}$$

Inserting Eq. (34) in Eq. (33), and again using the DFT matrix $\mathbf{F}$ to diagonalize $\hat{\mathbf{H}}_{ij}(n)$ in Eq. (33), it can be shown that $\hat{\xi}_{ij}(\omega,n)$ are obtained as the diagonal elements of the resulting matrix $\mathbf{F}\hat{\mathbf{H}}_{ij}(n)\mathbf{F}^H$, as

$$\begin{aligned}\hat{\xi}_{ij}(\omega,n) =& \left(1 - 2\mu(n)(1+|A_0(\omega)|^2)S_w(\omega)\right)\hat{\xi}_{ij}(\omega,n-1)\\ &+ L\mu^2(n)(1+|A_0(\omega)|^2)S_w(\omega)|A(\omega)|^2S_{x_{ij}}(\omega) + S_{\check{h}_{ij}}(\omega). \end{aligned} \tag{35}$$

Table 1: System behavior in terms of convergence rate (**CR**), steady-state error (**SSE**) and tracking error (**TE**) at frequency ω, for the traditional AFC approach (**T-AFC**), traditional probe noise AFC approach (**T-PN**), proposed probe noise approach I (**PN-I**) and the proposed probe noise approach II (**PN-II**). For reading convenience, we introduce $\Gamma_1 = \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)$ and $\Gamma_2 = \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega)$.

	CR	**SSE**	**TE**
T-AFC	$1-2\mu(n)S_u(\omega)$	$L\frac{\mu(n)}{2}\Gamma_1$	$\frac{\Gamma_2}{2\mu(n)S_u(\omega)}$
T-PN	$1-2\mu(n)S_w(\omega)$	$L\frac{\mu(n)}{2}\Gamma_1$	$\frac{\Gamma_2}{2\mu(n)S_w(\omega)}$
PN-I	$1-2\mu(n)S_w(\omega)$	$L\frac{\mu(n)}{2}\lvert A(\omega)\rvert^2\Gamma_1$	$\frac{\Gamma_2}{2\mu(n)S_w(\omega)}$
PN-II	$1-2\mu(n)(1+\lvert A_0(\omega)\rvert^2)S_w(\omega)$	$L\frac{\mu(n)}{2}\lvert A(\omega)\rvert^2\Gamma_1$	$\frac{\Gamma_2}{2\mu(n)(1+\lvert A_0(\omega)\rvert^2)S_w(\omega)}$

Finally, inserting Eq. (35) in Eq. (18), the resulting PTF $\hat{\xi}(\omega,n)$ is expressed by

$$\begin{aligned}\hat{\xi}(\omega,n) = &\left(1-2\mu(n)(1+|A_0(\omega)|^2)S_w(\omega)\right)\hat{\xi}(\omega,n-1)\\ &+L\mu^2(n)(1+|A_0(\omega)|^2)S_w(\omega)|A(\omega)|^2\sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{x_{ij}}(\omega)\\ &+\sum_{i=1}^{P}\sum_{j=1}^{P}G_i(\omega)G_j^*(\omega)S_{\check{h}_{ij}}(\omega). \qquad (36)\end{aligned}$$

3.3 Discussion

3.3.1 Resulting Expressions for All Approaches

Eqs. (22), (23), (29) and (36) are first-order difference equations in $\hat{\xi}(\omega,n)$ and determine the behavior of the corresponding systems. In particular, we determine the convergence rate describing the decay rate of $\hat{\xi}(\omega,n)$ per sample period, and the steady-state behavior $\lim_{n\to\infty}\hat{\xi}(\omega,n)$ which is the sum of steady-state and tracking errors upon convergence of $\hat{\xi}(\omega,n)$. The steady-state error describes the lowest possible steady-state value of $\hat{\xi}(\omega,n)$, whereas the tracking error is the additional error to that due to the variations in the acoustic feedback paths. The resulting expressions are given in Table 1, for ease of a comparison between the different approaches.

3.3.2 T-PN vs. T-AFC

It is seen from Table 1 that the only difference between the T-AFC approach and the T-PN approach is that for the convergence rate and the tracking error, $S_u(\omega)$ is replaced

by $S_w(\omega)$. Because $S_w(\omega)$ must generally be much lower than $S_u(\omega)$ to ensure the added probe noise is inaudible, the convergence rate in T-PN is reduced by the factor $S_u(\omega)/S_w(\omega)$, which is typically as large as 30, and the tracking error is increased by the same amount.

3.3.3 PN-I vs. T-PN

The only modification introduced by the PN-I approach is the scaling of the steady-state error by the factor of $|A(\omega)|^2$ compared to the T-PN approach. Thus, depending on $|A(\omega)|^2$, the PN-I approach has the capability of reducing the steady-state error compared to the T-PN approach, while maintaining the convergence rate and tracking error. This is obtained for $|A(\omega)|^2 < 1$, i.e. in frequency regions where the enhancement filter $\mathbf{a}$ can (partly) predict $e_i(n)$ based on its past samples, via $\hat{e}_i(n) = \sum_{k=D}^{L_a} a(k)e_i(n-k)$. Recall that the enhancement filter $\mathbf{a}$ is found as the MMSE long-term prediction error filter in Eq. (12), $e_i(n)$ and $-\hat{e}_i(n)$ can be considered as the direct and prediction part of the filtered error signal $\check{e}_i(n) = e_i(n) + \hat{e}_i(n)$. Thus, $\mathbf{a}$ is able to (partly) predict/remove the disturbing signals, e.g. the incoming signals $x_i(n)$, for the estimation of $\mathbf{h}_i(n)$, as long as the autocorrelation function $r_x(k) = E[x_i(n)x_i(n-k)]$ has nonzero lags for $|k| \geq D$. This typically occurs for tonal signals with clear spectral peaks, where $-\hat{e}_i(n)$ provides a precise estimate of $e_i(n)$, so that $E[\check{e}_i^2(n)] \ll E[e_i^2(n)]$ and $|A(\omega)| \ll 1$. Thus, a reduction in steady-state error is expected using the PN-I approach, particularly in frequency regions of $e_i(n)$ with distinct spectral peaks.

On the other hand, as shown in Appendices A and B, due to the structure of the enhancement filter $\mathbf{a}$, in particular, the constraint of D given in Eq. (11), $\mathbf{a}$ does not have any influence on the probe noise signal $w(n)$ either in the expected value $E[\hat{\mathbf{h}}_i(n)]$ nor in the covariance calculation of $\hat{\mathbf{H}}_{ij}(n)$. Thus, the enhancement filter $\mathbf{a}$ can be considered statistically transparent for the probe noise signal $w(n)$ in the estimation of $\mathbf{h}_i(n)$.

To summarize, in the PN-I approach, the different characteristics of the specifically designed enhancement filter $\mathbf{a}$ on the probe noise signal with limited correlation time and disturbing signals lead to a reduced steady-state error without sacrificing the convergence rate and tracking error. The degree of reduction in steady-state error depends on the capability of the enhancement filter to predict/remove the disturbing signals. Typically, a better prediction and thereby higher reduction in steady-state error can be obtained for tonal signals. Furthermore, it is possible to apply an increased step size $\mu(n)$ to obtain a higher convergence rate and lower tracking error in the PN-I approach, while still obtaining an unchanged steady-state error as in the T-PN approach. In this way, part of the drop in convergence rate associated with T-PN approach can be regained.

3.3.4 PN-II vs. PN-I

The idea behind the PN-II approach is similar to the PN-I approach, i.e. utilizing the long-term prediction characteristic of the enhancement filter $\mathbf{a}$. The copies of $\mathbf{a}$ to generate the filtered probe noise signal $\breve{w}(n) = w(n) + \hat{w}(n)$ makes it possible to achieve further improvements, where $w(n)$ and $-\hat{w}(n) = -\sum_{k=D}^{L_a} a(k)w(n-k)$ can be considered as the direct and prediction part of the signal $\breve{w}(n)$, however, $w(n)$ and $\hat{w}(n)$ are uncorrelated due to the constraint on D. The introduction of the extra enhancement filters applied to $w(n)$ means that instead of considering the terms involving $\mathbf{w}(n-1)\mathbf{a}^T\mathbf{W}(n-1) = \mathbf{w}(n-1)\mathbf{w}^T(n-1) + \mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)$ in Eq. (26), where $E[\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)] = \mathbf{0}$ as shown in Appendix A, we are now considering the terms involving $\mathbf{W}^T(n-1)\mathbf{a}\mathbf{a}^T\mathbf{W}(n-1)$ in Eq. (32) in the calculation of $\hat{\mathbf{H}}_{ij}(n)$ in the PN-II approach, and we get the additional contribution $\mathbf{A}_0^T\mathbf{R}_w(0)\mathbf{A}_0 \neq \mathbf{0}$ in Eq. (34). This corresponds to utilizing both the direct parts of signals $w(n)$ and $e_i(n)$, and the prediction parts $-\hat{w}(n)$ and $-\hat{e}_i(n)$ of the filtered signals $\breve{w}(n)$ and $\breve{e}_i(n)$ for the estimation of $\mathbf{h}_i(n)$.

In this way, in contrast to the PN-I approach, where the expected disturbing signal power $E[x_i^2(n)]$ is reduced and the expected probe noise power $E[w^2(n)]$ can be considered unchanged, the expected probe noise power $E[\breve{w}^2(n)] = E[w^2(n)] + E[\hat{w}^2(n)]$ for the estimation algorithm and thereby the probe noise to disturbing signal ratio is further increased in the PN-II approach. As the result, the convergence rate and tracking error are increased by the factor of $1 + |A_0(\omega)|^2$, with $|A_0(\omega)| \leq 1$ at the frequency ω where the enhancement filter $\mathbf{a}$ is able to make a reasonable prediction of $e_i(n)$ from its past samples $e_i(n-D), \ldots, e_i(n-L_a+1)$. Hence, in the PN-II approach, the convergence rate can be further increased by the factor $1 \leq 1 + |A_0(\omega)|^2 \leq 2$, and the tracking error is reduced by the same amount, while maintaining the steady-state error as in the PN-I approach.

Although the proposed probe noise approaches PN-I shown in Fig. 3 and especially PN-II shown in Fig. 4 are somewhat similar in structure to the decorrelating prefilter method [9], where prefilters are applied to the loudspeaker and error signals in a similar way to the enhancement filters, their goal and procedure are very different. The goal of the prefilters in [9] is to decorrelate the incoming signals $x_i(n)$ and the loudspeaker signal $u(n)$, whereas the goal of the enhancement filters is to increase the probe noise to disturbing signal ratio. Furthermore, the proposed approaches differ from the decorrelating prefilter method by using long-term prediction error filters as the enhancement filters.

3.4 Verification of Analysis Results

To complete the analytical analysis and discussion, we perform simple simulation experiments to verify the derived PTF expressions in Eqs. (23), (29) and (36), for the different probe noise approaches shown in Figs. 2-4, respectively, and to visually demonstrate the improvements.

The simulations are performed in a closed-loop AFC system in a hearing aid setup with $P = 2$ microphones, using a sampling frequency $f_s = 20$ kHz. The feedback paths $\mathbf{h}_1(n)$ and $\mathbf{h}_2(n)$ are measured from a behind-the-ear hearing aid with an order of about 50. Because the impulse responses $\mathbf{h}_i(n)$ are known, we can compute the true PTF $\xi(\omega, n)$ according to Eq. (17) to verify the derived expressions for $\hat{\xi}(\omega, n)$. We compute $\xi(\omega, n)$ as the average across $R = 100$ simulation runs, i.e. $\xi(\omega, n) \approx \frac{1}{R}\sum_{k=1}^{R} |\sum_{i=1}^{P} G_i(\omega)\tilde{H}_i^k(\omega, n)|^2$, where $\tilde{H}_i^k(\omega, n)$ is the result of the kth simulation run.

A simple beamformer is used, $\mathbf{g}_1 = \mathbf{g}_2 = \frac{1}{2}$. The forward path $\mathbf{f}(n)$ has a delay of 120 samples modeling a hearing aid processing delay of 6 ms, and it has a fixed amplification of approximately 29 dB so that the most critical frequency for system stability can be found at approximately 2.5 kHz, where the magnitude value of the open-loop transfer function is -1 dB and the phase is 0 rad.

The adaptive filters $\hat{\mathbf{h}}_i(n)$ have a length of $L = 64$ and are initialized as $\hat{\mathbf{h}}_i(0) = \mathbf{0}$. The true feedback paths $\mathbf{h}_i(n)$ are fixed during the first part of the simulation, whereas random walk variations with variances $\sigma^2_{h_1} = 4.844 \times 10^{-5}$ and $\sigma^2_{h_2} = 6.484 \times 10^{-5}$ are added during the last 15 s. Three different simulation experiments are carried out using the T-PN, PN-I and the PN-II approaches. The step size values are respectively chosen to be 2^{-24}, 2^{-21} and 2^{-21} for all three experiments, in order to obtain same steady-state errors but different convergence rates and tracking errors.

In each simulation run, new realizations of standard Gaussian stochastic sequences are drawn; the incoming signals $x_i(n)$ and the probe noise signal $w(n)$ are obtained as these sequences filtered by the inverse of the enhancement filter $\mathbf{a}$ and probe noise shaping filter $\mathbf{h}_w$, respectively. Both filters are known and fixed in this simulation experiment, because the goal of this experiment is to verify the derived expressions; we postpone simulation of the more practical situation where these filters are time-varying to the next section. The shaping filter $\mathbf{h}_w$ with a length $L_w = 13$ is created by first computing $S_u(\omega)$ as the PSD $S_{x_{ij}}(\omega)$ scaled by the forward path amplification of 29 dB, and then the PSD $S_w(\omega) = 0.25S_u(\omega)$ is computed as a scaled version of $S_u(\omega)$. Finally, the filter $\mathbf{h}_w$ is designed using the frequency sampling method. The power ratio between the signals $w(n)$ and $u(n)$ is thereby -12 dB; clearly, the probe noise will generally be audible in this case. In the next section, we demonstrate system performance when the noise is created to be inaudible. The enhancement filter $\mathbf{a}$ has a length of $L_a = 96$ with a value of $D = 76$ to fulfill the requirement of $D \geq L + L_w - 1$, and its magnitude response has a sharp notch at 2.5 kHz.

Fig. 5 shows the simulation results verifying the PTF prediction values, at the most critical frequency $\omega = 2\pi l/L$, where $l = 8$, corresponding to 2.5 kHz. In all cases, the values predicted from the derived expressions are successfully verified by the simulation results. Furthermore, the desired steady-state error of approximately -52 dB is obtained for all three approaches, but very clearly, the convergence rates and the tracking errors are completely different, as expected. Due to the difference in step sizes by a factor of 8, the convergence rate is increased and the tracking error is reduced by the same amount

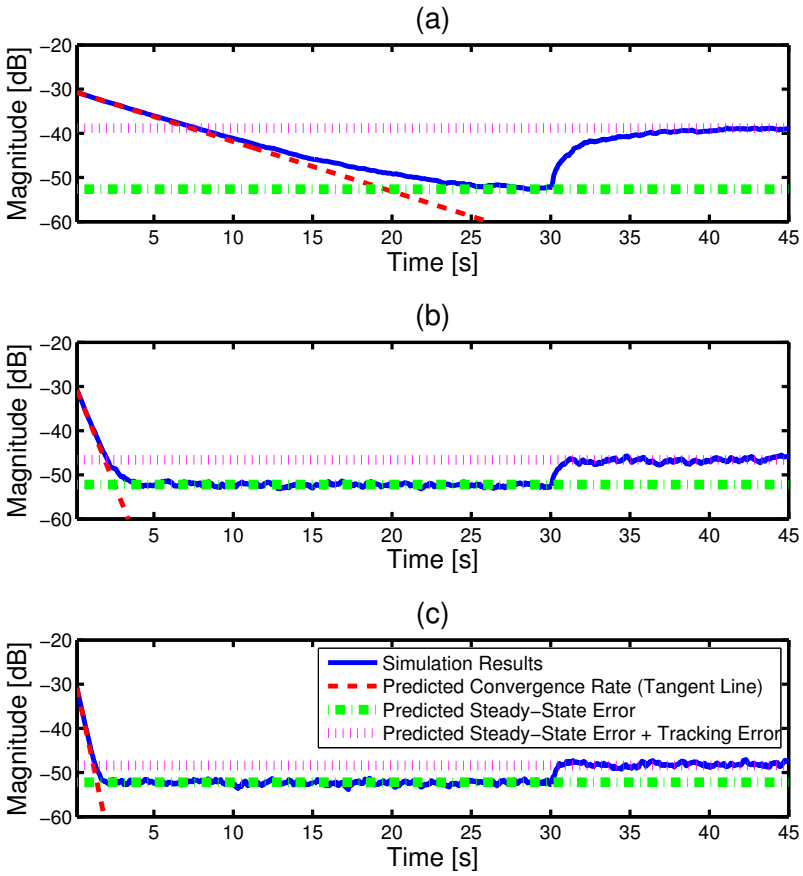


Fig. 5: Verification at the frequency of 2.5 kHz. (a) The traditional probe noise approach (T-PN). (b) The proposed probe noise approach I (PN-I). (c) The proposed probe noise approach II (PN-II).

for the PN-I approach compared to the T-PN approach. Furthermore, it is seen that by using an identical step size in the PN-II approach, the convergence rate and tracking error is further modified by a factor of approximately 1.8 due to the extra enhancement filters applied on the probe noise signal.

4 Demonstration in A Practical Application

In this section, we perform simulations using audio signals in a hearing aid AFC system with $P = 2$ microphones. The goal of the simulations is to show the improvements by

the proposed PN-I and PN-II approaches compared to the T-PN approach in a practical situation, where enhancement filters are time-varying and estimated based on available signals only, the probe noise signal $w(n)$ is generated using a spectral masking model to be inaudible in the presence of the original loudspeaker signal $u(n)$, and the feedback paths $\mathbf{h}_i(n)$ exhibit quick changes, e.g. corresponding to a telephone-to-ear situation, which is known to be a difficult scenario for hearing aid AFC systems. We show that whereas the T-AFC and the T-PN approaches fail to cancel the acoustic feedback, the proposed PN-I and PN-II approaches are efficient in doing so.

4.1 Acoustic Environment

The simulations are carried out using a sampling frequency of $f_s = 20$ kHz. In the following, we provide information of the true feedback paths and the audio signal used to generate the incoming signals in the simulations.

4.1.1 Acoustic Feedback Paths

The true acoustic feedback paths denoted as $\mathbf{h}_i(n)$ in Figs. 1-4 are obtained by measurements from a behind-the-ear hearing aid while worn by a test person. The hearing aid has two omnidirectional microphones and a loudspeaker. We divide the entire simulation into two different periods. In both periods, the true feedback paths $\mathbf{h}_i(n)$ are stationary. At the transition between the periods, we change the feedback paths momentarily to simulate a situation where the hearing aid user makes a phone call and places a telephone close to the ear and thereby the hearing aid. This change of feedback paths is usually very challenging for AFC systems, because sound reflected on the phone/hand back to the microphones increases the feedback path magnitude response by as much as 16 dB [29], almost momentarily, and the AFC system must adapt to the new acoustic feedback paths very quickly to prevent the system from becoming unstable.

The feedback paths used before the transition were measured without any obstacles in the close proximity of the hearing aid, whereas the feedback paths used after the transition were measured when a telephone is closely placed to the ear (less than 1 cm). Fig. 6 shows the impulse and frequency responses of the true feedback paths. It is clearly seen that the telephone-to-ear transition in this example increases the magnitude response in the order of $5 - 10$ dB for most frequencies.

4.1.2 Incoming Signals

The bias problem in general AFC systems typically occurs for tonal signals due to their long correlation time. Although the traditional probe noise approaches can be used to avoid the bias problem in this situation, other side effects such as decreased convergence rate would appear. Therefore, in order to make the demonstration most convincing, we choose an audio signal which has some significant spectral peaks. In particular, we

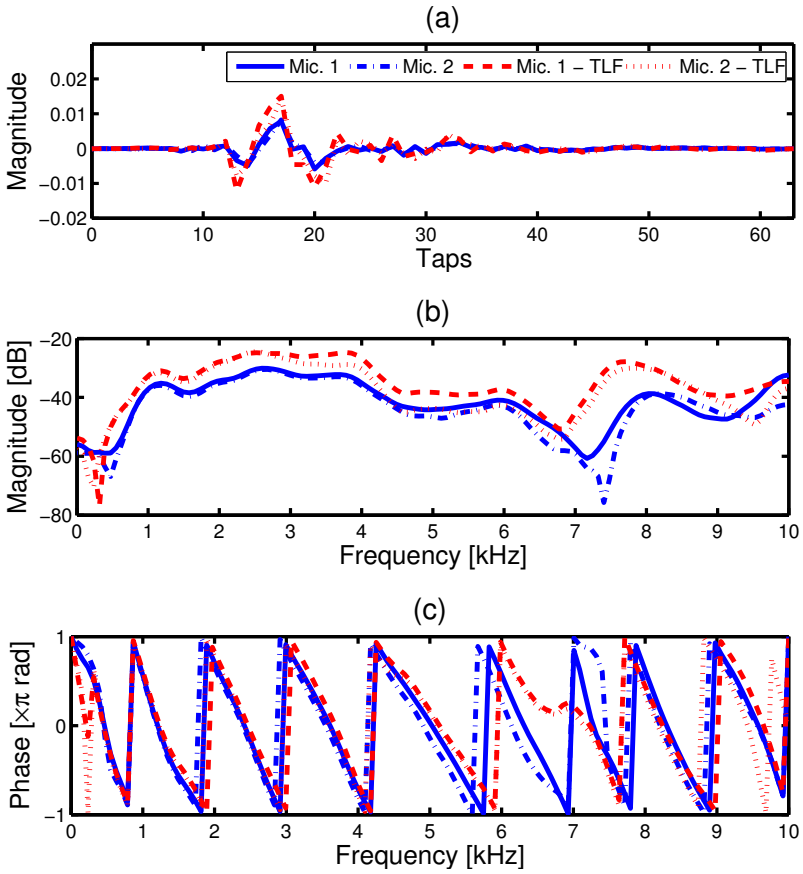


Fig. 6: The measured acoustic feedback paths without and with a telephone closely placed to the hearing aid. (a) Impulse response. (b) Magnitude response. (c) Phase response.

choose an audio signal with a very dominating flute sound around 2.5 kHz as shown in the spectrogram in Fig. 7.

The audio signal shown in Fig. 7 is used as a basis for the incoming signals $x_1(n)$ and $x_2(n)$. For a longer simulation, this audio signal is repeated. In order to perform beamforming, the two hearing aid microphones are typically aligned in the horizontal plane and in the same direction as the face of the hearing aid user; the distance between them is often about 15 mm. In the following simulations, we simply apply a delay to model the distance between the microphones. The audio signal of Fig. 7 is used as the incoming signal $x_1(n)$, whereas the incoming signal $x_2(n)$ is generated by delaying $x_1(n)$

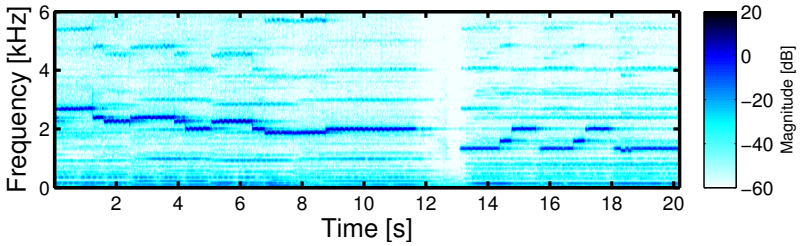


Fig. 7: The spectrogram of an audio signal used for generating the incoming signals. The window size is 512 samples with 50% overlap, and a Hanning window is applied. Furthermore, for reading convenience, we limit the frequency axis to $0-6$ kHz, because most frequency content of the signal are found in this range.

by one sample. This simulates the source signal coming from the frontal direction, with a distance between the two microphones of about 17 mm.

4.2 System Setup

4.2.1 Forward Path and Beamformer

Similar to the simulation experiment in Sec. 3.4, we apply a simple beamformer by setting $\mathbf{g}_1 = \mathbf{g}_2 = \frac{1}{2}$. Furthermore, a hearing aid input-to-output processing delay is typically around $4-8$ ms [30]; in this simulation experiment, we model this as a pure delay of 120 samples corresponding to 6 ms in the forward path $\mathbf{f}(n)$.

In contrast to the experiment in Sec. 3.4, the forward path $\mathbf{f}(n)$ in the present experiment provides a time-varying amplification using a single-channel fullband compressor [31]. The amplification over time is computed as a function of the power level of the signal $\bar{e}(n)$. The compressor provides, for all frequencies, an amplification of 29 dB when the estimated power level is below a certain point, and the amplification is reduced by the excess amount of the estimated power level above this point. With the chosen compressor settings and the acoustic feedback paths, the most critical frequency is found at about 2.5 kHz, where the magnitude of the open-loop transfer function is about -1 dB and the phase 0 rad at the beginning of the simulation; it means that the system initially is close to instability without an AFC system. At the feedback path transition, the worst-case magnitude value of the open-loop transfer function increases momentarily to about 4.5 dB without an AFC system, and the system would certainly become unstable without a properly working AFC system.

4.2.2 AFC using Delayless Subband Adaptive Filters

In practical applications, implementing AFC using a subband structure is often preferred for obtaining higher convergence rate and a reduction in computational complexity [32]. In this work, we apply a delayless subband adaptive filter (SAF) in a closed-loop structure [33, 34] to obtain $\hat{\mathbf{h}}_i(n)$ in the estimation blocks shown in Figs. 1-4.

The length of the fullband filter $\hat{\mathbf{h}}_i(n)$ is chosen to be $L = 64$. The subband NLMS step size for the PN-I and PN-II approaches is chosen as $\mu = 2^{-12}$ for all subbands except the lowest one, where the step size is set to 0. Thus, AFC is not performed below approximately 500 Hz, because there is generally no feedback problem at the lowest frequencies in hearing aid applications, as e.g. seen in Fig. 6(b). For the T-PN and T-AFC approaches, the step size is decreased by a factor of 6, so that the steady-state error is approximately the same for all approaches.

4.2.3 Probe Noise Generation

The probe noise signal $w(n)$ should be generated with the highest possible signal power at each frequency while being inaudible in the presence of the original loudspeaker signal $u(n)$. This can e.g. be achieved by using perceptual audio coding techniques, see e.g. [35] and the references therein, based on the masking effects of the human auditory system [36]. In this work, we generate the probe noise signal using a spectral masking model based on [37]. For a given loudspeaker signal $u(n)$, the model estimates a masking threshold $M(\omega, n)$; ideally, additive and uncorrelated noise shaped according to this threshold would be inaudible in the presence of $u(n)$.

The shaping filter $\mathbf{h}_w(n)$ with length $L_w = 128$ is created using the frequency sampling filter design method, based on $M(\omega, n)$. In order to verify that the generated probe noise $w(n)$ is essentially inaudible in the presence of $u(n)$, we performed control measurements, based on the perceptual evaluation of speech quality (PESQ) and perceptual evaluation of audio quality (PEAQ) models, described in [38] and [39]. More specifically, we use the MATLAB implementations of PESQ and PEAQ provided in [40] and [41] for our verifications. The explanations of the output scores from these PESQ and PEAQ implementations are given in Table 2. Both scores are related to the mean opinion scores [42].

For each noise induced test signal, PESQ or PEAQ values are computed. For comparison, we also evaluated test signals injected with white noise at different fullband signal-to-noise ratio (SNR) of 60 dB, 40 dB and 20 dB, respectively. The results are given in Table 3.

From Table 3, it is seen that the generated probe noise is rated somewhere between imperceptible and perceptible but not annoying, which is very satisfactory. On the other hand, using white noise as probe noise, the SNR must be somewhere between 40 and 60 dB in order to obtain similar sound quality. However, the fullband SNRs between the test signals and the perceptually generated probe noise signals are generally found

Table 2: The output scores given by the applied PESQ and PEAQ models.

PESQ	PEAQ	Quality	Description of Impairment
4.5	0	Excellent	Imperceptible
4	-1	Good	Perceptible but not annoying
3	-2	Fair	Slightly annoying
2	-3	Poor	Annoying
1	-4	Bad	Very annoying

Table 3: Several test signals with inserted probe noise are objectively evaluated using PESQ or PEAQ. The probe noises are either perceptually generated probe noise (**PGPB**) or white noise at the SNR 60 dB (**WN60**), 40 dB (**WN40**) and 20 dB (**WN20**).

	Test Signal	PGPB	WN60	WN40	WN20
PESQ	Danish Male Speech	4.11	4.49	4.06	2.50
	Eng. Female Speech 1	4.03	4.48	3.91	2.44
	Eng. Female Speech 2	4.23	4.48	3.75	2.43
	Eng. Male Speech	4.22	4.48	4.09	2.89
	Japanese Male Speech	4.18	4.49	4.19	3.12
PEAQ	Music - Classic	-0.33	-0.02	-0.95	-3.33
	Music - Flute	-0.89	-0.30	-2.44	-3.85
	Music - Jazz	-0.63	-0.04	-1.06	-3.52
	Music - Symphony	-0.68	0.00	-1.45	-3.52
	Music - Trumpet	-0.70	-0.79	-2.93	-3.84

to be $20 - 25$ dB. Thus, shaping the probe noise in a perceptual relevant manner, it is possible to inject an inaudible probe noise with higher signal power compared to using white noise as probe noise.

4.2.4 Enhancement Filter Estimation

In our simulations, the time-varying enhancement filter $\mathbf{a}(n)$ is estimated based on the error signal $e_1(n)$, according to Eq. (12). The estimated filter coefficients are then copied to different blocks indicated by $\mathbf{a}_i(n)$ in Figs. 1-4. The length of $\mathbf{a}(n)$ is chosen to be $L_a = 260$, and $D = 197$ is used. Thereby, the requirement of D in Eq. (11) is fulfilled.

For simplicity, we used the same SAF approach, as in Sec. 4.2.2, to estimate the nonzero part of the enhancement filter $\mathbf{a}(n)$ with a length-64 adaptive filter. The subband NLMS step size $\mu = 2^{-8}$ is used for all subbands except for the lowest one, where

the step size is set to 0.

4.3 Simulation Results and Discussions

Five simulation experiments are carried out. In the first experiment, we set $\mathbf{h}_i(n) = \hat{\mathbf{h}}_i(n) = \mathbf{0}$ in Fig. 1, this gives an ideal working situation for the hearing aid without acoustic feedback. In the remaining four experiments, the loudspeaker signal is fed back to the microphones through the acoustic feedback paths $\mathbf{h}_i(n)$ as shown in Fig. 6, and AFC is carried out using the different approaches illustrated in Figs. 1-4. The duration of the simulation is 150 s, and the transition of the feedback paths $\mathbf{h}_i(n)$ from the normal to the telephone situation takes place after 50 s. As mentioned, the step sizes for estimation of $\mathbf{h}_i(n)$ are adjusted so that the same steady-state error would be obtained in all approaches.

4.3.1 Howling Suppression

First, we evaluate the abilities to suppress howling by examining the loudspeaker signals from the different AFC approaches. In Fig. 8, the spectrograms are shown for a selected time period and frequency region of the loudspeaker signals from all five simulations. The selected time period includes the transition of the acoustic feedback paths $\mathbf{h}_i(n)$ after 50 s, and the selected frequency region $0-6$ kHz includes the most significant differences among the approaches. It is expected that the system would become unstable, and howling occurs, shortly after the transition, until the AFC system again stabilizes the system by adapting to the new acoustic feedback paths.

Comparing Fig. 8(b) to the reference loudspeaker signal in Fig. 8(a), it is seen that using the T-AFC approach shown in Fig. 1, severe sound distortions are introduced in the resulting loudspeaker signal. The distortion is present before the telephone-to-ear transition at 50 s, and it is caused by biased estimation of $\mathbf{h}_i(n)$, because the incoming signals have very dominant spectral peaks, especially around 2.5 kHz, which leads to a nonzero correlation between the loudspeaker signal and the incoming signal (despite the hearing aid processing delay of 6 ms). Furthermore, howling occurs after the feedback path transition at 50 s, reflected by the additional tonal components in the loudspeaker signal after the transition.

Comparing the results from the T-PN approach shown in Fig. 8(c) to the reference signal in Fig. 8(a), no severe sound distortions are observed before the feedback path transition at 50 s. This is a significant improvement compared to the traditional AFC approach shown in Fig. 8(b) and is achieved because the T-PN approach guarantees unbiased estimation, and because the true feedback paths are stationary, such that the slow convergence rate of T-PN approach is not revealed. However, the system becomes unstable after the transition, as seen by the additional tonal component found at approximately 2.5 kHz after 50 s in Fig. 8(c). The howling disappears over time, although it can not be seen in Fig. 8(c). The long howling time is caused by the slow

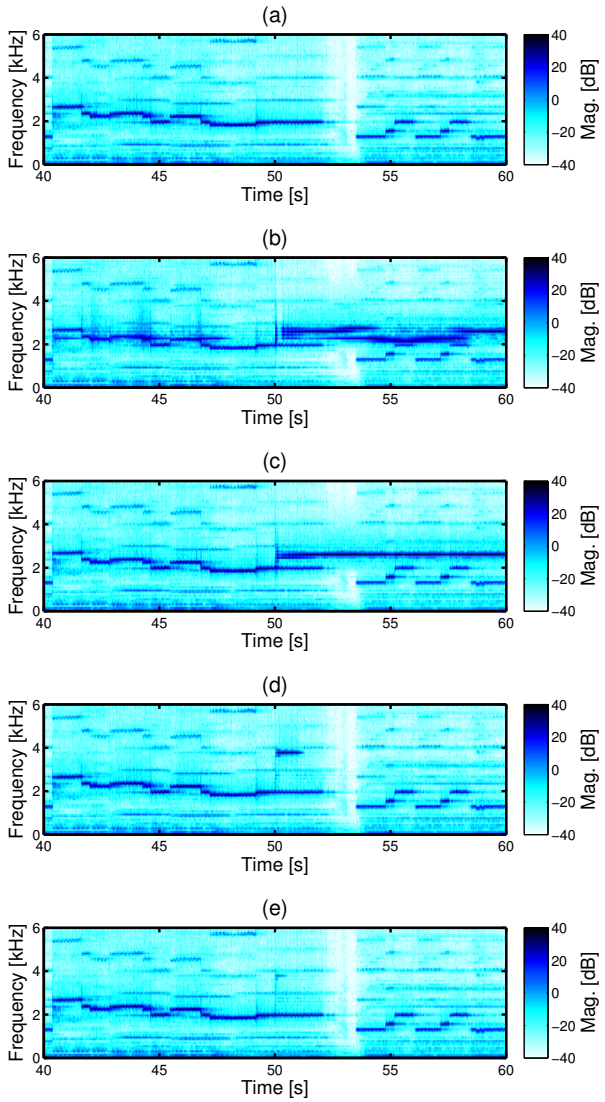


Fig. 8: The spectrograms of the loudspeaker signal in a hearing aid system. The window size is 512 samples with 50% overlap, and a Hanning window is applied. (a) The reference - without acoustic feedback and AFC. (b) The traditional AFC approach (T-AFC). (c) The traditional probe noise approach (T-PN). (d) The proposed probe noise approach I (PN-I). (e) The proposed probe noise approach II (PN-II).

convergence rate of the T-PN approach due to the low probe noise to disturbing signal ratio. This is an example where the T-PN approach faces significant difficulties in practical applications.

Using the PN-I approach, the howling after the feedback path transition is not completely eliminated, as seen in Fig. 8(d). However, it is canceled within 1 s by the AFC system. This is significantly shorter than the case for the T-PN approach. Otherwise, no noteworthy signal distortion is observed from this improved approach.

Finally, using the PN-II approach as shown in Fig. 8(e), the howling is almost avoided after the feedback path transition; the howling is only barely observed after the feedback path transition due to the further increased convergence rate in this approach.

4.3.2 Convergence over Frequencies

We evaluate further the different AFC approaches objectively by using a performance measure, similar to the PTF expression in Eq. (17), defined as

$$\xi'(\omega, n) = \left| \sum_{i=1}^{P} G_i(\omega) \tilde{H}_i(\omega, n) \right|. \tag{37}$$

The magnitude of the open-loop transfer function is given by $|\Theta(\omega, n)| = F(\omega, n)\xi'(\omega, n)$. To ensure system stability, the forward path gain $F(\omega, n)$ for each frequency ω and time index n can be limited to $F(\omega, n) < 1/\xi'(\omega, n)$, so that $|\Theta(\omega, n)| < 1$. This gain limit can be considered as an instantaneous gain margin, which provides the maximum possible gain in the forward path $\mathbf{f}(n)$ before the system might become unstable; obviously, a relatively large gain margin is desired.

In Fig. 9, we show $\xi'(\omega, n)$ at 2.5 kHz, where the incoming signals have the most spectral energy and the enhancement filter has most of its effect around this specific frequency. It is clear that $\xi'(\omega, n)$ has a high steady-state value when using the T-AFC approach due to the bias problem. Using the T-PN approach, $\xi'(\omega, n)$ converges over time, but only at a very slow speed. On the other hand, the convergence rate is significantly increased, by a factor of approximately 6 in this example using the PN-I approach; an additional improvement by a factor of more than 1.6 is obtained in the PN-II approach. The curves in Fig. 9 are computed based on a single simulation run and are therefore less smooth than the curves in Fig. 5, which are the average of 100 simulation runs.

5 Conclusion

In this work, we dealt with probe noise based acoustic feedback cancellation approaches in a multiple-microphone and single-loudspeaker audio system. Traditional probe noise approaches can be used to prevent the major problem of biased adaptive filter estimation

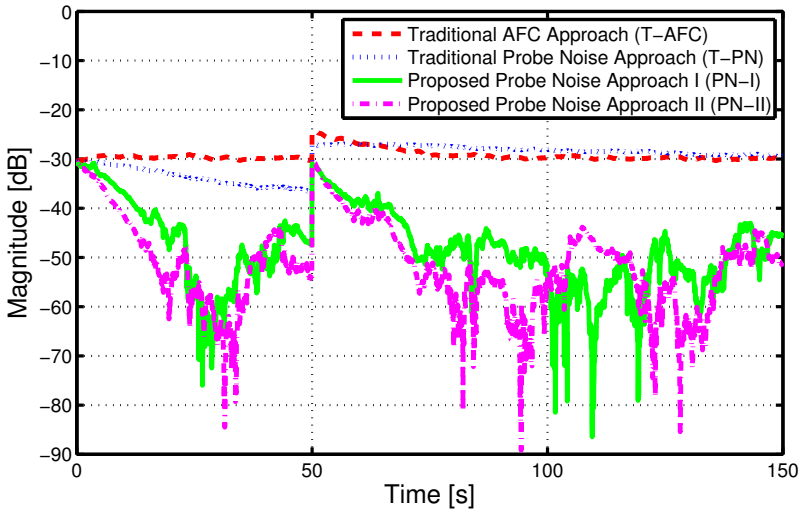


Fig. 9: Evaluation of different AFC approaches using the convergence of $\xi'(\omega, n)$ at 2.5 kHz.

in acoustic feedback cancellation by basing the estimation of acoustic feedback paths on a probe noise signal. However, the convergence rate is generally decreased since the added probe noise must have low power in order to be inaudible. In this paper, we presented and analyzed two probe noise based approaches. We showed that both approaches are capable of increasing the convergence rate significantly without compromising the desired steady-state error, by using *a combination of an inaudible probe noise signal with limited correlation time and the so-called probe noise enhancement filters designed as long-term prediction error filters*. This is verified by simulation experiments, where the proposed probe noise approach I increases the convergence rate by a factor of 6 compared to the traditional probe noise approach, and the proposed probe noise approach II increases the convergence rate further by a factor of 1.6, whereas the traditional acoustic feedback cancellation approach without probe noise completely fails due to the bias problem. Furthermore, we demonstrated through simulation experiments that these proposed approaches are applicable to acoustic feedback cancellation in a realistic hearing aid system.

We believe that the proposed probe noise approaches, which provide unbiased estimation with much higher convergence rate than the traditional probe noise approaches, bring us closer to a complete solution of the biased estimation problem in closed-loop hearing aid systems. The idea behind these approaches could also be applicable in other closed-loop applications such as public address systems and in open-loop acoustic echo cancellation systems. These are considered as future work, which also include a

comparison between the proposed approaches and existing AFC systems in terms of cancellation performance and computational complexity.

A Constraint on Enhancement Filter to Ensure Unbiased Estimation

In this appendix, we show that by using the constraint given in Eq. (11), an unbiased estimation of $\mathbf{h}_i(n)$ is guaranteed in the PN-I approach.

Recall that $w(n)$ is uncorrelated with the incoming signals $x_i(n)$ and the original loudspeaker signal $u(n)$. Then, using Eqs. (24) and (19), the expected value of $\hat{\mathbf{h}}_i(n)$ can be expressed by

$$\begin{aligned} E\left[\hat{\mathbf{h}}_i(n)\right] =& E\Big[\hat{\mathbf{h}}_i(n-1) + \mu(n)\mathbf{w}(n-1)\left(\mathbf{a}^T\mathbf{x}_i(n)\right. \\ & \left. -\mathbf{a}^T\mathbf{U}(n-1)\tilde{\mathbf{h}}_i(n-1) - \mathbf{a}^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)\right)\Big] \\ =& E\left[\hat{\mathbf{h}}_i(n-1) - \mu(n)\mathbf{w}(n-1)\mathbf{a}^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)\right] \\ =& E\left[\hat{\mathbf{h}}_i(n-1) - \mu(n)\mathbf{w}(n-1)\mathbf{w}^T(n-1)\tilde{\mathbf{h}}_i(n-1)\right] \\ & - E\left[\mu(n)\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)\right]. \end{aligned} \tag{38}$$

It is seen that the expectation term $E[\hat{\mathbf{h}}_i(n-1)-\mu(n)\mathbf{w}(n-1)\mathbf{w}^T(n-1)\tilde{\mathbf{h}}_i(n-1)]$ in Eq. (38) follows a standard LMS algorithm and therefore provides an unbiased estimation of $\mathbf{h}_i(n)$. However, we need to consider the last term of $E[\mu(n)\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)]$ in Eq. (38), which occurs due to the introduction of the enhancement filter $\mathbf{a}$, where the desired filtered probe noise signal $-\mathbf{w}(n-1)\tilde{\mathbf{h}}_i(n-1)$ can be modified by $\mathbf{a}_0$ and thereby may introduce a bias in $E[\hat{\mathbf{h}}_i(n)]$. Introducing the vector $\mathbf{w}_{h_i}(n) = \mathbf{W}(n)\tilde{\mathbf{h}}_i(n) = [w_{h_i}(0,n),\ldots,w_{h_i}(L_a-1,n)]^T$, its element is given by

$$w_{h_i}(l,n) = \sum_{k=0}^{L-1} \tilde{h}_i(k,n)w(n-l-k). \tag{39}$$

The last term in Eq. (38) can now be written as

$$E\left[\mu(n)\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)\right] = \mu(n)E\left[\mathbf{w}(n-1)\mathbf{w}_{h_i}^T(n-1)\right]\mathbf{a}_0. \tag{40}$$

The expected value $E[\mathbf{w}(n-1)\mathbf{w}_{h_i}^T(n-1)]$ is further expressed by Eq. (41), where we use the notation $r_w(k) = E[w(n)w(n-k)]$. It follows that $r_w(k) = 0\ \forall\ |k| \geq L_w$ because $w(n)$ is generated using an L_w-1 order shaping filter $\mathbf{h}_w(n)$. Thus, it can be seen from Eq. (41) that all entries in the columns $L+L_w-1$ through L_a of the matrix $E[\mathbf{w}(n-1)\mathbf{w}_{h_i}^T(n-1)]$ are equal to zero because these entries only involve the

$$E\left[\mathbf{w}(n-1)\mathbf{w}_{\tilde{h}_i}^T(n-1)\right] = \begin{bmatrix} \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k) & \dots \\ \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k-1) & \dots \\ \vdots & \dots \\ \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k-L+1) & \dots \end{bmatrix}$$
$$\begin{bmatrix} \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k+1) & \dots & \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k+L_a-1) \\ \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k) & \dots & \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k+L_a-2) \\ \vdots & \ddots & \vdots \\ \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k-L+2) & \dots & \sum_{k=0}^{L-1}\tilde{h}_i(k,n-1)r_w(k+L_a-L) \end{bmatrix} \quad (41)$$

autocorrelation values $r_w(k)\ \forall\ |k| \geq L_w$. It means that by imposing the constraint introduced in Eq. (11), the vector $E[\mu(n)\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)\tilde{\mathbf{h}}_i(n-1)]$ in Eqs. (40) and (38) equals a null-vector. It is now seen that Eq. (38) follows a standard LMS algorithm and thereby provides an unbiased estimation of $\mathbf{h}_i(n)$ [3].

B Influence of Enhancement Filter on Probe Noise

Under the constraint of $D \geq L + L_w - 1$ in Eq. (11), we show that $E[\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)] = \mathbf{0}$.

Define $r_w(k) = E[w(n)w(n-k)]$. It follows that

$$\begin{aligned} & E\left[\mathbf{w}(n-1)\mathbf{a}_0^T\mathbf{W}(n-1)\right] \\ =& E\left[\mathbf{w}(n-1)\bar{\mathbf{w}}^T(n-1)\right]\mathbf{A}_0 \\ =& \begin{bmatrix} r_w(0) & \dots & r_w(L_w-1) & \dots & r_w(L+L_a-2) \\ \vdots & \dots & \vdots & \dots & \vdots \\ r_w(1-L) & \dots & r_w(L_w-L) & \dots & r_w(L_a-1) \end{bmatrix}\mathbf{A}_0 \\ =& \mathbf{0}, \end{aligned} \quad (42)$$

where we defined $\bar{\mathbf{w}}(n)$ and $\mathbf{A}_0$ in Sec. 3.2.1. Eq. (42) is valid because the first $D-1$ samples of $\mathbf{a}_0 = [a_0(0), \dots, a_0(L_a-1)]^T$ are zeros and $r_w(k) = 0\ \forall\ |k| \geq L_w$.

Acknowledgment

The authors would like to thank the reviewers and the associate editor for their valuable suggestions and comments.

Furthermore, M. Guo would like to thank T. B. Elmedyb for discussing the initial idea of this work.

References

[1] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Heidelberg, Germany: Springer, May 2001.

[2] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

[3] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, US: Prentice Hall, Sep. 2001.

[4] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, US: Wiley, Jun. 2003.

[5] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Elsevier J. Franklin Inst.*, vol. 343, no. 6, pp. 545–573, Sep. 2006.

[6] M. R. Schroeder, "Improvement of acoustic-feedback stability by frequency shifting," *J. Acoust. Soc. Am.*, vol. 36, no. 9, pp. 1718–1724, Sep. 1964.

[7] H. A. L. Joson, F. Asano, Y. Suzuki, and T. Sone, "Adaptive feedback cancellation with frequency compression for hearing aids," *J. Acoust. Soc. Am.*, vol. 94, no. 6, pp. 3248–3254, Dec. 1993.

[8] J. Hellgren, "Analysis of feedback cancellation in hearing aids with filtered-X LMS and the direct method of closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 2, pp. 119–131, Feb. 2002.

[9] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.

[10] J. M. Kates, "Feedback cancellation in hearing aids: Results from a computer simulation," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 553–562, Mar. 1991.

[11] J. A. Maxwell and P. M. Zurek, "Reducing acoustic feedback in hearing aids," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 4, pp. 304–313, Jul. 1995.

[12] P. M. J. Van Den Hof and R. J. P. Schrama, "An indirect method for transfer function estimation from closed loop data," *Elsevier Automatica*, vol. 29, no. 6, pp. 1523–1527, Nov. 1993.

[13] N. A. Shusina and B. Rafaely, "Unbiased adaptive feedback cancellation in hearing aids by closed-loop identification," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 658–665, Mar. 2006.

[14] T. van Waterschoot and M. Moonen, "Assessing the acoustic feedback control performance of adaptive feedback cancellation in sound reinforcement systems," in *Proc. 17th European Signal Process. Conf.*, Aug. 2009, pp. 1997–2001.

[15] J. Hellgren and U. Forssell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 906–913, Nov. 2001.

[16] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "On acoustic feedback cancellation using probe noise in multiple-microphone and single-loudspeaker systems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 283–286, May 2012.

[17] M. Guo, S. H. Jensen, and J. Jensen, "An improved probe noise approach for acoustic feedback cancellation," in *Proc. 7th IEEE Sensor Array Multichannel Signal Process. Workshop*, Jun. 2012, pp. 497–500.

[18] C. P. Janse and C. C. Tchang, "Acoustic feedback suppression," Int. Patent Application, WO 2005/079109 A1, Aug. 2005.

[19] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of adaptive feedback and echo cancelation algorithms in a general multiple-microphone and single-loudspeaker system," in *Proc. 2011 IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 433–436.

[20] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification – a survey," *Elsevier Automatica*, vol. 26, no. 1, pp. 7–21, Jan. 1990.

[21] S. Gunnarsson and L. Ljung, "Frequency domain tracking characteristics of adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 1072–1089, Jul. 1989.

[22] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Heidelberg, Germany: Springer, Jun. 2001.

[23] A. Lombard, K. Reindl, and W. Kellermann, "Combination of adaptive feedback cancellation and binaural adaptive filtering in hearing aids," *EURASIP J. Advances Signal Process.*, vol. 2009, pp. 1–15, Apr. 2009.

[24] J. R. Deller, Jr., J. G. Proakis, and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*. New York City, NY, US: Macmillan Publishers, Apr. 1993.

[25] H. Nyquist, "Regeneration theory," *Bell System Tech. J.*, vol. 11, pp. 126–147, 1932.

[26] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of acoustic feedback/echo cancellation in multiple-microphone and single-loudspeaker systems using a power transfer function method," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5774–5788, Dec. 2011.

[27] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory*. Cambridge, MA, US: MIT Press, 1984.

[28] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. Hanover, MA, US: Now Publishers Inc., Jan. 2006.

[29] J. Hellgren, T. Lunner, and S. Arlinger, "Variations in the feedback of hearing aids," *J. Acoust. Soc. Am.*, vol. 106, no. 5, pp. 2821–2833, Nov. 1999.

[30] A. Spriet, M. Moonen, and J. Wouters, "Evaluation of feedback reduction techniques in hearing aids based on physical performance measures," *J. Acoust. Soc. Am.*, vol. 128, no. 3, pp. 1245–1261, Sep. 2010.

[31] H. Dillon, *Hearing Aids*. Stuttgart, Germany: Thieme, May 2001.

[32] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

[33] D. R. Morgan and J. C. Thi, "A delayless subband adaptive filter architecture," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1819–1830, Aug. 1995.

[34] J. Huo, S. Nordholm, and Z. Zang, "New weight transform schemes for delayless subband adaptive filtering," in *Proc. IEEE Global Telecommunications Conf.*, vol. 1, Nov. 2001, pp. 197–201.

[35] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. IEEE*, vol. 88, no. 4, pp. 451–513, Apr. 2000.

[36] B. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. Bingley, UK: Emerald Group Publishing Limited, Apr. 2003.

[37] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 2, pp. 314–323, Feb. 1988.

[38] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, Int. Telecommun. Union, 2001.

[39] ITU-R Recommendation BS.1387-1, *Method for objective measurements of perceived audio quality*, Int. Telecommun. Union, 1998.

[40] P. C. Loizou, *Speech Enhancement: Theory and Practice.* Boca Raton, FL, US: CRC Press, Jun. 2007.

[41] P. Kabal, *Perceptual Evaluation of Audio Quality (PEAQ)*, http://www-mmsp.ece.mcgill.ca/documents/, 2004.

[42] ITU-T Recommendation P.800, *Methods for subjective determination of transmission quality*, Int. Telecommun. Union, 1996.

Paper E

On the Use of a Phase Modulation Method for Decorrelation in Acoustic Feedback Cancellation

Meng Guo, Søren Holdt Jensen, Jesper Jensen, and Steven L. Grant

Published in
Proc. 20th European Signal Process. Conf., Aug. 2012, pp. 2000–2004.

The layout has been revised.

On the Use of a Phase Modulation Method for Decorrelation in Acoustic Feedback Cancellation

Meng Guo, Søren Holdt Jensen, Jesper Jensen, and Steven L. Grant

Abstract

A major problem in using an adaptive filter in acoustic feedback cancellation systems is that the loudspeaker signal is correlated with the signals entering the microphones of the audio system, leading to biased filter estimates. One possible solution for reducing this problem is by means of decorrelation. In this work, we study a subband phase modulation method, which was originally proposed for decorrelation in multichannel acoustic echo cancellation systems. We determine if this method is effective for decorrelation in acoustic feedback cancellation systems by comparing it to a structurally similar frequency shifting decorrelation method. We show that the phase modulation method is suitable for decorrelation in a hearing aid acoustic feedback cancellation system, although the frequency shifting method is in general slightly more effective.

1 Introduction

Adaptive filters have been widely used in both acoustic echo cancellation (AEC) for audio and communication systems and acoustic feedback cancellation (AFC) for sound reinforcement systems. The goal of the adaptive filters in both cases is to model the acoustic signal paths from loudspeakers to microphones of audio systems.

A major problem when using adaptive filters in stereo and/or multichannel AEC systems is the so-called non-uniqueness problem due to the fact that the loudspeaker signals are strongly correlated [1]. It can be shown that the adaptive filter estimates do not converge correctly to the true acoustic echo paths. In AFC systems, on the other hand, the main problem in using adaptive filters is the biased adaptive filter estimation of the acoustic feedback paths [2], which is caused by the nonzero correlation between the loudspeaker signals and the signals entering the microphones.

In both cases, the biased filter estimation is due to undesired and unavoidable signal correlations in audio systems, although the causes of these signal correlations are different. Many decorrelation methods have been proposed for both stereo AEC and AFC systems in the past. A simple method is to introduce nonlinear distortions to loudspeaker signals as firstly proposed for stereo AEC systems [3] and later studied for AFC

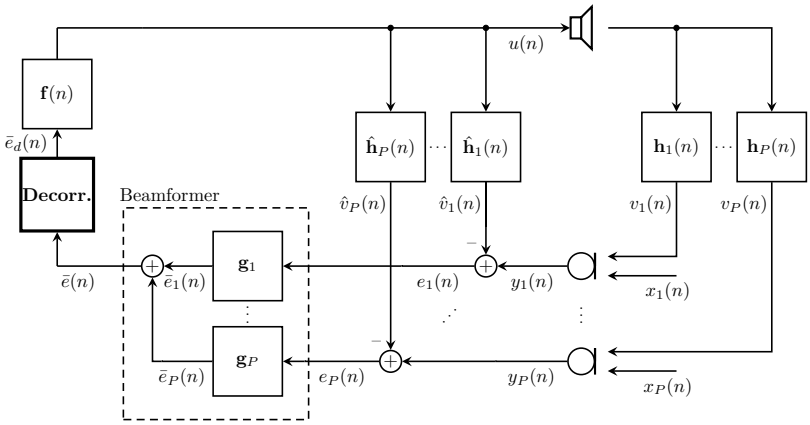


Fig. 1: A general AFC system with a decorrelation function.

systems [4]. Another widely used decorrelation method for both stereo AEC and AFC systems is performed by adding uncorrelated noise to the loudspeaker signals; the added noise is preferably generated such that it is inaudible in the presence of the loudspeaker signals, see e.g. [5–7]. Some other proposed decorrelation methods include introducing time-variable delays on the loudspeaker signals [8], using variable all-pass filtering on the loudspeaker signals to introduce phase shifts [9], applying decorrelation prefilters to the signals used for the adaptive filter estimation [10], and using frequency shifting of the loudspeaker signals [11]. Generally, all these methods might introduce sound quality degradations. Thus, an important compromise in using these methods is sound quality versus decorrelation ability and thereby cancellation performance improvement.

In this work, we study decorrelation methods in an AFC system as shown in Fig. 1, where the AFC is carried out by adaptive filters $\hat{\mathbf{h}}_i(n)$, where n is the time index, $i = 1, ..., P$, and P is the number of microphones. The goal of $\hat{\mathbf{h}}_i(n)$ is to cancel the effects of the true acoustic feedback paths $\mathbf{h}_i(n)$. Furthermore, beamformer filters $\mathbf{g}_i$ are performing a spatial filtering on the feedback compensated signals $e_i(n)$. The block "Decorr." denotes the applied decorrelation function, and the decorrelated signal $\bar{e}_d(n)$ is modified by the forward path $\mathbf{f}(n)$ to form the loudspeaker signal $u(n)$.

More specifically, we study a perceptually motivated decorrelation method by means of subband phase modulation for AFC systems. This phase modulation method was originally introduced for stereo and multichannel AEC systems in [12], where it provided a good cancellation performance without significant sound quality degradation. Here we will determine if this phase modulation method is useful for AFC systems, since not every decorrelation method suitable for AEC would necessarily be appropriate for AFC systems [4].

This phase modulation method is in structure very similar to a frequency shifting

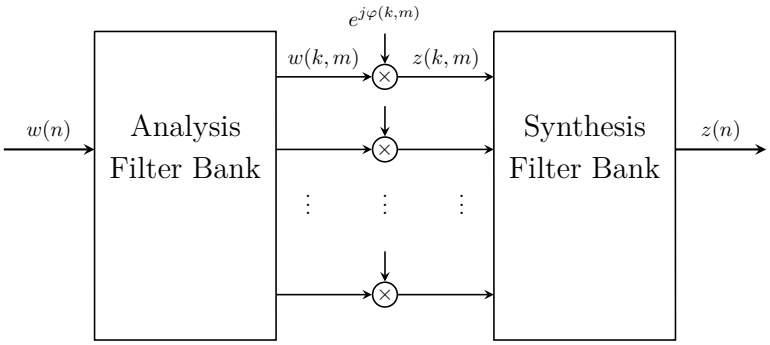


Fig. 2: Decorrelation with subband phase modifications.

decorrelation method [4] carried out in a subband implementation. Thus, we find it obvious to compare both methods. Since the frequency shifting has already been evaluated among other decorrelation methods for AFC systems [4], this comparison will also reveal the effectiveness of the phase modulation method. In particular, we determine analytically the differences between these two methods, before we evaluate AFC performance by simulations, given that sound quality distortions are at same levels for both methods.

2 Analysis of Decorrelation Methods

In this section, we provide details on the subband phase modulation and frequency shifting decorrelation methods. Furthermore, we discuss the differences between them.

Both decorrelation methods are carried out in filter bank subbands, as shown in Fig. 2. An over-sampled analysis filter bank with a decimation factor D divides the input signal $w(n)$ into M subbands with subband index $m = 0, 1, \ldots, M-1$. A complex exponential function $e^{j\varphi(k,m)}$ is then multiplied on each filter bank subband signal $w(k,m)$ to create $z(k,m) = w(k,m)e^{j\varphi(k,m)}$, where $k = 0, 1, 2, \ldots$ is the subband time index with the corresponding fullband time index $n = 0, D, 2D, \ldots$. A synthesis filter bank recombines the processed subband signals $z(k,m)$ to a fullband signal $z(n)$. The difference between these two decorrelation methods is the choice of complex exponential functions $e^{j\varphi(k,m)}$.

The subband structure allows the use of different phase functions $\varphi(k,m)$ over subbands, and $\varphi(k,m)$ can be chosen based on human auditory perception to minimize sound quality degradation. In the phase modulation method proposed in [12], a smooth phase function $\varphi_p(t,m)$ to provide decorrelation with minor sound distortions was sug-

gested as

$$\varphi_p(t,m) = \alpha(m)\sin\left(2\pi f_m t\right), \tag{1}$$

where t denotes continuous time, $\alpha(m)$ is the phase amplitude for the mth subband, and f_m is the modulation frequency. In [12], the optimal values of $\alpha(m)$ and f_m were found by a listening procedure, so that effects of $\varphi_p(t,m)$ would be perceptually insignificant. In particular, a modulation frequency $f_m = 0.75$ Hz was suggested, whereas the phase amplitudes $\alpha(m)$ varied from 10 degrees at low frequencies to 90 degrees above 2.5 kHz. Furthermore, complex conjugated phase functions $\varphi_p(t,m)$ were applied on both microphone channels in a stereo system.

Eq. (1) can also be expressed in the subband time index k using $t = \frac{k}{f_s/D}$, where f_s is the fullband sampling rate, as

$$\varphi_p(k,m) = \alpha(m)\sin\left(2\pi f_m k\frac{D}{f_s}\right). \tag{2}$$

On the other hand, the frequency shifting method is carried out using the complex exponential function $e^{j\varphi_f(k,m)}$ in filter bank subbands, and the phase function $\varphi_f(k,m)$ is given by

$$\varphi_f(k,m) = 2\pi f_0(m)k\frac{D}{f_s}, \tag{3}$$

where $f_0(m)$ denotes the amount of frequency shifting in subband m. Thus, for $f_0(m) > 0$, $\varphi_f(k,m)$ increases linearly with increasing k. Furthermore, using the modulus operator mod(), the wrapped version of $\varphi_f(k,m) > 0$ is expressed by

$$\varphi'_f(k,m) = \mathrm{mod}\left(2\pi f_0(m)k\frac{D}{f_s}, 2\pi\right) - c, \tag{4}$$

where $c = 0$ if $\mathrm{mod}(2\pi f_0(m)k\frac{D}{f_s}, 2\pi) \leq \pi$, and $c = 2\pi$, otherwise.

Fig. 3 shows the phase functions $\varphi_p(k,m)$ and $\varphi'_f(k,m)$ given by Eqs. (2) and (4) for the phase modulation and frequency shifting methods, respectively. The time is computed as $t = \frac{k}{f_s/D}$. There are some obvious similarities, both functions are periodic with a certain frequency and amplitude. In the phase modulation case, the frequency and amplitude of $\varphi_p(k,m)$ are determined by f_m and $\alpha(m)$, respectively. In the frequency shifting case, the frequency of $\varphi'_f(k,m)$ is determined by $f_0(m)$, whereas the maximum and minimum amplitude values are always $\pm\pi$.

An important difference between these phase functions is that $\varphi'_f(k,m)$ is an increasing function in its unwrapped form $\varphi_f(k,m)$ given by Eq. (3), whereas $\varphi_p(k,m)$ is identical to its wrapped function and it is always periodic. Since a frequency shift is introduced proportionally to the temporal derivative of the phase function $\varphi(k,m)$, a constant frequency shift is obtained by using the phase function $\varphi_f(k,m)$, whereas with $\varphi_p(k,m)$ the amount of frequency shift is time-varying with an average of zero.

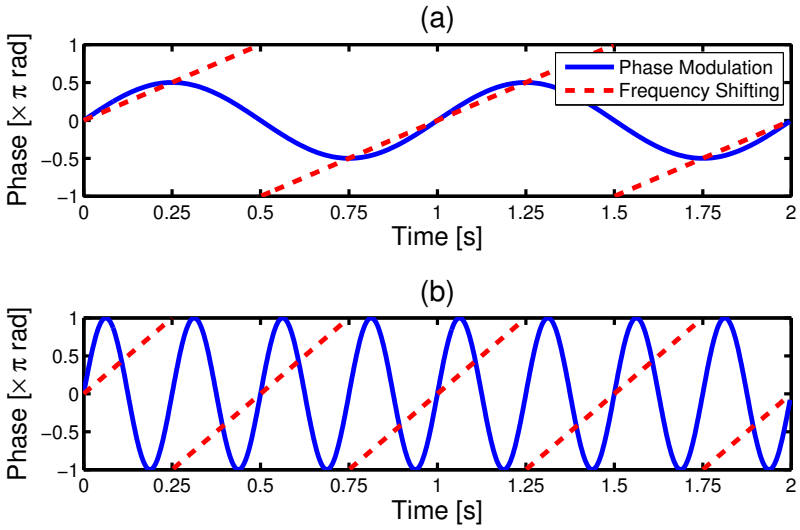


Fig. 3: Phase functions $\varphi_p(k,m)$ and $\varphi'_f(k,m)$ with different parameter values. (a) $\alpha(m) = \frac{\pi}{2}$, $f_m = 1$ Hz, and $f_0(m) = 1$ Hz. (b) $\alpha(m) = \pi$, $f_m = 4$ Hz, and $f_0(m) = 2$ Hz.

3 Sound Quality Considerations

In this section, we choose parameters $f_0(m)$, $\alpha(m)$ and f_m for both decorrelation methods, so that they only introduce insignificant and somewhat equal sound quality degradations.

In many applications, a fullband frequency shifting factor f_0 is commonly chosen as $0 < f_0 \leq 10$ Hz to avoid significant sound quality distortions, as e.g. suggested in [4]. In this work, we let f_0 be subband dependent as $f_0(m)$, and we only perform frequency shifting at higher frequencies to further preserve sound quality, especially for speech signals. We use filter banks with $M = 64$ complex conjugated subbands with a fullband sampling rate of $f_s = 20$ kHz and a decimation factor $D = 8$. Table 1 shows the chosen shifting factors $f_0(m)$ for subbands $m = 0, \ldots, 31$.

The phase function $\varphi_p(k,m)$ for the phase modulation method has two parameters, $\alpha(m)$ and f_m. In this work, we use the phase amplitude values $\alpha(m)$ as given in Table 2. We chose these amplitude values to match the introduced phase amplitude differences between microphone channels in [12]. Furthermore, as demonstrated in Fig. 3, the phase modulation frequency f_m determines the frequency of the periodic phase function $\varphi_p(k,m)$ as the shifting frequency $f_0(m)$ does for $\varphi'_f(k,m)$. From this relation, we choose $f_m = 10$ Hz in an attempt to obtain similar sound qualities in both methods.

We now perform objective sound quality measurements to verify our parameter

Table 1: The frequency shifting factor $f_0(m)$ for subbands $m = 0, \ldots, 31$. The subband bandwidth is 312.5 Hz.

Subband m	0	1	2	3	≥ 4
$f_0(m)$ **[Hz]**	0	0	0	0	10

Table 2: The phase amplitude $\alpha(m)$ for subbands $m = 0, \ldots, 31$. The subband bandwidth is 312.5 Hz.

Subband m	$0-3$	4	5	6	≥ 7
$\alpha(m)$ **[Degrees]**	20	40	70	90	180

Table 3: The output scores from PESQ and PEAQ models.

PESQ	**PEAQ**	**Sound Quality/Description**
4.5	0	Excellent/Imperceptible
4	-1	Good/Perceptible but not annoying
3	-2	Fair/Slightly annoying
2	-3	Poor/Annoying
1	-4	Bad/Very annoying

choices, by using the MATLAB implementations [13, 14] of perceptual evaluation of speech quality (PESQ) and perceptual evaluation of audio quality (PEAQ) models, described in [15] and [16], respectively. Table 3 provides descriptions of the output scores from both models.

Although both models were originally developed to assess relatively mild coding artifacts, we use them to evaluate the relatively small degradations from both decorrelation methods. We compare several test signals $w(n)$ with their processed versions $z(n)$ obtained as shown in Fig. 2 using both decorrelation methods. Table 4 shows the mean, standard deviation, and median values of the determined sound quality scores based on a total of 18 speech and music test signals.

The quality scores in Table 4 show that similar sound quality degradations can be expected from both decorrelation methods, with the chosen parameters of $f_0(m)$, $\alpha(m)$ and f_m. Furthermore, the sound quality degradations are limited in both cases, especially for speech signals. From the determined PESQ scores we can classify the degradation as only slightly perceptual but not annoying. This is because none or only minor modifications are carried out in frequency regions below approximately 1.5 kHz. For music signals, however, the sound quality degradations are more severe since they

Table 4: The mean, standard deviation, and median values of the PESQ and PEAQ scores based on 9 speech signals and 9 music signals processed by the frequency shifting (**FS**) and phase modulation (**PM**) decorrelation methods.

Statistics	**PESQ**		**PEAQ**	
	FS	**PM**	**FS**	**PM**
Mean	4.26	4.30	-1.29	-1.34
Stdv.	0.04	0.11	0.69	0.59
Median	4.27	4.29	-1.00	-1.05

generally have more high frequency contents. Nevertheless, the introduced degradations can still be roughly characterized as perceptible but not annoying. Furthermore, the result from the objective sound quality evaluation was also confirmed by a few experienced listeners.

4 Simulation Experiments

In this section, we perform simulations to evaluate both decorrelation methods in a hearing aid AFC system as shown in Fig. 1 with two microphones ($P = 2$). We determine how effective the phase modulation method is compared to the frequency shifting method, when both methods have sound quality degradations at similar levels as determined in Sec. 3.

The true acoustic feedback paths $\mathbf{h}_i(n)$ remain time invariant in simulations, i.e. $\mathbf{h}_i(n) = \mathbf{h}_i$, and they are obtained by measurements from a behind-the-ear hearing aid with two microphones. Fig. 4 shows the impulse responses $\mathbf{h}_i$ for both microphone channels, where the sampling rate is 20 kHz. From Fig. 4 we observe that the effective length, which covers the nonzero values of $\mathbf{h}_i$, is approximately 50 taps.

The feedback path estimates $\hat{\mathbf{h}}_i(n)$ are obtained using a delayless subband adaptive filter approach [17, 18] with 32 subbands and a decimation factor of 16, where the corresponding fullband adaptive filters $\hat{\mathbf{h}}_i(n)$ have 64 taps, and the subband NLMS algorithm utilizes a step size of $\mu = 2^{-11}$ for subbands above approximately 1.25 kHz, and $\mu = 0$ for low frequency subbands. This is motivated by the fact that AFC is usually not necessary for the lowest frequencies in a hearing aid application. Furthermore, the regularization parameter $\delta = 2^{-14}$ is used in the subband NLMS algorithm.

We evaluate the AFC performance using the coefficient misalignment criterion $\epsilon(n)$ defined as

$$\epsilon(n) = \frac{\sum_{i=1}^{P} \|\mathbf{h}_i - \hat{\mathbf{h}}_i(n)\|^2}{\sum_{i=1}^{P} \|\mathbf{h}_i\|^2}. \tag{5}$$

In simulations, the incoming signal $x_1(n)$ is either a speech or music signal, and we

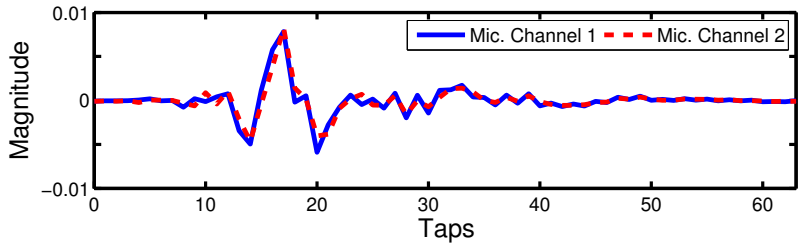


Fig. 4: Measured acoustic feedback paths from a behind-the-ear hearing aid with a sampling rate of $f_s = 20$ kHz.

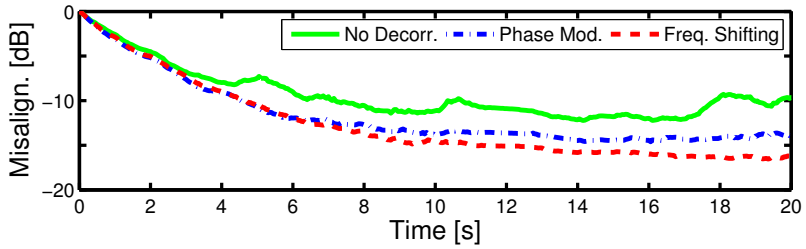


Fig. 5: A representative example simulation result showing the misalignments $\epsilon(n)$ for three different AFC systems.

simply use a delay of one sample between $x_2(n)$ and $x_1(n)$ to model the distance between the microphones. This delay corresponds to a microphone distance of approximately 17 mm, assuming the microphones are aligned in front/rear positions in a horizontal plane, and the sound signal is coming from the front direction.

We use a simple beamformer setup as $\mathbf{g}_1 = \mathbf{g}_2 = \frac{1}{2}$, whereas the forward path $\mathbf{f}(n)$ consists of a delay of 120 samples, corresponding to a hearing aid processing delay of 6 ms. Furthermore, $\mathbf{f}(n)$ consists of a single-channel fullband compressor to provide a maximum amplification of 29.3 dB, and the most critical closed-loop magnitude value without $\hat{\mathbf{h}}_i(n)$ becomes -1 dB at approximately 2.5 kHz.

We compare simulation results in terms of $\epsilon(n)$ between a reference AFC system without applying any decorrelation method and two AFC systems using each decorrelation method. We performed many simulation trials with different speech and music signals as the incoming signals $x_i(n)$. Fig. 5 shows a representative example result for a music signal. We observe that a smaller misalignment $\epsilon(n)$ is already obtained after about 1 s, by using either decorrelation method compared to the system without decorrelation, and the improvement is more than 6 dB in the frequency shifting case at the end of the simulation.

Table 5 shows statistics of the steady-state misalignments from all simulations. For

Table 5: Simple statistics for steady-state misalignments [dB] in AFC systems without decorrelation (**None**), with phase modulation (**PM**) and frequency shifting (**FS**), respectively.

Stat.	**Speech**			**Music**		
	None	**PM**	**FS**	**None**	**PM**	**FS**
Mean	-14.6	-16.0	-16.0	-7.6	-13.6	-14.8
Stdv.	1.4	0.5	0.9	3.6	2.4	2.7
Median	-14.3	-16.0	-16.2	-9.2	-14.5	-16.2

speech signals, the steady-state misalignments $\epsilon(n)$ are close (within 1.4 dB) between all three systems. In fact, $\epsilon(n)$ generally followed each other closely over time (not shown). Hence, there is only small improvements by using either decorrelation method in this case. This is because speech signals in general only cause a limited correlation problem for the estimation of $\hat{\mathbf{h}}_i(n)$ at the higher frequencies above approximately 1.5 kHz, so a decorrelation in this region can not improve AFC performance much further.

On the other hand, the correlation problem is generally more severe for most music signals in the frequency region above 1.5 kHz, and we can expect a more significant AFC improvement when using decorrelation. This is confirmed by our simulations. In Table 5, we observe the average steady-state misalignment improvements for music signals are 6 dB and 7.2 dB in the phase modulation and frequency shifting cases, respectively. In some simulations, these improvements were found to be more than 12 dB. Furthermore, Table 5 also reveals that similar AFC performance to speech signals can be achieved for music signals when using decorrelation.

We should emphasize that the AFC performance improvements achieved from both decorrelation methods depend on the compromise made in sound quality. Thus, we conclude based on the chosen parameter values and sound quality evaluations done in Sec. 3, that the phase modulation decorrelation method is effective in improving hearing aid AFC performance, especially for music signals. However, the structurally very similar frequency shifting decorrelation method is generally slightly better for doing so.

5 Conclusion

In this work, we studied a subband phase modulation decorrelation method originally proposed for stereo and multichannel AEC systems. We compared it to a similar frequency shifting decorrelation method in an AFC system. We determined analytically the differences between these two decorrelation methods. Furthermore, we showed that by choosing appropriate parameter values in the phase modulation method, it is capable of improving the AFC performance without introducing significant sound quality degradations. However, by controlling sound quality degradations at similar and insignificant

levels in both methods, the frequency shifting decorrelation method gives slightly better overall AFC performance, which probably makes it the preferred method, especially when taking its simplicity into account.

References

[1] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic acoustic echo cancellation – an overview of the fundamental problem," *IEEE Signal Process. Lett.*, vol. 2, no. 8, pp. 148–151, Aug. 1995.

[2] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Elsevier J. Franklin Inst.*, vol. 343, no. 6, pp. 545–573, Sep. 2006.

[3] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 2, pp. 156–165, Mar. 1998.

[4] T. van Waterschoot and M. Moonen, "Assessing the acoustic feedback control performance of adaptive feedback cancellation in sound reinforcement systems," in *Proc. 17th European Signal Process. Conf.*, Aug. 2009, pp. 1997–2001.

[5] T. Gänsler and P. Eneroth, "Influence of audio coding on stereophonic acoustic echo cancellation," in *Proc. 1998 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, May 1998, pp. 3649–3652.

[6] A. Gilloire and V. Turbin, "Using auditory properties to improve the behaviour of stereophonic acoustic echo cancellers," in *Proc. 1998 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, May 1998, pp. 3681–3684.

[7] M. Guo, S. H. Jensen, and J. Jensen, "An improved probe noise approach for acoustic feedback cancellation," in *Proc. 7th IEEE Sensor Array Multichannel Signal Process. Workshop*, Jun. 2012, pp. 497–500.

[8] A. Sugiyama, Y. Joncour, and A. Hirano, "A stereo echo canceler with correct echo-path identification based on an input-sliding technique," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2577–2587, Nov. 2001.

[9] M. Ali, "Stereophonic acoustic echo cancellation system using time-varying all-pass filtering for signal decorrelation," in *Proc. 1998 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, May 1998, pp. 3689–3692.

[10] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.

[11] M. R. Schroeder, "Improvement of acoustic-feedback stability by frequency shifting," *J. Acoust. Soc. Am.*, vol. 36, no. 9, pp. 1718–1724, Sep. 1964.

[12] J. Herre, H. Buchner, and W. Kellermann, "Acoustic echo cancellation for surround sound using perceptually motivated convergence enhancement," in *Proc. 2007 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, Apr. 2007, pp. 17–20.

[13] P. C. Loizou, *Speech Enhancement: Theory and Practice*. Boca Raton, FL, US: CRC Press, Jun. 2007.

[14] P. Kabal, *Perceptual Evaluation of Audio Quality (PEAQ)*, http://www-mmsp.ece.mcgill.ca/documents/, 2004.

[15] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, Int. Telecommun. Union, 2001.

[16] ITU-R Recommendation BS.1387-1, *Method for objective measurements of perceived audio quality*, Int. Telecommun. Union, 1998.

[17] D. R. Morgan and J. C. Thi, "A delayless subband adaptive filter architecture," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1819–1830, Aug. 1995.

[18] J. Huo, S. Nordholm, and Z. Zang, "New weight transform schemes for delayless subband adaptive filtering," in *Proc. IEEE Global Telecommunications Conf.*, vol. 1, Nov. 2001, pp. 197–201.

Paper F

Evaluation of State-of-the-Art Acoustic Feedback Cancellation Systems for Hearing Aids

Meng Guo, Søren Holdt Jensen, and Jesper Jensen

To be published in
J. Audio Eng. Soc., 2013.

The layout has been revised.

Evaluation of State-of-the-Art Acoustic Feedback Cancellation Systems for Hearing Aids

Meng Guo, Søren Holdt Jensen, and Jesper Jensen

Abstract

In this work, we evaluate four state-of-the-art acoustic feedback cancellation systems for hearing aid applications. We show that significant improvements in cancellation performance can be made over traditional systems by allowing small alterations in the loudspeaker signal, and a computational complexity increase by a factor of $2-3$*. The evaluation is based on a listening test and objective assessments of simulation results.*

1 Introduction

Acoustic feedback occurs when the output signal of an audio device returns to its microphone and thereby forms an acoustic feedback loop. The typical consequences of acoustic feedback are sound quality degradation and, in the worst-case, howling. Acoustic feedback occurs typically in a sound reinforcement system such as public address systems and hearing aids; especially, it is very likely to occur in a hearing aid, due to the closely placed microphones and loudspeaker, typically only a few millimeters to a few centimeters apart depending on the hearing aid style. In this work, we focus on hearing aid systems.

Acoustic feedback cancellation (AFC) using adaptive filter techniques in a system identification configuration [1] has become the state-of-the-art method for reducing the effect of acoustic feedback [2]. Fig. 1 illustrates a simple hearing aid system with an AFC system, where an adaptive filter $\hat{\mathbf{h}}(n)$ models and cancels the acoustic feedback path $\mathbf{h}(n)$ from the hearing aid loudspeaker to the microphone. The hearing aid processing is represented by the generally time-varying forward path impulse response $\mathbf{f}(n)$, and the microphone signal $y(n)$ consists of the desired incoming signal $x(n)$ and the undesired but unavoidable feedback signal $v(n)$, whereas the loudspeaker signal is denoted by $u(n)$.

In a traditional fullband (FB) AFC system [2], the cancellation is carried out by updating the adaptive filter $\hat{\mathbf{h}}(n)$ using e.g. normalized least mean square (NLMS), affine projection (AP), or recursive least squares (RLS) algorithms [1]. One of the challenges in using this traditional AFC system is that whenever the correlation time of the incoming signal $x(n)$ is longer than the system latency (from the microphones to

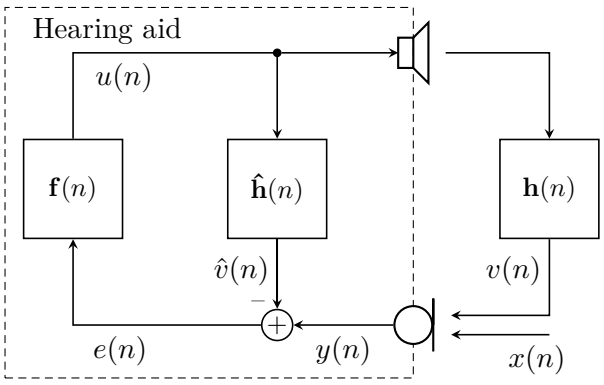


Fig. 1: A simple single microphone hearing aid system. All signals are denoted as discrete-time signals for notational convenience.

the loudspeaker) of the hearing aid, the signals $x(n)$ and $u(n)$ become correlated, and the adaptive filter estimate $\hat{\mathbf{h}}(n)$ is biased [3, 4].

There are different techniques to compensate for this biased estimation problem. The adaptive filter update can be carried out in a non-continuous and open-loop manner as suggested in [5, 6]. The indirect approaches [7, 8] utilize a probe noise (PN) signal to obtain an estimate $\hat{u}(n)$ of $u(n)$, where an unbiased estimation of $\hat{\mathbf{h}}(n)$ based on $\hat{u}(n)$ is possible since $\hat{u}(n)$ is uncorrelated with $x(n)$. The prediction error method (PEM) [9, 10] utilizes a prefilter to whiten the component of the incoming signal $x(n)$ entering the adaptive filter estimation and thereby decorrelate it from $u(n)$. The frequency shifting (FS) technique [11, 12] shifts the spectral contents of the microphone signal $y(n)$ to create the loudspeaker signal $u(n)$; ideally, it decorrelates $x(n)$ and $u(n)$ and thereby facilitates unbiased estimation of $\hat{\mathbf{h}}(n)$. Furthermore, unbiased estimation of $\hat{\mathbf{h}}(n)$ can be achieved by basing the adaptive filter estimation on a probe noise signal added to the loudspeaker signal $u(n)$, see e.g. [2] and the recently proposed probe noise approach in [13].

Ideally, all the above mentioned techniques eliminate the biased estimation problem. However, to reach this goal, they either rely on certain signal models of incoming signals $x(n)$ which are not always valid in practice (e.g. the PEM based techniques), or they introduce modifications to the loudspeaker signal $u(n)$ to decorrelate it from $x(n)$, which might cause sound quality degradations (e.g. the frequency shifting and probe noise based approaches). However, allowing a *minor* sound quality degradation from the decorrelation process can very often improve AFC performance greatly, and it is thereby possible to obtain a much better *overall* sound quality in $u(n)$. Furthermore, for practical applications such as hearing aids, another important issue is the computational complexity. Hence, the trade-off in practice is often between getting a better

AFC performance by obtaining an unbiased estimation of $\hat{\mathbf{h}}(n)$, the degree of introduced sound quality degradations for obtaining this unbiased estimation, and computational complexity.

Some recent studies have compared the performance of different AFC systems. A comparison of different sound reinforcement AFC systems was presented in [14], where the cancellation performance and sound quality degradation are both evaluated objectively. In [15], a comparison based on physical measurements of different commercial hearing aids is reported; in this study, however, it is not clear which AFC system is used in each individual commercial hearing aid, and these AFC systems might not be state-of-the-art. Hence, no reported work exists for comparing different *state-of-the-art AFC algorithms/systems for hearing aids*, in terms of AFC performance, sound quality degradation, and computational complexity. Moreover, the recent introduction of a novel probe noise based AFC system [13] makes such a comparison even more appealing.

Therefore, in this work, we compare four different state-of-the-art AFC systems to a traditional fullband AFC system in terms of their abilities to cancel acoustic feedback, the sound quality degradations they might introduce to decorrelate the loudspeaker signal from the incoming signal, and their computational complexity. This comparison is performed using simulation experiments with realistic hearing aid setups and objective performance measures. To ensure a high sound quality in $u(n)$, the distortions (if any) introduced in $u(n)$ for decorrelation must be controlled; this is carried out by choosing appropriate parameters in different systems, and we evaluate the sound quality using both objective quality measures and a subjective listening test.

The rest of this paper is organized as follows. In Sec. 2, we provide an overview of all considered AFC systems. In Sec. 3, we evaluate the sound quality for two systems which introduce additional sound degradations for decorrelation. In Sec. 4, we perform simulation experiments, using audio signals, to compare cancellation performance for all systems. Sec. 5 outlines the computational complexity of the systems. Finally, we conclude this work in Sec. 6.

2 Overview of Different AFC Systems

In this work, we consider two different structures for adaptive filter estimation: A fullband NLMS algorithm and a subband (SB) NLMS algorithm. Other adaptive algorithms such as AP and RLS can be used for the adaptive estimation of $\hat{\mathbf{h}}(n)$, but the NLMS algorithm is often chosen in hearing aid applications for its simplicity. A subband system is more complex in structure than the fullband system, but it has more freedom in its configuration since one step size parameter per subband is available for the adaptive algorithm compared to a single step size parameter for all frequencies in the fullband system. In addition, due to the subband processing, the subband spectrum is generally flatter than the fullband spectrum, and an increased convergence rate of the adaptive

Table 1: The combinations of adaptive filter structures and decorrelation methods. In this work, we focus on five selected systems.

	No Decor.	**PEM**	**FS**	**PN**
Fullband	✓	✓	–	–
Subband	✓	–	✓	✓

Table 2: Abbreviations and descriptions for AFC systems.

Sys.	**Abbreviation**	**Description**
I	F-AFC	FB AFC system
II	PEM-AFC	FB PEM-based AFC system
III	S-AFC	SB AFC system
IV	FS-AFC	SB FS-based AFC system
V	PN-AFC	SB PN-based AFC system

estimation is thereby possible [16]. Furthermore, in contrast to the fullband system, a subband system is often less sensitive to the biased estimation problem, which would only affect a few subbands in the frequency regions with strong signal correlations between $u(n)$ and $x(n)$, while a correct estimation would still be possible in the remaining frequency regions.

Each adaptive filter structure can be further combined with different decorrelation methods to form a complete AFC system. We focus on the prediction error method because it has been shown to have superior performance compared to many other decorrelation methods [14], and it does not introduce sound degradations in $u(n)$. Furthermore, we focus on the frequency shifting decorrelation method for its performance [14] and simplicity [17]. Finally, we consider the recently proposed probe noise approach in combination with probe noise enhancement [13].

Clearly, there are many AFC systems, if all possible combinations of adaptive filter structures and decorrelation methods were considered. To limit our study to a practical size, we focus on the AFC systems outlined in Table 1. We choose the traditional fullband AFC system without any decorrelation as a reference; the fullband PEM is chosen as a basic version of various PEM systems; the subband system without decorrelation is chosen for a direct adaptive filter structure comparison to the reference fullband AFC system; we choose the subband system with frequency shifting to evaluate a simple alternative to the subband system with probe noise for decorrelation as described in [13]. We evaluate these five systems in a hearing aid application with realistic parameter settings.

In the following, we describe each of the five AFC systems. Table 2 provides descriptions and abbreviations for the five systems. Fig. 2 shows a compact block diagram for

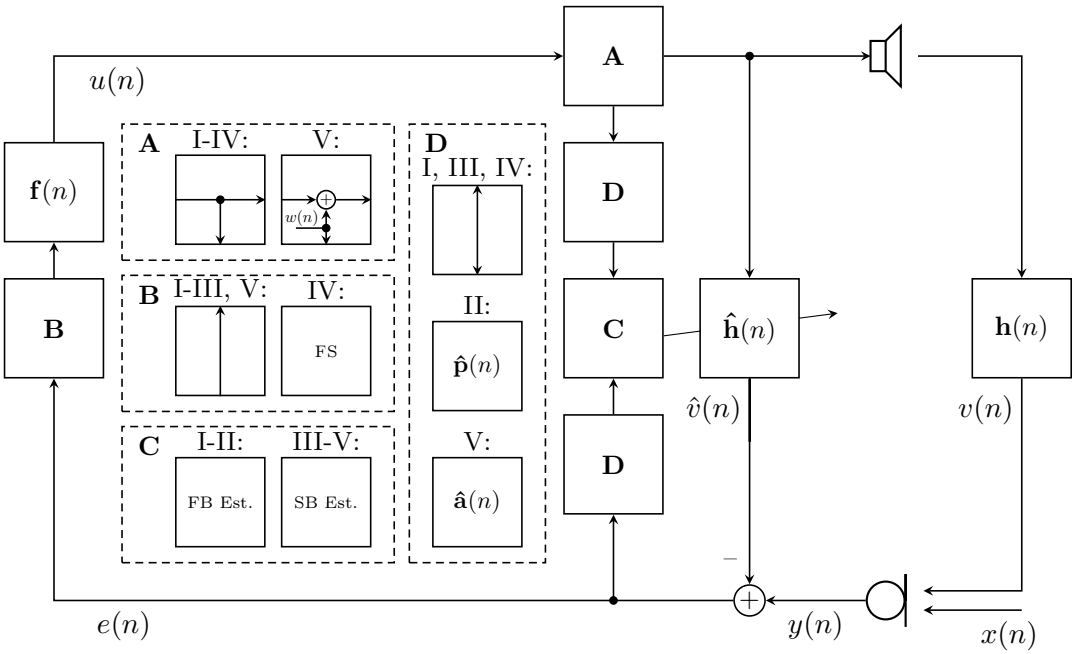


Fig. 2: An overview of the AFC systems I-V. For simplicity, we only consider a hearing aid system with a single microphone. For each system I-V, the blocks **A**-**D** in the block diagram should be accordingly substituted by their replacement blocks given in the dashed boxes **A**-**D**.

all five AFC systems. For each system I-V, the blocks **A**-**D** in the block diagram should be accordingly substituted by their replacement blocks given in the dashed boxes **A**-**D**.

2.1 System I: F-AFC System

In the F-AFC system [2], the adaptive filter $\hat{\mathbf{h}}(n)$ is estimated to minimize the mean square error $E[e^2(n)]$. An NLMS algorithm is used to estimate $\hat{\mathbf{h}}(n)$. As we will demonstrate in Sec. 4, one of the major limitations in this system is the biased estimation of $\hat{\mathbf{h}}(n)$ whenever $u(n)$ is correlated with $x(n)$.

2.2 System II: PEM-AFC System

The PEM-AFC systems [9, 10] assume that the incoming signal $x(n)$ can be modeled well as a white noise sequence filtered through an autoregressive model $1/\sum_{k=0}^{L_p-1} p_k(n)z^{-k}$, where z^{-1} is the unit delay operator. Let $\mathbf{p}(n) = [p_0(n), p_1(n), \ldots, p_{L_p-1}(n)]^T$, the

prefilter $\hat{\mathbf{p}}(n) = [\hat{p}_0(n), \hat{p}_1(n), \ldots, \hat{p}_{L_p-1}(n)]^T$ and cancellation filter $\hat{\mathbf{h}}(n)$ are then jointly estimated to minimize the mean square error $E[e_p^2(n)]$, where $e_p(n) = \sum_{k=0}^{L_p-1} \hat{p}_k(n)e(n-k)$ is the filtered signal of $e(n)$.

For an ideal estimation, $\hat{\mathbf{p}}(n) = \mathbf{p}(n)$, the filtered error signal $e_p(n)$ only contains the filtered signal $v_p(n)$ of $v(n)$ and samples of the white noise excitation sequence, which are uncorrelated with the loudspeaker signal $u(n)$ and the filtered signal $u_p(n)$ of $u(n)$, respectively. In this way, the PEM-AFC ideally provides an unbiased estimation of $\hat{\mathbf{h}}(n)$ for autoregressive incoming signals $x(n)$.

In different versions of the PEM-AFC system, $x(n)$ is modeled in different ways, and various schemes to estimate $\hat{\mathbf{p}}(n)$ and $\hat{\mathbf{h}}(n)$ have been proposed. For more details on the PEM-AFC system, we refer to [4, 9, 10].

2.3 System III: S-AFC System

The S-AFC system facilitates adaptive filter estimation in filter bank subbands. We employ the delayless subband structure proposed in [18] and refined in [19]. The greatest advantage by using this structure in contrast to other subband adaptive filter structures e.g. [20] is that no additional delay is introduced in the signal path from $y(n)$ to $u(n)$. In particular, uniform filter banks are used to divide the fullband error signal $e(n)$ and loudspeaker signal $u(n)$ to subband signals $E(m,k)$ and $U(m,k)$ with the subband frequency index m and time index k, respectively. The adaptive filter estimation of the frequency response $H(m,k)$ is then performed in the subband domain using e.g. a low-order NLMS algorithm based on $E(m,k)$ and $U(m,k)$. The fullband adaptive filter estimate $\hat{\mathbf{h}}(n)$ is obtained by the inverse discrete Fourier transform of the frequency response estimate $\hat{H}(m,k)$.

The subband structure also makes it easy to use different parameter settings across frequency regions, e.g. the subband step size parameter. For hearing aid applications, AFC systems are often only necessary for frequencies higher than above approximately 1 kHz [9]. On the other hand, the signal correlation between $u(n)$ and $x(n)$ is often most dominant at low frequency regions, e.g. for speech signals. Therefore, having AFC systems running in the lowest frequencies might introduce artifacts due to the biased estimation problem rather than solving a small acoustic feedback problem. In this subband structure, AFC systems at lower frequencies can easily be turned off by setting the corresponding subband step sizes to zero.

2.4 System IV: FS-AFC System

The FS-AFC system utilizes the same subband AFC system as the S-AFC system. In addition, a frequency shifting [17] is applied before the hearing aid processing unit $\mathbf{f}(n)$. It is known that frequency shifting decorrelates the loudspeaker signal $u(n)$ from the incoming signal $x(n)$ and can thereby reduce the biased estimation problem [11, 12, 14].

2.5 System V: PN-AFC System

The PN-AFC system also makes use of the subband AFC system, and it is described in detail in [13]. Instead of modifying the loudspeaker signal $u(n)$ via frequency shifting, this system obtains an unbiased estimation of adaptive filter $\hat{\mathbf{h}}(n)$ by adding a generally non-stationary probe noise signal $w(n)$, which is uncorrelated with $x(n)$, to the loudspeaker signal $u(n)$. The probe noise signal $w(n)$ is generated using a masking model to ensure that $w(n)$ is inaudible in the presence of the loudspeaker signal $u(n)$.

Furthermore, enhancement filters $\hat{\mathbf{a}}(n)$ are used for filtering the probe noise signal $w(n)$ and the error signal $e(n)$ before they entry the adaptive estimation algorithm. The enhancement filters have a specific structure which allows them to improve the probe noise to disturbing signal power ratio and then improve the estimation of $\hat{\mathbf{h}}(n)$. We refer to [13] for more details.

3 Sound Quality Evaluation of Decorrelation Methods

Among the five AFC systems, the FS-AFC and PN-AFC systems might introduce additional audible artifacts because they modify the loudspeaker signal $u(n)$ to decorrelate it from $x(n)$. More specifically, for the FS-AFC system, the loudspeaker signal $u(n)$ is frequency shifted, whereas the PN-AFC system adds a probe noise signal $w(n)$ to the loudspeaker signal $u(n)$. These modifications of the loudspeaker signal $u(n)$ might be perceived as a sound quality degradation. Depending on the exact system parameter settings, frequency shifting might introduce a vibrato-like sound degradation for both speech and music signals, whereas probe noise might add an audible hiss to the original signal. Therefore, it is important to understand how these modifications of the loudspeaker signal $u(n)$ affect the sound quality. Ideally, only minor and nonirritating audible artifacts should be allowed to preserve sound quality. Furthermore, the degradations should ideally be at similar levels for making the AFC performance comparison in Sec. 4 more straightforward. To achieve these, we perform a sound quality evaluation.

3.1 Evaluation Method

We carry out a listening test to evaluate the sound quality degradation due to the decorrelation in terms of frequency shifting and probe noise injection in the FS-AFC and PN-AFC systems, respectively. We initially adjust system parameters in both decorrelation methods so that they are effective for decorrelation, and the sound quality degradations are in the range of intermediate to small (verified by informal listening tests). Moreover, we want to evaluate the sound quality for both speech and music signals. To achieve this, we use the "MUlti Stimulus test with Hidden Reference and Anchor (MUSHRA)" specified in [21] with one reference signal, five test signals including

Table 3: Sound signals used for the MUSHRA.

Speech 1	Danish Female Reading
Speech 2	English Male Reading
Music 1	Piano + Female Singing
Music 2	Female Singing
Music 3	Singing Female Voice

a hidden reference, a hidden anchor, two test signals processed with frequency shifting variants, and one test signal processed with probe noise injection.

Furthermore, we choose to evaluate the sound quality using normal hearing test subjects. The underlying hypothesis is that if the sound quality degradation is not significant nor annoying for normal hearing persons, then it is also not annoying for hearing impairment persons, and it would work for all hearing aid users in contrast to obtaining an acceptance level of artifacts based on a specific group of test subjects with hearing impairments.

3.2 Processing of Test Signals

We focus on the artifacts of the frequency shifting and probe noise injection, without any interaction with AFC systems; it makes the interpretation of this sound quality evaluation results more straightforward. Moving back to Fig. 2, it corresponds to removing the feedback path $\mathbf{h}(n)$ and its estimate $\hat{\mathbf{h}}(n)$, i.e. $\mathbf{h}(n) = \mathbf{0}$ and $\hat{\mathbf{h}}(n) = \mathbf{0}$.

The reference signal in the MUSHRA is created by passing an unprocessed high quality sound signal through the hearing aid processing with only an amplification and delay. Test signals are created by passing the same high quality sound signal through the hearing aid processing with the same amplification and delay, and an additional frequency shifting or probe noise injection; in this way, the hearing aid loudspeaker signals $u(n)$ are presented to test subjects in this listening test[1]. The anchor signal is created in a similar way as the test signals, but with both frequency shifting and probe noise injection. Table 3 shows the five sound signals chosen for the MUSHRA. Details on the processing conditions are given in Table 4.

[1]We considered to present instead a mixture of the loudspeaker signal $u(n)$ and the incoming signal $x(n)$ because this would be closer to reality for some hearing aid styles [22]. However, with the mixed signal, an undesired comb filter effect is probable to occur [22]. The comb filter effect makes the listening test more complicated, since test subjects have to assess and judge one type artifact (comb filter effect) found in the reference signal from other types of artifacts (comb filter effect + frequency shifting or probe noise) found in the test signals. A potential risk of doing such a test is that test subjects would non-deliberately pay more attention to the comb filter effect instead of artifacts from frequency shifting and probe noise. Therefore, to avoid this, we decided to present the loudspeaker signals $u(n)$ in the MUSHRA.

Table 4: Processing of test signals for each test sound. All test signals are processed with identical amplification and delay.

Test Signal	**Description**
Reference	Amplification + Delay.
FS-500	Amplification + Delay + Frequency shifting of 10 Hz above 500 Hz.
FS-1000	Amplification + Delay + Frequency shifting of 10 Hz above 1000 Hz.
PN-500	Amplification + Delay + Probe noise above 500 Hz with noise level computed based on a spectral masking threshold $M(m,k)$ for the mth frequency at time index k, see [13] for more details.
Anchor	Amplification + Delay + Frequency shifting of 25 Hz above 500 Hz and probe noise injection above 500 Hz with noise level $2 \cdot M(m,k)$.

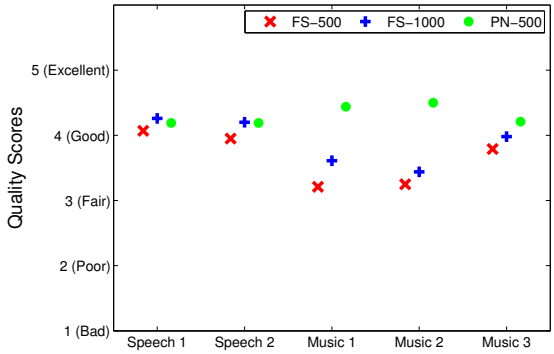


Fig. 3: Initial verifications of test signals for the MUSHRA, using respectively PESQ/PEAQ scores for speech/music signals.

To ensure that the chosen parameters for frequency shifting and probe noise injection given in Table 4 do not introduce significant sound quality degradations, we initially verified these parameter choices via informal listening test and using the objective evaluation measures PESQ [23] and PEAQ [24] for respectively speech and music signals (although these objective evaluation measures were not specifically designed for evaluating the artifacts of either frequency shifting nor probe noise). Fig. 3 shows the PESQ/PEAQ prediction results.

The PESQ/PEAQ predictions were in line with the observations made via the initial informal listening, and they indicate that the parameter choices actually provide nonsignificant sound quality degradations, and the minor sound quality degradations from both methods are at a similar level, which makes the AFC performance evaluation results in Sec. 4 more directly comparable.

3.3 Training and Test Procedure

In subjective tests including the MUSHRA, training of test subjects is important to obtaining reliable results. Preliminary tests and experiences gathered through the development of the FS-AFC system show that the training for this listening test is difficult.

Specifically, we noticed that the training period in the frequency shifting case might be as long as days or weeks of listening, before the artifacts become clearly noticeable for even experienced test subjects. However, once these artifacts become noticeable, they appear to be very annoying for some test subjects. On the other hand, we did not notice that test subjects became more annoyed with probe noise injection over time.

For practical reasons in this work, we were not able to conduct a preparation over weeks for a large group of test subjects. Therefore, we decided to have two training sessions prior to the actual test session, where the artifacts of frequency shifting and probe noise are exaggerated in the first training session in the hope that this would shorten the training period. The three test signals are processed by frequency shifting of 20 Hz and probe noise threshold $1.5 \cdot M(m,k)$ in contrast to the nominal 10 Hz and $M(m,k)$ as given in Table 4, which are used for the second training session and the test session. In this way, the test consists of three sessions: two training sessions and a test session.

3.4 Listening Test

The listening test is conducted during a period of three days in a quiet room, and a total of 16 normal-hearing test subjects participated in the test. The test signals are presented diotically via a headphone (Sennheiser EH2270) connected to a computer, which runs the MUSHRA interface, and the sound level is adjusted individually by the subjects. The rating of each test signal is recorded by the computer.

We instructed test subjects to find the hidden reference and anchor for each test trial, and rate them with the quality scores 100 and 0, respectively. The remaining test signals should be rated with respect to the reference and anchor. We did not inform test subjects which kind of artifacts nor the processing algorithms they were listening to. For most subjects, the entire listening test took about $20 - 40$ min. including the initial instruction, two training sessions, the test session, and a short debriefing at the end of the test.

3.5 Test Results

Only the results from the test session are used for determining test statistics. The statistics determined from all subjects are shown in Fig. 4. Since all test subjects were able to correctly find the hidden reference and anchor, the results of these are not presented. Seven test subjects were already familiar with both type of artifacts before

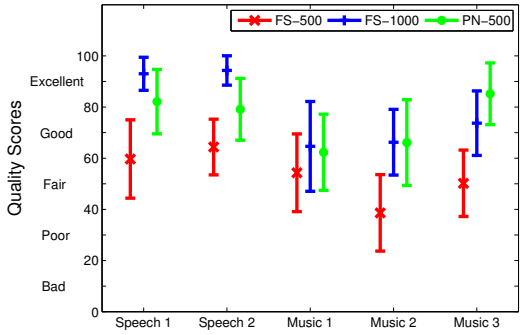


Fig. 4: Mean and 95%-confidence interval of test results for all test subjects (sample size = 16).

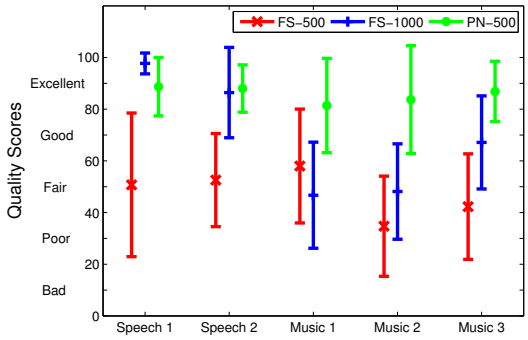


Fig. 5: Mean and 95%-confidence interval of test results for expert test subjects (sample size = 7).

the listening test, and can therefore be considered as expert listeners for this test. Fig. 5 shows the statistics of test results from this expert panel.

3.6 Discussions

From Figs. 4 and 5 we conclude that the frequency shifting of 10 Hz with a cut-off frequency at 1000 Hz (FS-1000) and probe noise with a cut-off frequency at 500 Hz (PN-500) introduce relatively small degradations, especially for speech signals. Moreover, we conclude that the frequency shifting with a cut-off frequency at 500 Hz (FS-500) generally produces signals with sound quality in the range of poor to fair. Hence, we consider the FS-500 processing to degrade sound quality too much compared to the FS-1000 and PN-500 processing, and it is thereby not recommended for practical use and we exclude it from the following discussion.

The PN-500 processing seems to have a higher degradation level for speech signals compared to the FS-1000, although it can be shown that the difference is not statistically significant based on a paired t-test with a 5% level of significance. Nevertheless, the sound quality of speech signals is still perceived as almost excellent for the PN-500. This trend is just opposite for music signals, where the PN-500 seems to provide a slightly better sound quality than the FS-1000. However, it should be noted that the PN-500 has a decorrelation effect in the frequency region of $500 - 1000$ Hz, which is not the case for the FS-1000, and better AFC performance can thereby be expected for the PN-500.

Interestingly, the expert test subjects rated the test signals differently than the general test subjects. By comparing Figs. 4 and 5, it is clear that the expert subjects generally rated the frequency shifting processed test signals (FS-1000) lower, and the probe noise processed test signals (PN-500) higher compared to the general test subjects, although it is not statistically significant due to the relatively small sample size in the test. This can be explained by the hypothesis that a long training period is necessary for test subjects to be aware of the frequency shifting artifact, and they tend to dislike it once noticing it. Fig. 5 suggests that the sound quality degradation with FS-1000 is somewhat higher than with PN-500, especially for music signals, and that only two training sessions are not adequate for general test subjects. Moreover, these results shown in Fig. 5 can also be considered as upper-bounds for the sound quality that can be achieved in an ideal AFC system when using these decorrelation methods, that is when $\hat{\mathbf{h}}(n) = \mathbf{h}(n)$.

Furthermore, test subjects typically used words such as "vibration", "ringing sounds", "modulation", "metallic", and "additional tones" to describe the frequency shifting artifacts; whereas for the probe noise artifacts, the most common describing words are "noisy" and "hiss". Test subjects also suggested that the level of sound quality degradation depends on test signals, and speech signals generally have higher quality. Most test subjects reported that they get more annoyed by the modulation artifacts from the frequency shifting, because they do not "fit into" the sound, whereas the noise from the probe noise processing is better "hidden" in the original signals.

Finally, a comparison of Figs. 3 and 5 shows that there is a great similarity between the sound quality prediction using PESQ/PEAQ and the subjective evaluated sound quality using the MUSHRA with expert subjects. Similar trends are found in both the PESQ/PEAQ measures and the MUSHRA results, e.g. that the speech signals have generally higher sound quality scores than music signals, and the PN-500 provides a similar if not better overall sound quality compared to the FS-1000 processing. The absolute values of sound quality scores might not be directly comparable between these two evaluations, since the sound quality scores in the MUSHRA depend on the anchor signal; increasing the quality of the anchor signal would somewhat lower the sound quality perception for the test signals and vice versa. However, it is still clear that the PESQ/PEAQ measures seem to provide an accurate relative sound quality prediction.

In addition to the sound quality consideration, another important concern of these

two decorrelation methods regards the preservation of the localization cues for binaural reproductive systems upon modifications of loudspeaker signals. This is a topic for future work.

4 AFC Performance Evaluation

In the previous section, evaluations are carried out to ensure that the signal distortions due to decorrelation are insignificant and at similar levels for the FS-AFC and PN-AFC systems. In this section, we conduct simulation experiments for comparing the ability to cancel feedback in all five AFC systems.

4.1 Objective Performance Measures

To quantify objectively the cancellation performance, we use two performance measures to be described in the following. The difference between the feedback path impulse response $\mathbf{h}(n)$ and its estimate $\hat{\mathbf{h}}(n)$ is directly linked to system stability via the open-loop transfer function $\Theta(\omega, n)$. For the general system (ignoring the frequency shifting) shown in Fig. 2, $\Theta(\omega, n)$ is expressed by

$$\Theta(\omega, n) = F(\omega, n) \left(H(\omega, n) - \hat{H}(\omega, n) \right), \tag{1}$$

where $F(\omega, n)$ is the frequency response of the hearing aid forward path impulse response $\mathbf{f}(n)$, $H(\omega, n)$ and $\hat{H}(\omega, n)$ are frequency responses of $\mathbf{h}(n)$ and $\hat{\mathbf{h}}(n)$, respectively. The open-loop transfer function $\Theta(\omega, n)$ determines system stability according to the Nyquist stability criterion [25], which states that a linear and time-invariant closed-loop system becomes unstable whenever the following two criteria are both fulfilled:

$$1.\ |\Theta(\omega, n)| \geq 1; \tag{2}$$

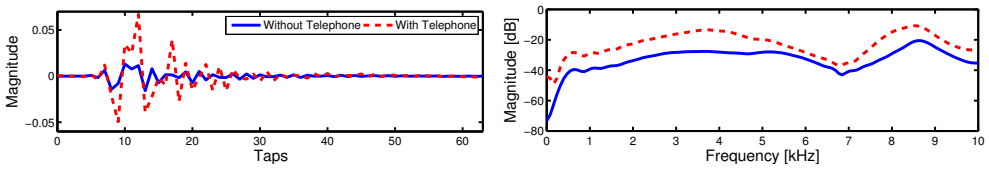
$$2.\ \angle\Theta(\omega, n) = l2\pi,\ l = \mathbb{Z}. \tag{3}$$

Eq. (2) forms the basis for different distance measures between $\mathbf{h}(n)$ and $\hat{\mathbf{h}}(n)$ [1]. A direct measure of the convergence rate, tracking error, and the steady-state error of the adaptive filter $\hat{\mathbf{h}}(n)$ is the frequency domain coefficient misalignment measure $\varepsilon(\omega, n)$ defined as

$$\varepsilon(\omega, n) = \left| H(\omega, n) - \hat{H}(\omega, n) \right|. \tag{4}$$

Focusing on the most critical value of $\varepsilon(\omega, n)$ in Eq. (4) over frequencies, we define a simplified measure referred to as the maximum coefficient misalignment (MCM) $\varepsilon_{\mathcal{F}}(n)$; it is determined by the maximum of $\varepsilon(\omega, n)$ across a frequency region denoted by $\mathcal{F}$, as

$$\varepsilon_{\mathcal{F}}(n)\ [\text{dB}] = 20 \log_{10} \max_{\omega \in \mathcal{F}} \varepsilon(\omega, n). \tag{5}$$



(a) (b)

Fig. 6: Measured acoustic feedback paths with and without a telephone placed next to the ear of a hearing aid user. (a) Impulse Response. (b) Magnitude Response.

Furthermore, assuming that Eq. (3) is fulfilled for all frequencies, a conservative maximum fullband ($\omega \in [0, \pi]$) gain to ensure system stability referred to as the maximum stable gain (MSG) $\mathcal{M}(n)$ is determined as

$$\mathcal{M}(n) \text{ [dB]} = -20 \log_{10} \max_{\omega} \varepsilon(\omega, n). \tag{6}$$

In this work, we use $\varepsilon_{\mathcal{F}}(n)$ and $\mathcal{M}(n)$ for the objective performance evaluation.

4.2 Test Setups and Signals

4.2.1 Acoustic Feedback Path

The acoustic feedback paths $\mathbf{h}(n)$ used in the simulation experiments are obtained from measurements on hearing aids. During simulations, the feedback path $\mathbf{h}(n)$ remains fixed for most of the time. However, halfway through each simulation run, a feedback path change is simulated by a complete and momentary change of $\mathbf{h}(n)$. This change models a situation, where a hearing aid user places a telephone close to the ear and thereby the hearing aid, and it is known to be a very challenging situation for hearing aid AFC systems [26].

Fig. 6 shows an example of measured acoustic feedback paths $\mathbf{h}(n)$ before and after a telephone is placed next to the ear of a hearing aid user. The sampling rate is 20 kHz for the measurement. Placing the telephone next to the ear increases the feedback path magnitude response by 15 dB at certain frequencies.

4.2.2 Forward Path

For simplicity, the forward path $\mathbf{f}(n)$ only consists of an amplification and delay for all systems. A single-channel compressor [22] is used to provide a fullband, time-varying, and signal dependent amplification. The maximum amplification of the compressor is limited so that the magnitude of the open-loop transfer function $|\Theta(\omega, n)|$ is -1 dB without acoustic feedback cancellation, i.e. when $\hat{\mathbf{h}}(n) = \mathbf{0}$.

Table 5: Test signals used in simulation experiments.

Speech	Music
Danish Female Speech	Classic
Japanese Female Speech	Female Singing
Norwegian Male Speech	Violin + Female Singing
English Female Speech 1	Flute
English Female Speech 2	Organ
English Male Speech	Piano + Female Singing

4.2.3 Test Signals

We use speech and music signals as test signals $x(n)$ in the simulation experiments. The duration of each test signal is 60 s, and the sampling rate is 20 kHz.

In this work, we choose speech signals spoken by both male and female speakers and in different languages. In contrast to speech, for which the autocorrelation time is often several tens of ms, music signals might have correlation time of several hundreds or even thousands of ms, and it could easily cause significantly biased estimation of $\hat{\mathbf{h}}(n)$. Therefore, we choose some music signals with clear sustaining tonal components for relatively long time periods. With these test signals and the already introduced acoustic feedback path change, we determine cancellation performance of different AFC systems in a very demanding but realistic situation. Table 5 summarizes the test signals used in the simulations.

4.2.4 AFC Systems

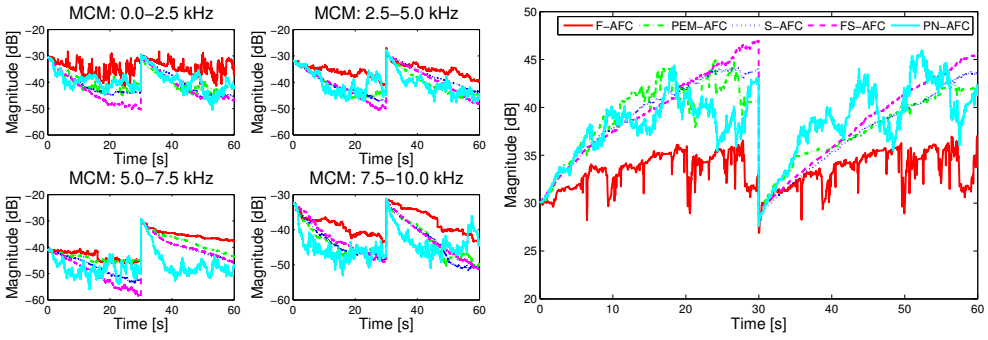
For all AFC systems, the length of $\hat{\mathbf{h}}(n)$ is 64 at a sampling rate of 20 kHz. Furthermore, we choose the parameters of the various systems so that a similar and relatively fast convergence is achieved for all systems for speech signals.

For the F-AFC system, an NLMS algorithm with a step size of 2^{-9} and a regularization parameter of 2^{-12} is used to estimate $\hat{\mathbf{h}}(n)$.

For the PEM-AFC system, we use the Levinson-Durbin recursion [27] to compute the prefilter $\hat{\mathbf{p}}(n)$ of order 20 based on the error signal $e(n)$, whereas we use the same NLMS algorithm and parameter settings for the estimation of $\hat{\mathbf{h}}(n)$ as the F-AFC system.

For the S-AFC system, uniform filter banks with 32 complex conjugated subbands and a decimation factor of 16 divide fullband signals $e(n)$ and $u(n)$ into subbands for the subband estimation. The subband NLMS step size is chosen to be 2^{-12} for all subbands except the lowest one (DC band), which is set to zero, so that the AFC is not performed under approximately 500 Hz. The subband NLMS regularization parameter is 2^{-16}.

For the FS-AFC system, we use the same S-AFC system to estimate $\hat{\mathbf{h}}(n)$. In addition, a frequency shifting of 10 Hz is performed in the forward path $\mathbf{f}(n)$ for frequencies



(a) (b)

Fig. 7: Results for an example speech signal. (a) Maximum coefficient misalignment $\varepsilon_{\mathcal{F}}(n)$. (b) Maximum stable gain $\mathcal{M}(n)$. See Fig. 7(b) for legend.

above 1000 Hz, because performing frequency shifting at the lowest frequencies would not be beneficial since the AFC system is already turned off, and it could significantly degrade sound quality, as shown in Sec. 3.

For the PN-AFC system, we again use the S-AFC system to estimate $\hat{\mathbf{h}}(n)$. Furthermore, the probe noise signal is generated using a masking model, and the probe noise enhancement filter $\mathbf{a}(n)$ is estimated using the same subband adaptive filter structure for estimating $\hat{\mathbf{h}}(n)$, with the subband step size of 2^{-2} for all subbands.

4.3 Test Results and Discussions

In this section, we evaluate the test results in terms of the maximum coefficient misalignment $\varepsilon_{\mathcal{F}}(n)$ and the maximum stable gain $\mathcal{M}(n)$. We consider $\varepsilon_{\mathcal{F}}(n)$ in four equally divided frequency regions from 0 to 10 kHz.

4.3.1 Speech Signals

Fig. 7 shows test results for an example speech signal in terms of $\varepsilon_{\mathcal{F}}(n)$ and $\mathcal{M}(n)$. As expected, $\varepsilon_{\mathcal{F}}(n)$ decreases and $\mathcal{M}(n)$ increases over time as the AFC systems adaptively improve the feedback path estimate; it continues until the feedback path undergoes a momentary change after 30 s. The result of this is a mismatch between the true and estimated feedback paths, which is seen as the jump/drop in $\varepsilon_{\mathcal{F}}(n)$ and $\mathcal{M}(n)$, respectively.

We observed that for this example speech signal, the AFC cancellation performance in terms of $\varepsilon_{\mathcal{F}}(n)$ and $\mathcal{M}(n)$ are very similar for almost all AFC systems except the F-AFC, which performs poorly at the lower frequencies $0-2.5$ kHz, where the correlation

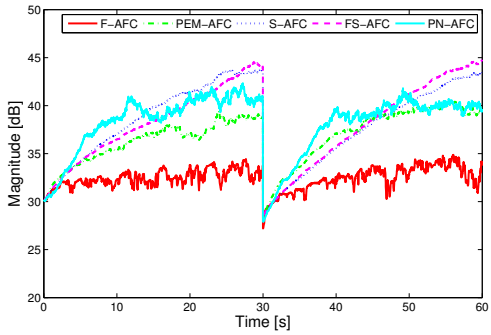


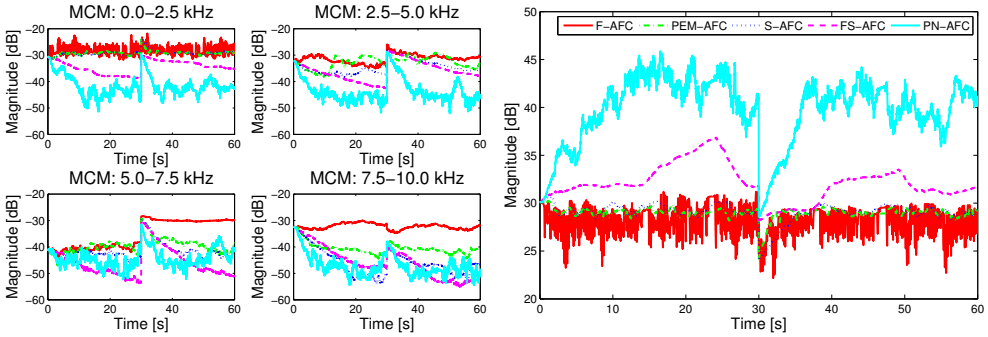
Fig. 8: Average maximum stable gain for speech signals.

in speech signals is generally strong. The consequence of this is that a relatively small $\mathcal{M}(n)$ less than 35 dB (as an additional gain less than 5 dB added to the initial value of $\mathcal{M}(0) \approx 30$ dB) will be available in the system, as seen in Fig. 7(b), even though the desired amplification in hearing aids can be more than 50 dB [22]. In this example case, the improvement of the maximum stable gains can be more than 10 dB in the other AFC systems. Furthermore, Fig. 8 shows the averaged maximum stable gains obtained from simulations of all speech signals. Again, it is seen that the F-AFC performs much worse than the other AFC systems, which have similar behaviors.

The PEM-AFC system works generally well in this case, since the assumption of autoregressive incoming signals $x(n)$ is reasonably valid especially for the unvoiced regions of speech signals [28]. The S-AFC system transforms the fullband signals $u(n)$ and $e(n)$ to subbands, this transform generally decorrelates the signals [1], and an improvement is thereby obtained. In the FS-AFC and PN-AFC systems, the frequency shifting and probe noise are applied in addition to the subband transform, it seems that any further improvement is very limited compared to the S-AFC system. Furthermore, the PEM-AFC and PN-AFC seem to have a slightly faster convergence but lower maximum stable gain in the steady-state with the chosen simulation parameters.

4.3.2 Music Signals

Fig. 9 shows similar simulation results for an example music signal. It is very clear that the F-AFC does not work properly, due to the much stronger correlation, over the entire frequency spectrum, in the music signal. Interestingly, performance of PEM-AFC and S-AFC is not much better than F-AFC in terms of MSG $\mathcal{M}(n)$ and MCM $\varepsilon_{\mathcal{F}}(n)$ below 5 kHz. The PEM-AFC system is not performing well due to the assumption of the incoming signals $x(n)$ to be autoregressive is violated in the music signal case, and the prefilters in the PEM-AFC system can thereby not decorrelate the loudspeaker signal



(a) (b)

Fig. 9: Results for an example music signal. (a) Maximum coefficient misalignment $\varepsilon_{\mathcal{F}}(n)$. (b) Maximum stable gain $\mathcal{M}(n)$. See Fig. 9(b) for legend.

$u(n)$ from the incoming signal $x(n)$. The subband transform in the S-AFC system still has a decorrelation effect, but it is no longer adequate.

Furthermore, it is clear that the FS-AFC and PN-AFC systems perform significantly better than other systems in this case. Fig. 9(a) shows that the FS-AFC system is very effective to provide unbiased estimation at higher frequencies above 5 kHz, but it has only limited effects in lower frequencies, due to the stronger correlation in signals. On the other hand, the PN-AFC is able to provide unbiased estimation in the entire frequency range, and it provides the highest maximum stable gain as seen in Fig. 9(b).

The same trend is observed in Fig. 10, which shows the averaged maximum stable gain obtained from all simulations of music signals. It is clear that the F-AFC system fails to increase the maximum stable gain over time, the PEM-AFC and S-AFC can only provide minimal increments, but the FS-AFC improves it by more than 6 dB, and the PN-AFC improves it by more than 12 dB.

4.3.3 Speech Versus Music Signals

To complete the observations, we computed the slopes of the average maximum stable gain curves based on the first 5 seconds of all simulation results. Table 6 shows the computed slopes. These data demonstrate that all AFC systems except the F-AFC perform well for speech signals, but only the FS-AFC and especially the PN-AFC are robust against music signals.

Furthermore, an interesting and somewhat unexpected observation is that the PEM-AFC has similar performance as the S-AFC system for speech signals, whereas the S-AFC system is slightly more robust against correlation in music signals (compared to this particular version of the PEM-AFC). A similar conclusion is found in a comparison

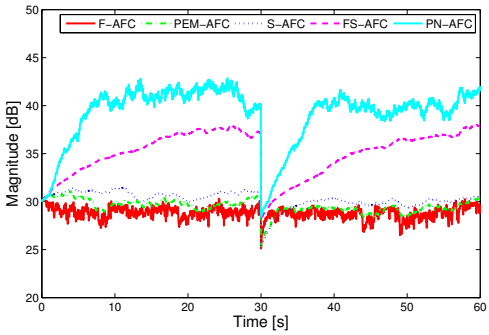


Fig. 10: Average maximum stable gain for music signals.

Table 6: Slopes [dB/s] of average maximum stable gain curves.

System	Speech	Music
F-AFC	0.41	-0.25
PEM-AFC	0.75	0.11
S-AFC	0.73	0.22
FS-AFC	0.72	0.54
PN-AFC	0.83	1.23

between an DCT domain AFC system and a PEM based AFC system in a recent work [29].

5 Computational Complexity Evaluation

In addition to the sound quality and AFC performance improvement evaluations, another important consideration is the computational complexity and memory usage for each AFC system. In this section, we make a rough complexity estimation of each system by counting the number of required real multiplications; we do not take into account that specific optimizations and modifications of algorithms could reduce the complexity significantly for a particular system. Hence, our estimates should be considered as upper-bounds for the complexity.

We estimated the complexity for each AFC system in a two microphone channel hearing aid system. The exact number of multiplications are computed based on the chosen parameters for simulations in this work. For convenience, we normalized the results relative to the F-AFC reference system. Furthermore, we divided each AFC system into functional subgroups and computed the complexity in percentage for each

Table 7: Complexity factors (CF) for different AFC systems.

System	CF	Subgroup Complexity	
F-AFC	1	NLMS update	67%
		Filtering	33%
PEM-AFC	1.27	Prefilter estimation	9%
		Prefilter filtering	12%
		F-AFC System	79%
S-AFC	1.55	Subband transform	12%
		Subband NLMS update	28%
		Time-domain filter	38%
		Filtering	22%
FS-AFC	1.68	Frequency shifting	8%
		S-AFC System	92%
PN-AFC	2.78	Probe noise generation	8%
		Probe noise enhancement	36%
		S-AFC System	56%

subgroup. The results are given in Table 7.

The F-AFC system has the lowest computational complexity, the adaptive filter estimation is done cheaply using the NLMS algorithm. The PEM-AFC needs an additional prefilter estimation and the subsequent filtering on error and reference signals for the adaptive filter estimation. The S-AFC system is more complex due to the need of filter banks to divide fullband signals into subbands, where a subband NLMS algorithm works on complex signals rather than real signals in the F-AFC system, and the frequency domain feedback path estimate has to be transformed back to a time-domain filter. The FS-AFC needs an additional frequency shifting algorithm compared to the S-AFC system. Finally, the PN-AFC system needs probe noise generation, probe noise enhancement including filtering of error and reference signals similar to the PEM-AFC system, in addition to the subband adaptive filter estimation as in the S-AFC system.

Table 7 shows, perhaps as expected, that the systems with best AFC performance as demonstrated in Sec. 4 also have higher computational complexity.

6 Conclusion

We conducted an evaluation of several state-of-the-art acoustic feedback cancellation systems for hearing aids in terms of the cancellation performance, sound quality degradation, and computational complexity. In particular, we compared a traditional fullband system to a prediction error method based fullband system, a subband system, a subband system with frequency shifting, and a recently proposed subband system with a

novel probe noise usage. By allowing a perceptually noticeable but small sound quality degradation in the loudspeaker signal to decorrelate it from the incoming signal, we evaluated the cancellation performance in terms of maximum coefficient misalignment of the adaptive filters and the maximum stable gain in hearing aid simulations. All systems outperformed the traditional fullband system in cancellation performance. Especially the subband system with probe noise provided an improvement of more than 12 dB in maximum stable gain in the most difficult situations, but it is also the most computationally complex system, roughly 2.8 times more complex than the traditional system. The subband system with frequency shifting improved the maximum stable gain by more than 6 dB with a complexity increment by a factor of 1.7 for the same situations. Furthermore, we showed that the subband system had a slightly larger improvement than the prediction error method based fullband system, which had the lowest computational complexity increment by a factor of 1.3 compared to the traditional system. In this way, the systems providing largest improvements are also more computationally complex. Hence, choosing an appropriate system for a practical application, among these evaluated, is a clear compromise between performance and computational cost. However, for a price of $2-3$ times the reference complexity, a cancellation system can be realized which is robust in even the most challenging feedback situations. With the increasing computational power available in hearing aids, these improved cancellation systems can be realistically implemented in the near future.

References

[1] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, US: Prentice Hall, Sep. 2001.

[2] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

[3] M. G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 4, pp. 443–453, Jul. 2000.

[4] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Elsevier J. Franklin Inst.*, vol. 343, no. 6, pp. 545–573, Sep. 2006.

[5] J. M. Kates, "Feedback cancellation in hearing aids: Results from a computer simulation," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 553–562, Mar. 1991.

[6] J. A. Maxwell and P. M. Zurek, "Reducing acoustic feedback in hearing aids," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 4, pp. 304–313, Jul. 1995.

[7] P. M. J. Van Den Hof and R. J. P. Schrama, "An indirect method for transfer function estimation from closed loop data," *Elsevier Automatica*, vol. 29, no. 6, pp. 1523–1527, Nov. 1993.

[8] N. A. Shusina and B. Rafaely, "Unbiased adaptive feedback cancellation in hearing aids by closed-loop identification," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 658–665, Mar. 2006.

[9] J. Hellgren, "Analysis of feedback cancellation in hearing aids with filtered-X LMS and the direct method of closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 2, pp. 119–131, Feb. 2002.

[10] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.

[11] H. A. L. Joson, F. Asano, Y. Suzuki, and T. Sone, "Adaptive feedback cancellation with frequency compression for hearing aids," *J. Acoust. Soc. Am.*, vol. 94, no. 6, pp. 3248–3254, Dec. 1993.

[12] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "On the use of a phase modulation method for decorrelation in acoustic feedback cancellation," in *Proc. 20th European Signal Process. Conf.*, Aug. 2012, pp. 2000–2004.

[13] M. Guo, S. H. Jensen, and J. Jensen, "Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancements," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2549–2563, Nov. 2012.

[14] T. van Waterschoot and M. Moonen, "Assessing the acoustic feedback control performance of adaptive feedback cancellation in sound reinforcement systems," in *Proc. 17th European Signal Process. Conf.*, Aug. 2009, pp. 1997–2001.

[15] A. Spriet, M. Moonen, and J. Wouters, "Evaluation of feedback reduction techniques in hearing aids based on physical performance measures," *J. Acoust. Soc. Am.*, vol. 128, no. 3, pp. 1245–1261, Sep. 2010.

[16] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation.* Berlin, Heidelberg, Germany: Springer, May 2001.

[17] S. Wardle, "A Hilbert-transform frequency shifter for audio," in *Proc. Digit. Audio Effects Workshop*, Nov. 1998, pp. 25–29.

[18] D. R. Morgan and J. C. Thi, "A delayless subband adaptive filter architecture," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1819–1830, Aug. 1995.

[19] J. Huo, S. Nordholm, and Z. Zang, "New weight transform schemes for delayless subband adaptive filtering," in *Proc. IEEE Global Telecommunications Conf.*, vol. 1, Nov. 2001, pp. 197–201.

[20] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

[21] ITU-R Recommendation BS.1534-1, *Method for subjective assessment of intermediate quality level of coding systems*, Int. Telecommun. Union, 2001.

[22] H. Dillon, *Hearing Aids*. Stuttgart, Germany: Thieme, May 2001.

[23] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, Int. Telecommun. Union, 2001.

[24] ITU-R Recommendation BS.1387-1, *Method for objective measurements of perceived audio quality*, Int. Telecommun. Union, 1998.

[25] H. Nyquist, "Regeneration theory," *Bell System Tech. J.*, vol. 11, pp. 126–147, 1932.

[26] J. Hellgren, T. Lunner, and S. Arlinger, "Variations in the feedback of hearing aids," *J. Acoust. Soc. Am.*, vol. 106, no. 5, pp. 2821–2833, Nov. 1999.

[27] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing – Principles, Algorithms, and Applications*, 4th ed. Upper Saddle River, NJ, US: Pearson Education, Apr. 2006, ch. 13.

[28] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, Apr. 1975.

[29] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Transform domain prediction error method for improved acoustic echo and feedback cancellation," in *Proc. 20th European Signal Process. Conf.*, Aug. 2012, pp. 2422–2426.

Paper G

Analysis of Closed–Loop Acoustic Feedback Cancellation Systems

Meng Guo, Søren Holdt Jensen, Jesper Jensen, and Steven L. Grant

To be published in
IEEE Int. Conf. Acoust., Speech, Signal Process., May 2013.

The layout has been revised.

Analysis of Closed-Loop Acoustic Feedback Cancellation Systems

Meng Guo, Søren Holdt Jensen, Jesper Jensen, and Steven L. Grant

Abstract

In a previous study, the performance of an acoustic feedback/echo cancellation system was analyzed using a power transfer function method. Whereas the analysis result provides very accurate performance predictions in open-loop acoustic echo cancellation systems, it is less accurate in closed-loop acoustic feedback cancellation systems if there is a strong *correlation between the loudspeaker signal and the signals entering the microphones. This work extends the performance analysis to include the effects of the nonzero correlation on the adaptive filters. Simulation results verify that this extension provides much more accurate performance predictions in closed-loop acoustic feedback cancellation systems.*

1 Introduction

Acoustic feedback problems arise when a microphone of an audio system picks up part of its acoustic output signal from the loudspeaker. Acoustic feedback cancellation using adaptive filters [1–3] in a system identification setup [4, 5] has evolved to be a state-of-the-art solution [6–12]. Much work has been done to analyze/characterize [13–19] and improve [20–25] these adaptive algorithms in terms of robustness, stability bounds, convergence rate, steady-state behavior, complexity, etc.

In [26], an analysis is performed to describe the frequency domain performance characteristics for acoustic feedback cancellation (AFC) and/or acoustic echo cancellation (AEC) in a multiple-microphone and single-loudspeaker (MMSL) system, illustrated in Fig. 1, in terms of the concept of power transfer function (PTF). The AFC/AEC is carried out by adaptive filters $\hat{\mathbf{h}}_i(n)$, where n is the time index, and $i = 1, ..., P$, where P is the number of microphones, and the beamformer filters $\mathbf{g}_i$ are performing a spatial filtering on the feedback/echo compensated signals $e_i(n)$. The PTF analysis in [26] determined a simple and accurate approximation $\hat{\xi}(\omega, n)$ of the expected magnitude-squared transfer function from point A to B in Fig. 1, where ω is the discrete frequency index. This approximation allowed prediction of the convergence rate, steady-state behavior, and the tracking ability of AFC/AEC systems without knowing the true acoustic

feedback/echo paths $\mathbf{h}_i(n)$.

For simplicity, the analysis in [26] was performed in an open-loop system by omitting the forward path $\mathbf{f}(n)$ in Fig. 1, and the loudspeaker signal $u(n)$ was assumed to be uncorrelated with the incoming signals $x_i(n)$. Hence, whereas the results from [26] are very accurate for open-loop AEC systems, these results have certain limitations in closed-loop AFC systems.

Specifically, the most significant limitation occurs when the incoming signals $x_i(n)$ have long tails in their autocorrelation functions (compared to the system latency from microphone to loudspeaker), such as in most music and alarm signals. The loudspeaker signal $u(n)$ is then correlated with $x_i(n)$. This leads to a biased estimation of $\hat{\mathbf{h}}_i(n)$ [7], and it violates the assumption of uncorrelated $u(n)$ and $x_i(n)$ for the PTF prediction. Thus, for strongly correlated incoming signals, the derived PTF expressions in [26] provide poor predictions, although the expressions are relatively accurate when the incoming signals were speech signals as demonstrated in [26].

Another important application of the PTF approximation $\hat{\xi}(\omega, n)$ in AFC systems is to ensure system stability. The true PTF $\xi(\omega, n)$ is the unknown part of the expected magnitude-squared open-loop transfer function $E[|\mathrm{OLTF}(\omega, n)|^2]$ of the MMSL system expressed by $E\left[|\mathrm{OLTF}(\omega, n)|^2\right] = |F(\omega, n)|^2 \xi(\omega, n)$, where $F(\omega, n)$ is the, generally known, frequency response of $\mathbf{f}(n)$. If $|\mathrm{OLTF}(\omega, n)| < 1$, system stability is guaranteed [27]. However, when the estimation of $\hat{\mathbf{h}}(n)$ is biased due to the correlation between $u(n)$ and $x_i(n)$, $\hat{\xi}(\omega, n)$ determined in [26] would generally be too small. Even if the forward path gain $|F(\omega, n)|$ was chosen as $|F(\omega, n)| < 1/\sqrt{\hat{\xi}(\omega, n)}$, stability could not be guaranteed.

In this work, we derive an extended PTF approximation that includes the influence of potentially biased estimation of $\hat{\mathbf{h}}_i(n)$. In particular, this is done by allowing the correlation function between the loudspeaker signal $u(n)$ and the incoming signals $x_i(n)$ to be nonzero.

2 Review of Power Transfer Function

The PTF describes the expected magnitude-squared transfer function from point A to B in Fig. 1, where the frequency responses $H_i(\omega, n)$ of the true feedback paths $\mathbf{h}_i(n)$ are unknown and considered stochastic. Hence, as in [26], we define the exact PTF of the MMSL system as $\xi(\omega, n) = E[|\sum_{i=1}^{P} G_i(\omega)\tilde{H}_i(\omega, n)|^2]$, where $G_i(\omega)$ is the frequency response of $\mathbf{g}_i$, and $\tilde{H}_i(\omega, n) = \hat{H}_i(\omega, n) - H_i(\omega, n)$ is the frequency response of $\tilde{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n) - \mathbf{h}_i(n)$. Clearly, $\xi(\omega, n) = \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)\xi_{ij}(\omega, n)$, where * denotes complex conjugation and $\xi_{ij}(\omega, n) = E[\tilde{H}_i(\omega, n)\tilde{H}_j^*(\omega, n)]$.

In general, however, we can not calculate the PTF $\xi(\omega, n)$ directly because $H_i(\omega, n)$ is unknown. In [26], an approximation $\hat{\xi}_{ij}(\omega, n) \approx E[\tilde{H}_i(\omega, n)\tilde{H}_j^*(\omega, n)]$ was introduced,

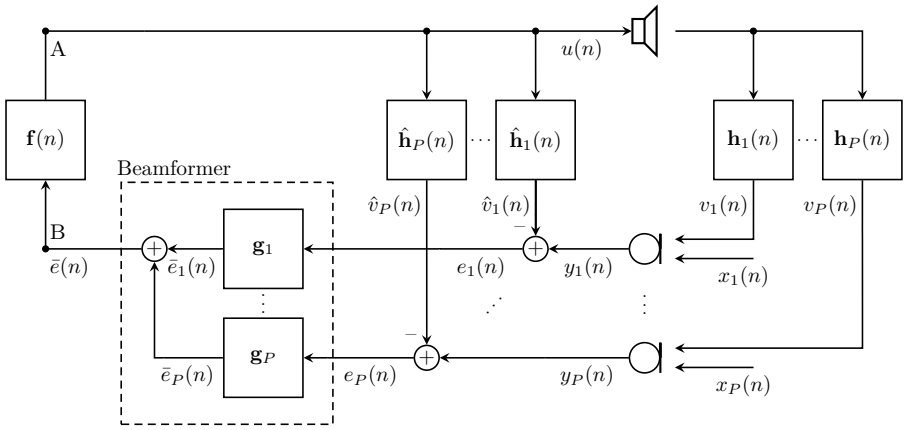


Fig. 1: A multiple-microphone and single-loudspeaker system.

where $\hat{\xi}_{ij}(\omega, n)$ is expressed by a relatively simple function, leading to an approximate PTF $\hat{\xi}(\omega, n) = \sum_{i=1}^{P} \sum_{j=1}^{P} G_i(\omega) G_j^*(\omega) \hat{\xi}_{ij}(\omega, n)$.

In [26], we derived PTF approximations $\hat{\xi}(\omega, n)$ for several adaptive algorithms for estimating the feedback/echo paths $\mathbf{h}_i(n)$. In this work, to limit our scope, we focus on $\hat{\xi}(\omega, n)$ for the least mean square (LMS) adaptive algorithm. Under the assumptions of uncorrelated $u(n)$ and $x_i(n)$, the LMS step size $\mu(n) \to 0$, the length of the adaptive filter $L \to \infty$, and $r_{x_{ij}}(k) = E[x_i(n) x_j(n+k)] = 0 \ \forall \ |k| > k_0 \in \mathbb{N}$, the PTF could be approximated as [26],

$$\begin{aligned} \hat{\xi}(\omega, n) = &\left(1 - 2\mu(n) S_u(\omega)\right) \hat{\xi}(\omega, n-1) \\ &+ L\mu^2(n) S_u(\omega) \sum_{i=1}^{P} \sum_{j=1}^{P} G_{ij}(\omega) S_{x_{ij}}(\omega) + \sum_{i=1}^{P} \sum_{j=1}^{P} G_{ij}(\omega) S_{\check{h}_{ij}}(\omega), \end{aligned} \tag{1}$$

where $S_u(\omega)$ denotes the power spectrum density (PSD) of the loudspeaker signal $u(n)$, $S_{x_{ij}}(\omega)$ denotes the cross(auto) PSDs of the incoming signals $x_i(n)$ and $x_j(n)$, $G_{ij}(\omega) = G_i(\omega) G_j^*(\omega)$, and $S_{\check{h}_{ij}}(\omega)$ is the PSD of the feedback/echo path variations over time.

Furthermore, system behavior in terms of the convergence rate, steady-state error, and the tracking error can be determined using Eq. (1), we refer to [26] for details.

3 Extended PTF in Closed-Loop Systems

In this section, we derive an extended PTF approximation for the case where $u(n)$ and $x_i(n)$ may be correlated. We do this based on the bias of the Wiener solution of the adaptive filter estimation for AFC systems.

3.1 Definition of Extended PTF

We consider the adaptive filter estimate $\hat{\mathbf{h}}_i(n)$ as $\hat{\mathbf{h}}_i(n) = \bar{\mathbf{h}}_i(n) + \breve{\mathbf{h}}_i(n)$, where $\bar{\mathbf{h}}_i(n)$ denotes the unbiased estimate from the adaptive algorithms if $u(n)$ was uncorrelated with $x_i(n)$, i.e. $E[\bar{\mathbf{h}}_i(n)] = \mathbf{h}_i(n)$, and $\breve{\mathbf{h}}_i(n)$ is the additional bias vector due to the correlation between $u(n)$ and $x_i(n)$.

The expected value of the adaptive filter estimate $\hat{\mathbf{h}}_i(n)$ which minimizes $E[e_i^2(n)]$ in MMSL systems as in Fig. 1 can be shown to be $E[\hat{\mathbf{h}}_i(n)] = \mathbf{h}_i(n) + E[\breve{\mathbf{h}}_i(n)]$, and $E[\breve{\mathbf{h}}_i(n)]$ is the Wiener solution of the bias vector $\breve{\mathbf{h}}_i(n)$ given by

$$E\left[\breve{\mathbf{h}}_i(n)\right] = E\left[\mathbf{u}(n)\mathbf{u}^T(n)\right]^{-1} \cdot E\left[\mathbf{u}(n)x_i(n)\right], \tag{2}$$

where $\mathbf{u}(n) = [u(n), u(n-1), \ldots, u(n-L+1)]^T$. An unbiased solution $E[\hat{\mathbf{h}}_i(n)] = \mathbf{h}_i(n)$ is obtained if $E[\mathbf{u}(n)x_i(n)] = \mathbf{0}$. However, this is usually not the case for AFC systems.

We now study the impact of $E[\breve{\mathbf{h}}_i(n)]$ on the PTF. The frequency response of $\hat{\mathbf{h}}_i(n)$ is denoted by $\hat{H}_i(\omega, n) = \bar{H}_i(\omega, n) + \breve{H}_i(\omega, n)$. We express the exact extended PTF $\breve{\xi}_{\text{exact}}(\omega, n)$ as

$$\breve{\xi}_{\text{exact}}(\omega, n) = E\left[\left|\sum_{i=1}^{P} G_i(\omega)\left(\hat{H}_i(\omega, n) - H_i(\omega, n)\right)\right|^2\right]. \tag{3}$$

Eq. (3) can be simplified. For a small step size $\mu(n)$, in principle $\mu(n) \to 0$, the fluctuation of the unbiased adaptive estimate $\bar{H}_i(\omega, n)$ tends to 0, i.e. $\bar{H}_i(\omega, n) - H_i(\omega, n) \to 0$. We can thereby neglect the cross-term $E[\breve{H}_i(\omega, n)(\bar{H}_i^*(\omega, n) - H_i^*(\omega, n))]$ compared to the auto-term $E[\breve{H}_i(\omega, n)\breve{H}_i^*(\omega, n)]$ when evaluating Eq. (3), which becomes

$$\begin{aligned}\breve{\xi}_{\text{exact}}(\omega, n) =& E\left[\left|\sum_{i=1}^{P} G_i(\omega)\left(\bar{H}_i(\omega, n) - H_i(\omega, n)\right)\right|^2\right] + E\left[\left|\sum_{i=1}^{P} G_i(\omega)\breve{H}_i(\omega, n)\right|^2\right] \\ =& \xi(\omega, n) + \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega)E\left[\breve{H}_i(\omega, n)\breve{H}_j^*(\omega, n)\right].\end{aligned} \tag{4}$$

Thus, the correlation between $x_i(n)$ and $u(n)$ leads to a nonzero bias term (the last term) in Eq. (4), which in turn leads to an increase in the extended PTF $\breve{\xi}_{\text{exact}}(\omega, n)$, over the PTF $\xi(\omega, n)$ which would have been achieved, had $x_i(n)$ and $u(n)$ been uncorrelated. To further simplify Eq. (4), we replace $\breve{H}_i(\omega, n)$ with its expected value $E[\breve{H}_i(\omega, n)]$, since we are studying the steady-state effects of $E[\breve{\mathbf{h}}_i(n)]$ in Eq. (2) on the PTF, and $\breve{H}_i(\omega, n) \to E[\breve{H}_i(\omega, n)]$ in steady-state for the LMS step size $\mu(n) \to 0$ [28]. Furthermore, by replacing the PTF $\xi(\omega, n)$ in Eq. (4) with its approximation $\hat{\xi}(\omega, n)$, we get the extended

PTF approximation $\breve{\xi}(\omega, n)$, as

$$\breve{\xi}(\omega, n) = \hat{\xi}(\omega, n) + \sum_{i=1}^{P} \sum_{j=1}^{P} G_i(\omega) G_j^*(\omega) E\left[\breve{H}_i(\omega, n)\right] \cdot E\left[\breve{H}_j^*(\omega, n)\right]. \tag{5}$$

The last term in Eq. (5) is independent of the step size parameter in applied adaptive algorithms, since the expected value of adaptive filter bias $E[\breve{\mathbf{h}}_i(n)]$ only depends on the incoming signals x_i(n) and the loudspeaker signal $u(n)$ as given in Eq. (2). Furthermore, this bias term can be considered as an additional error contribution to the steady-state error given by $\hat{\xi}(\omega, n)$. Therefore, while the convergence of $\breve{\xi}(\omega, n)$ is only determined by $\hat{\xi}(\omega, n)$, the steady-state behavior of $\breve{\xi}(\omega, n)$ is determined by both $\hat{\xi}(\omega, n)$ and this bias term. In the following sections, we study the influences of this bias term.

3.2 Extended PTF Analysis

We model the forward path $\mathbf{f}(n)$ as $\mathbf{f}(n) = \mathbf{f}_0(n) * \delta(n-d)$, where $*$ denotes convolution, i.e. a filtering part $\mathbf{f}_0(n)$ and a delay of $d > 0$ samples. Let $\mathbf{F}_0(n) \in \mathbb{R}^{L \times L}$ and $\mathbf{G}_i \in \mathbb{R}^{L \times L}$ be the Toeplitz structured convolution matrices of the forward path filter $\mathbf{f}_0(n)$ and the beamformer filter $\mathbf{g}_i$, respectively. We define the incoming signal vector as $\mathbf{x}_i(n) = [x_i(n), x_i(n-1), \ldots, x_i(n-L+1)]^T$, and by considering the case that $H_i(\omega, n) - \hat{H}_i(\omega, n)$ is relatively small in a steady-state situation, i.e. the adaptive filter provides a relatively precise estimate $\hat{H}_i(\omega, n)$ despite an eventual bias term $\breve{H}_i(\omega, n) \neq 0$, we can neglect the closed-loop effect, given by the transfer function $1/(1 - F_0(\omega, n) G_i(\omega)(H_i(\omega, n) - \hat{H}_i(\omega, n)))$, on the loudspeaker signal vector $\mathbf{u}(n)$, which is now simply expressed by

$$\mathbf{u}(n) = \sum_{i=1}^{P} \mathbf{F}_0(n) \mathbf{G}_i \mathbf{x}_i(n-d). \tag{6}$$

The correspondence of the resulting theory and the simulation results presented later showed that Eq. (6) is reasonable for even relatively large values of the bias term $\breve{H}_i(\omega, n)$ and thereby $H_i(\omega, n) - \hat{H}_i(\omega, n)$. Inserting Eq. (6) in Eq. (2), $E[\breve{\mathbf{h}}_i(n)]$ can be written as

$$E\left[\breve{\mathbf{h}}_i(n)\right] = \left(\sum_{p=1}^{P} \sum_{q=1}^{P} \mathbf{F}_0(n) \mathbf{G}_p \mathbf{R}_{x_{pq}}(0) \mathbf{G}_q^T \mathbf{F}_0^T(n) \right)^{-1} \cdot \sum_{p=1}^{P} \mathbf{F}_0(n) \mathbf{G}_p \mathbf{r}_{x_{pi}}(d), \tag{7}$$

where $\mathbf{R}_{x_{ij}}(k) = E[\mathbf{x}_i(n) \mathbf{x}_j^T(n+k)]$ and $\mathbf{r}_{x_{ij}}(k) = E[\mathbf{x}_i(n) x_j(n+k)]$.

We compute the frequency response $E[\breve{H}_i(\omega, n)]$ of $E[\breve{\mathbf{h}}_i(n)]$ using the discrete Fourier transform (DFT) matrix $\mathbf{D} \in \mathbb{C}^{L \times L}$. Since the DFT matrix $\mathbf{D}$ diagonalizes any Toeplitz matrix asymptotically, as $L \to \infty$ [29], and the matrices $\mathbf{F}_0(n)$, $\mathbf{G}_i$, and $\mathbf{R}_{x_{ij}}(0)$ are all

asymptotically Toeplitz matrices, each element $E[\breve{H}_i(\omega,n)]$ of the frequency response vector $\mathbf{D}E[\breve{\mathbf{h}}_i(n)]$ can be shown to be

$$E\left[\breve{H}_i(\omega,n)\right] = \frac{\sum_{p=1}^{P} F_0(\omega,n)G_p(\omega)\Gamma_{x_{pi}}(\omega)}{\sum_{p=1}^{P}\sum_{q=1}^{P} F_0(\omega,n)G_p(\omega)S_{x_{pq}}(\omega)G_q^*(\omega)F_0^*(\omega,n)}. \tag{8}$$

$\Gamma_{x_{pi}}(\omega)$ are elements of the vector $\mathbf{D}\mathbf{r}_{x_{pi}}(d)$, which is the DFT of the autocorrelation tail sequence $r_{x_{pi}}(d), r_{x_{pi}}(d+1), \ldots, r_{x_{pi}}(d+L-1)$. Furthermore, $S_{x_{pq}}(\omega)$ are the diagonal entries of the matrix $\frac{1}{L}\mathbf{D}\mathbf{R}_{x_{pq}}(0)\mathbf{D}^H$, these approach the DFT of the autocorrelation sequence $r_{x_{pq}}(0), r_{x_{pq}}(1), \ldots, r_{x_{pq}}(L-1)$, as $L \to \infty$. Finally, inserting Eq. (8) in Eq. (5), we get the expression for the extended PTF approximation $\breve{\xi}(\omega,n)$ as

$$\begin{aligned}\breve{\xi}(\omega,n) =& \hat{\xi}(\omega,n) + \sum_{i=1}^{P}\sum_{j=1}^{P} G_i(\omega)G_j^*(\omega) \\ &\cdot \left(\frac{\sum_{p=1}^{P} F_0(\omega,n)G_p(\omega)\Gamma_{x_{pi}}(\omega)}{\sum_{p=1}^{P}\sum_{q=1}^{P} F_0(\omega,n)G_p(\omega)S_{x_{pq}}(\omega)G_q^*(\omega)F_0^*(\omega,n)}\right) \\ &\cdot \left(\frac{\sum_{p=1}^{P} F_0(\omega,n)G_p(\omega)\Gamma_{x_{pj}}(\omega)}{\sum_{p=1}^{P}\sum_{q=1}^{P} F_0(\omega,n)G_p(\omega)S_{x_{pq}}(\omega)G_q^*(\omega)F_0^*(\omega,n)}\right)^*. \end{aligned} \tag{9}$$

4 Discussions

Generally, Eq. (9) is not easily interpreted. However, for a single-microphone and single-loudspeaker (SMSL) system ($P = 1$), the extended PTF approximation $\breve{\xi}(\omega,n)$ given by Eq. (9) simply reduces to

$$\breve{\xi}(\omega,n) = \hat{\xi}(\omega,n) + \frac{|\Gamma_x(\omega)|^2}{|F_0(\omega,n)|^2 S_x^2(\omega)}. \tag{10}$$

In general, $\breve{\xi}(\omega,n) > \hat{\xi}(\omega,n)$ since $|\Gamma_x(\omega)| > 0$ in Eq. (10). However, for incoming signals $x(n)$ fulfilling $r_x(k) = 0 \;\forall\; |k| > k_0 \in \mathbb{N}$, increasing the forward path delay d would generally decorrelate $u(n)$ from $x(n)$, see e.g. [30–32], and for a large value of $d > k_0$, we get $\Gamma_x(\omega) = 0$ leading to $\breve{\xi}(\omega,n) = \hat{\xi}(\omega,n)$ in Eq. (10).

Furthermore, Eq. (10) reveals that increasing the forward path gain $|F_0(\omega,n)|$ leads to a smaller bias term in $\breve{\xi}(\omega,n)$. Intuitively this can be explained by the fact when the forward path gain $|F_0(\omega,n)|$ gets larger, the larger is $u(n)$ compared to $x(n)$, see e.g. Eq. (6), the amplitude of each element in the expected bias vector $E[\breve{\mathbf{h}}_i(n)]$ would thereby be smaller as also seen from Eq. (2). However, although a large forward path gain $|F_0(\omega,n)|$ leads to a small $\breve{\xi}(\omega,n)$ in Eq. (10), $|F_0(\omega,n)|$ is still a compromise between

Table 1: Common parameters for simulation experiments.

Symbol	Value	Description
D_s	80000	Duration of simulation.
R	100	Number of sim. runs.
μ	2^{-11}	LMS step size.
L	32	Length of $\hat{\mathbf{h}}(n)$.
$\mathbf{g}$	$0.1 \cdot [10, 3, -2.5, 1, 0.5]^T$	Beamformer filter.
$\mathbf{h}(0)$	$0.01 \cdot [6, 0.84, -1.38]^T$	Initial values of $\mathbf{h}(n)$.
$N(\mu_h, \sigma_h^2)$	$N(0, 0.0019^2)$	Feedback path variation.
$\mathbf{h}_x$	$0.01 \cdot [10, -3, 4, -1, 0.5]^T$	Shaping filter for $x(n)$.
$\mathbf{f}_0$	$[10]^T$	Forward path filter.
d	1	Forward path delay.

being large enough to ensure a reasonably small bias, and being small enough to maintain system stability. In particular, $|F_0(\omega, n)|$ should be chosen according to $|F_0(\omega, n)| < 1/\sqrt{\breve{\xi}(\omega, n)}$ to ensure stability. Furthermore, using Eq. (10), a lower bound for the magnitude-squared open-loop transfer function $|\text{OLTF}(\omega, n)|^2 = |F_0(\omega, n)|^2 \breve{\xi}(\omega, n)$ is obtained as $|\Gamma_x(\omega)|^2 / S_x^2(\omega)$ for $\hat{\xi}(\omega, n) \to 0$, and it is interesting to note that it is actually independent of $|F_0(\omega, n)|$.

5 Simulation Verifications

In this section, we perform simulation experiments to verify the extended PTF approximation $\breve{\xi}(\omega, n)$ in Eq. (9) and show the improvements by comparing to the PTF approximation $\hat{\xi}(\omega, n)$ in Eq. (1). We consider an SMSL system ($P = 1$) with a known feedback path $\mathbf{h}(n)$, which remains fixed during the first half of the simulation experiment, but undergoes variations in the second half. We use the same procedure as described in Sec. VI-A of [26] for these simulation experiments. We refer to [26] for details. The only difference is that $u(n)$ is no longer independently generated, but it is rather computed in the closed-loop system as the error signal $\bar{e}(n)$ filtered through a time invariant forward path $\mathbf{f}$. Table 1 shows the general simulation parameters; the beamformer filter $\mathbf{g}$, initial feedback path $\mathbf{h}(0)$, and the forward path filter $\mathbf{f}_0$ are chosen to be lower order filters for reproducibility in these experiments.

In the first simulation experiment, we verify that the extended PTF expression in Eq. (9) can accurately predict biased steady-state values for an SMSL AFC system using an LMS algorithm. A biased estimation of $\mathbf{h}(n)$ is expected due to the choices of the forward path delay d and the shaping filter $\mathbf{h}_x$, which is used to generate the incoming signals $x(n)$ by convolving $\mathbf{h}_x$ with a white noise sequence.

Fig. 3 shows the results at two representative example frequency bins $l = 3, 7$. The

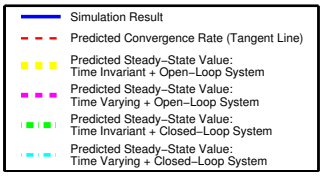


Fig. 2: Legends for Figs. 3-5.

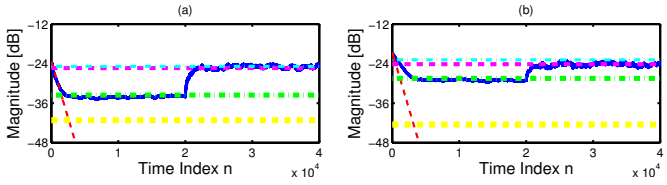


Fig. 3: PTF values for frequency bins (a) $l = 3$. (b) $l = 7$. See Fig. 2 for legend.

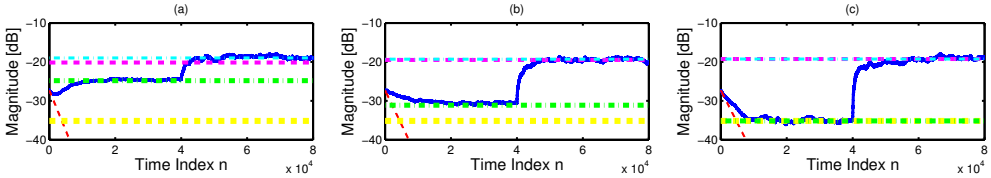


Fig. 4: PTF values for frequency bin 3. The forward path delay is (a) $d = 1$ sample. (b) $d = 3$ samples. (c) $d = 5$ samples. See Fig. 2 for legend.

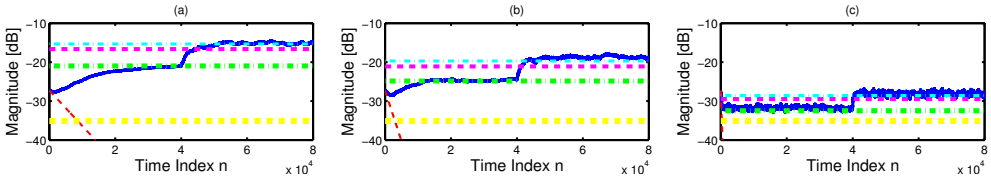


Fig. 5: PTF values for frequency bin 3. The forward path gain is (a) $|F_0(\omega)| = 6$ dB. (b) $|F_0(\omega)| = 10$ dB. (c) $|F_0(\omega)| = 20$ dB. See Fig. 2 for legend.

true PTF values can be calculated in simulations because $\mathbf{h}(n)$ is known. These true PTF values confirm the PTF prediction values using Eqs. (1) and (10) for computing the convergence rate and steady-state behaviors for the open-loop (see details in [26]) and closed-loop systems, respectively. Furthermore, the closed-loop PTF values for biased estimation of $\hat{\mathbf{h}}(n)$ are generally found at higher levels than the open-loop PTF values without biased estimation.

In the second simulation experiment, we show the dependence between forward path delay d and the PTF values. All simulation parameters are the same as given in Table 1, except the forward path delay, which is $d = 1, 3, 5$ in three different simulations. Fig. 4 shows the simulation results. The shaping filter $\mathbf{h}_x$ is a fourth-order filter, it means that $r_x(k) = 0 \ \forall \ |k| > 4$. Clearly, with a forward path delay $d = 1, 3$, $\breve{\xi}(\omega) > \hat{\xi}(\omega)$ due to biased estimation of $\hat{\mathbf{h}}(n)$. For a longer enough delay $d = 5$, $\breve{\xi}(\omega) = \hat{\xi}(\omega)$.

In the last simulation experiment, we show the dependence between the forward path gain $|F_0(\omega)|$ and the PTF values. We again use the general simulation parameters

given in Table 1, only the forward path filter $\mathbf{f}_0$ varies so that the $|F_0(\omega)| = 6, 10, 20$ dB, respectively. Fig. 5 shows the simulation results for different forward path gains $|F_0(\omega)|$. As expected, a higher forward gain $|F_0(\omega)|$ gives lower steady-state PTF values, as expressed in Eq. (10).

6 Conclusion

In previous work, PTF approximations have been derived for different adaptive algorithms in open-loop MMSL systems. In this work, we derived extensions to these PTF expressions for closed-loop MMSL systems. We showed that this extension provides more accurate and useful performance predictions of closed-loop AFC systems, especially if there are strong correlations between the loudspeaker and incoming signals, and the adaptive filter estimates therefore become heavily biased. The results also showed the relations between the forward path delay/gain in closed-loop systems and the biased adaptive filter estimates and thereby the PTF prediction values. This knowledge is important in designing AFC systems and provides a very useful upper-limit for the forward path gain to guarantee system stability in closed-loop AFC systems.

7 Relations to Prior Work

This work is an extension of the power transfer function analysis of a multiple-microphone and single-loudspeaker cancellation system introduced in [26]. The extended analysis provides accurate cancellation performance predictions in closed-loop systems even if there is a *strong* correlation between the loudspeaker and incoming signals, see details in the Introduction. The power transfer function analysis is inspired by the studies in [33, 34] of tracking characteristics for an open-loop single-microphone and single-loudspeaker cancellation system, which is a special case of the presented framework.

References

[1] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *IRE WESCON Conv. Record, Part 4*, pp. 96–104, 1960.

[2] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, US: Prentice Hall, Sep. 2001.

[3] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, US: Wiley, Jun. 2003.

[4] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ, US: Prentice Hall, Dec. 1998.

[5] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*. Hoboken, NJ, US: Wiley, May 2001.

[6] M. G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 4, pp. 443–453, Jul. 2000.

[7] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Elsevier J. Franklin Inst.*, vol. 343, no. 6, pp. 545–573, Sep. 2006.

[8] N. A. Shusina and B. Rafaely, "Unbiased adaptive feedback cancellation in hearing aids by closed-loop identification," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 658–665, Mar. 2006.

[9] C. Boukis, D. P. Mandic, and A. G. Constantinides, "Toward bias minimization in acoustic feedback cancellation systems," *J. Acoust. Soc. Am.*, vol. 121, no. 3, pp. 1529–1537, Mar. 2007.

[10] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.

[11] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Dual microphone solution for acoustic feedback cancellation for assistive listening," in *Proc. 2012 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2012, pp. 149–152.

[12] M. Guo, S. H. Jensen, and J. Jensen, "Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancements," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2549–2563, Nov. 2012.

[13] G. Long, F. Ling, and J. G. Proakis, "The LMS algorithm with delayed coefficient adaptation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 9, pp. 1397–1405, Sep. 1989.

[14] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification – a survey," *Elsevier Automatica*, vol. 26, no. 1, pp. 7–21, Jan. 1990.

[15] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2811–2825, Sep. 1993.

[16] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2937–2944, Nov. 1994.

[17] A. W. H. Khong and P. A. Naylor, "Selective-tap adaptive filtering with performance analysis for identification of time-varying systems," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 5, pp. 1681–1695, Jul. 2007.

[18] V. H. Nascimento, M. T. M. Silva, L. A. Azpicueta-Ruiz, and J. Arenas-García, "On the tracking performance of combinations of least mean squares and recursive least squares adaptive filters," in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 3710–3713.

[19] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Acoustic feedback and echo cancellation strategies for multiple-microphone and single-loudspeaker systems," in *Proc. 45th Asilomar Conf. Signals, Syst., Comput.*, Nov. 2011, pp. 556–560.

[20] J. Cioffi and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 304–337, Apr. 1984.

[21] D. T. M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 92–114, Jan. 1991.

[22] S. L. Gay and S. Tavathia, "The fast affine projection algorithm," in *Proc. 1995 IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, May 1995, pp. 3023–3026.

[23] S. Koike, "A class of adaptive step-size control algorithms for adaptive filters," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1315–1326, Jun. 2002.

[24] Y. Avargel and I. Cohen, "Adaptive system identification in the short-time Fourier transform domain using cross-multiplicative transfer function approximation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 1, pp. 162–173, Jan. 2008.

[25] C. Schuldt, F. Lindstrom, and I. Claesson, "A low-complexity delayless selective subband adaptive filtering algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 5840–5850, Dec. 2008.

[26] M. Guo, T. B. Elmedyb, S. H. Jensen, and J. Jensen, "Analysis of acoustic feedback/echo cancellation in multiple-microphone and single-loudspeaker systems using a power transfer function method," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5774–5788, Dec. 2011.

[27] H. Nyquist, "Regeneration theory," *Bell System Tech. J.*, vol. 11, pp. 126–147, 1932.

[28] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing – Principles, Algorithms, and Applications*, 4th ed. Upper Saddle River, NJ, US: Pearson Education, Apr. 2006, ch. 13.

[29] R. M. Gray, *Toeplitz and Circulant Matrices: A Review.* Hanover, MA, US: Now Publishers Inc., Jan. 2006.

[30] J. Hellgren and U. Forssell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 906–913, Nov. 2001.

[31] B. Rafaely, N. A. Shusina, and J. L. Hayes, "Robust compensation with adaptive feedback cancellation in hearing aids," *Elsevier Speech Commun.*, vol. 39, no. 1-2, pp. 163–170, Jan. 2003.

[32] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.

[33] S. Gunnarsson and L. Ljung, "Frequency domain tracking characteristics of adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 1072–1089, Jul. 1989.

[34] S. Gunnarsson, "On the quality of recursively identified FIR models," *IEEE Trans. Signal Process.*, vol. 40, no. 3, pp. 679–682, Mar. 1992.