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

Meng Guo, *Student Member, IEEE*, Thomas Bo Elmedyb, Søren Holdt Jensen, *Senior Member, IEEE*, 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.**

***Index Terms*—Adaptive echo cancellation, adaptive feedback cancellation, convergence rate, steady-state behavior.**

I. 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 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 that 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 electroacoustic 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

Manuscript received March 16, 2011; revised July 08, 2011; accepted August 24, 2011. Date of publication September 19, 2011; date of current version November 16, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Maciej Niedzwiecki.

M. Guo is with the Department of Electronic Systems, Aalborg University, DK-9220, Denmark, and also with Oticon A/S, DK-2765 Smørum, Denmark (e-mail: guo@oticon.dk).

T. B. Elmedyb is with Oticon A/S, DK-2765 Smørum, Denmark (e-mail: tbn@oticon.dk).

S. H. Jensen is with the Department of Electronic Systems, Aalborg University, DK-9220 Aalborg, Denmark (e-mail: shj@es.aau.dk).

J. Jensen is with Oticon A/S, DK-2765 Smoerum, Denmark (e-mail: jsj@oticon.dk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2011.2168523

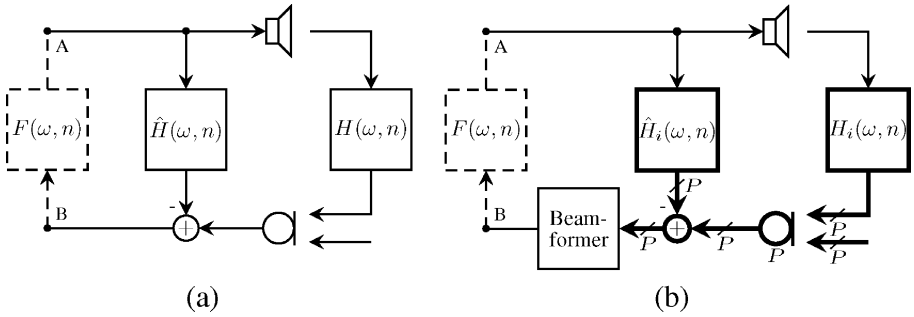

Fig. 1. Systems with acoustic feedback/echo and the cancellation. (a) A basic single-microphone and single-loudspeaker system. (b) A multiple-microphone and single-loudspeaker system with a beamformer, where $i = 1, \ldots, P$, and P denotes the number of microphone channels.

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:

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

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 (1), there are two criteria for the system stability. However, we ignore the phase information in (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] and [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] and [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 Section II, we introduce the system under analysis. We define the exact PTF $\xi(\omega,n)$ and its approximation in Section III. In Section IV, we present the detailed derivations of the PTF. After that, we discuss and verify the derived expressions through simulations in Sections V and VI, respectively. Finally, we give conclusions in Section VII.

II. System Description

A detailed overview of the MMSL system under analysis is given in Fig. 2. For convenience, we have expressed the feed-

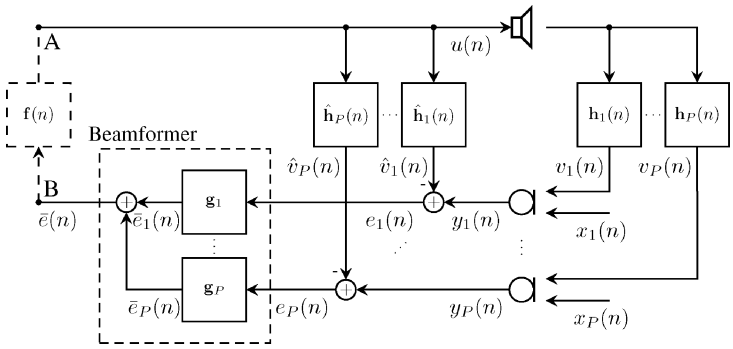

Fig. 2. A general multiple-microphone and single-loudspeaker system. In this paper, we focus on the power transfer function from point A to B.

back 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.

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 Section VI. The loudspeaker signal vector $\mathbf{u}(n)$ is defined as

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

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 $N-1$,

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

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

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

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].

III. 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

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

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] and [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 Section V.

The exact analytic expression for (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

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

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

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

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 Section VI, the derived results also provide accurate approximations in a closed-loop hearing aid AFC system with a realistic delay in $\mathbf{f}(n)$.

IV. System Analysis

A. 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],

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

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

$$\begin{aligned}\tilde{\mathbf{h}}_i(n) = &\left(\mathbf{I} - \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)\right)\tilde{\mathbf{h}}_i(n-1) \\ &+\mu(n)\mathbf{u}(n)x_i(n) - \check{\mathbf{h}}_i(n) \quad (23)\end{aligned}$$

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 (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

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

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 (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. Equation (24) can now be simplified as

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

Equation (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

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

Using (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

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

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

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

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

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

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

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

approaches a diagonal matrix, as $L \to \infty$, with the diagonal elements $\hat{\xi}_{ij}(\omega,n)$ as given in (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 (29) in (30) and using that $\frac{1}{L}\mathbf{F}^H\mathbf{F} = \mathbf{I}$, the matrix $\hat{\mathbf{\Xi}}_{ij}(n)$ is expressed by

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

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

$$\begin{aligned}\hat{\xi}_{ij}(\omega,n) = &\left(1 - 2\mu(n)S_u(\omega)\right)\hat{\xi}_{ij}(\omega,n-1)\\ &+ L\mu^2(n)S_u(\omega)S_{x_{ij}}(\omega) + S_{\check{h}_{ij}}(\omega) \end{aligned} \quad (32)$$

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

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

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

B. PTF for NLMS Algorithm

We can use the methodology from Section IV-A 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],

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

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, (34) can be rewritten as

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

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

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

Inserting (36) in (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

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

C. PTF for RLS Algorithm

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

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

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

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

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 Section IV-A 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

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

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 (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 (39) can be approximated by $\mathbf{Z}(n) \approx \mathbf{P}(n)$,[1] and the matrix $\mathbf{P}(n)$ in (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 (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 (44) in (41), the PTF approximation $\hat{\xi}(\omega,n)$ for the RLS algorithm is finally expressed by

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

V. 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.

Equations (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

$$\mathrm{CR}\left[\frac{\mathrm{dB}}{\mathrm{iteration}}\right] = \frac{d}{dn}10\log_{10}(\beta\cdot|\alpha|^n) = 10\log_{10}(|\alpha|). \tag{46}$$

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

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

The steady-state behavior is described through evaluation of

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

A. System Behavior for LMS Algorithm

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

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

Using (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 (33) in (48), the steady-state behavior is

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

Equation (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.

[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 $n < 1000$.

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

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

Using (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 (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 (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 (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 (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.

B. System Behavior for NLMS Algorithm

From (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 (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

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

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

Although the structure of (54)–(56) for the NLMS algorithm is similar to (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 (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 (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 (56) would instantly follow any variations in the frequency responses $G_i(\omega)$ and $G_j^*(\omega)$.

C. System Behavior for RLS Algorithm

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

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

and according to (47), stability is ensured for

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

The steady-state behavior is expressed by

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

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

$$\lambda = \frac{1 + 10^{\frac{\mathrm{CR}\left[\frac{\mathrm{dB}}{\mathrm{iteration}}\right]}{10}}}{2} \tag{62}$$

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

From (59)–(61), we observe some major differences in the RLS algorithm compared to the LMS and NLMS algorithms. For the

TABLE I
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
	$\uparrow S_{x_{ij}}(\omega)$	-	$\uparrow$	-
	$\uparrow S_{\check{h}_{ij}}(\omega)$	-	-	$\uparrow$

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 (61).

D. Summary of System Behavior

In Table I, 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 I 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 (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 (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 (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 subband structure with NLMS adaptation carried out in each subband, 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.

E. 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 (50) is smaller than the upper-bound defined in (65). This difference is due to the approximation made from (24) to (25), where we, under the assumption of $\mu(n) \to 0$, ignored the second-order term $E[\mu^2(n)\mathbf{u}(n)\mathbf{u}^T(n)\hat{\mathbf{h}}_i(n-1)\tilde{\mathbf{h}}_j^T(n-1)\mathbf{u}(n)\mathbf{u}^T(n)]$ in (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 (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 (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 (66)–(67) to the steady-state error of (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.

In our analysis, we used a model for the variations of feedback/echo paths identical to [4], so a comparison between (51) and (66) is therefore most meaningful. The eigenvalues $\lambda_{\check{h},k}$ and $\lambda_{u,k}$ in (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 (51) lead to the expression for the tracking error in (66). Similar expression of the tracking error is observed in (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] and [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 (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 $\hat{\xi}(\omega, \infty) = \lim_{n\to\infty} \frac{\bar{\mu}(n)}{2}$, which is the result presented in [37].

VI. 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.

A. 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 (19) using (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 simulation 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 $h_i(k, n) = h_i(k, 0) + \sum_{m=\frac{D_s}{2}}^{n} \epsilon_{h_i}(m)$, 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 $\sigma_{h_i}^2$. 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 II. A minimum order of various filters are used to be able to provide the numerical values of all coefficients.

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 (49) and (51) at two representative frequencies $\omega = \frac{2\pi l}{L}, l = 7, 11$ are shown in Fig. 3(a) and (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 indexes are computed

[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.

TABLE II
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 shown in Table III. 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 (52)–(53), we calculated the step size $\mu \approx 0.0012$ and $\mu \approx 0.0008$ to achieve a desired convergence rate of $-0.01 \frac{\text{dB}}{\text{iteration}}$ 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 and predicted results at frequency bin $l = 7$ are shown in Fig. 3(c) and (d). Again, the simulation results support the theory.

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

[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 $\delta = 0$.

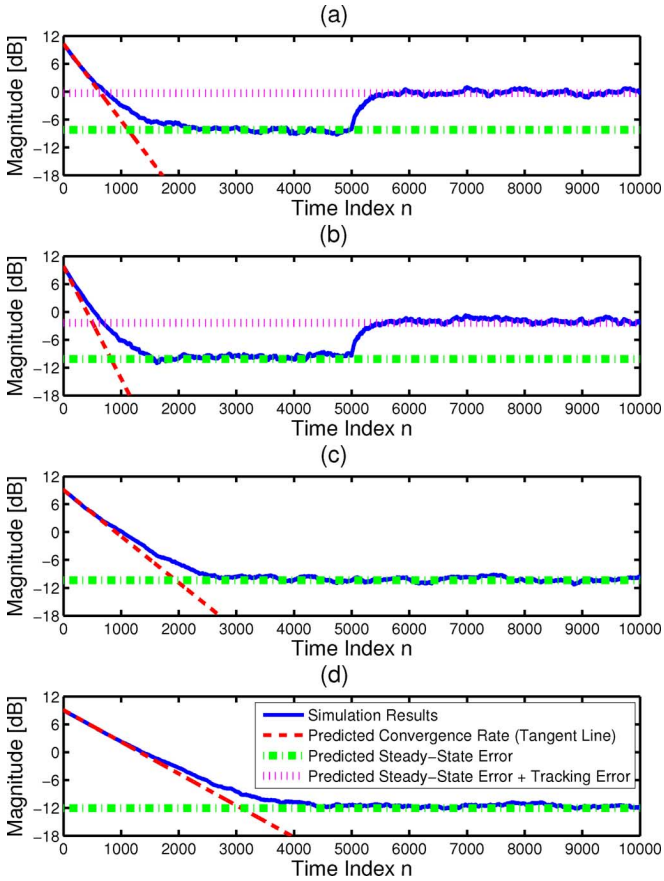

Fig. 3. LMS algorithm: the simulation results based on 100 simulation runs and the predicted values (a)-(b) Using (49) and (51) at frequency bin $l = 7$ and $l = 11$. (c)-(d) Using (52)–(53) at frequency bin $l = 7$. (a) PTF at frequency bin $l = 7$. (b) PTF at frequency bin $l = 11$. (c) Desired Convergence Rate $= -0.01 \frac{\text{dB}}{\text{iteration}}$. (d) Desired Steady-State Error $= -12$ dB.

TABLE III
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

In the second part of the experiment, using (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 $-0.01 \frac{\text{dB}}{\text{iteration}}$ and a steady-state error of -12 dB, respectively. In Fig. 4(c) and (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 (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

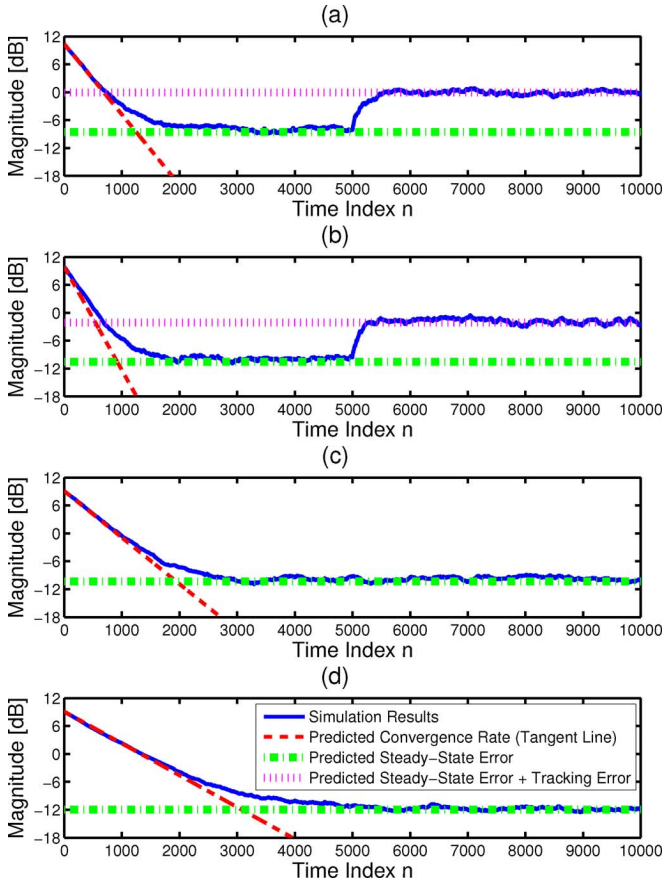

Fig. 4. NLMS algorithm: the simulation results based on 100 simulation runs and the predicted values (a)-(b) Using (54) and (56) at frequency bin $l = 7$ and $l = 11$. (c)-(d) Using (57)–(58) at frequency bin $l = 7$. (a) PTF at frequency bin l $= 7$. (b) PTF at frequency bin l $= 11$. (c) Desired Convergence Rate $= -0.01\ \frac{\text{dB}}{\text{iteration}}$. (d) Desired Steady-State Error $= -12$ dB.

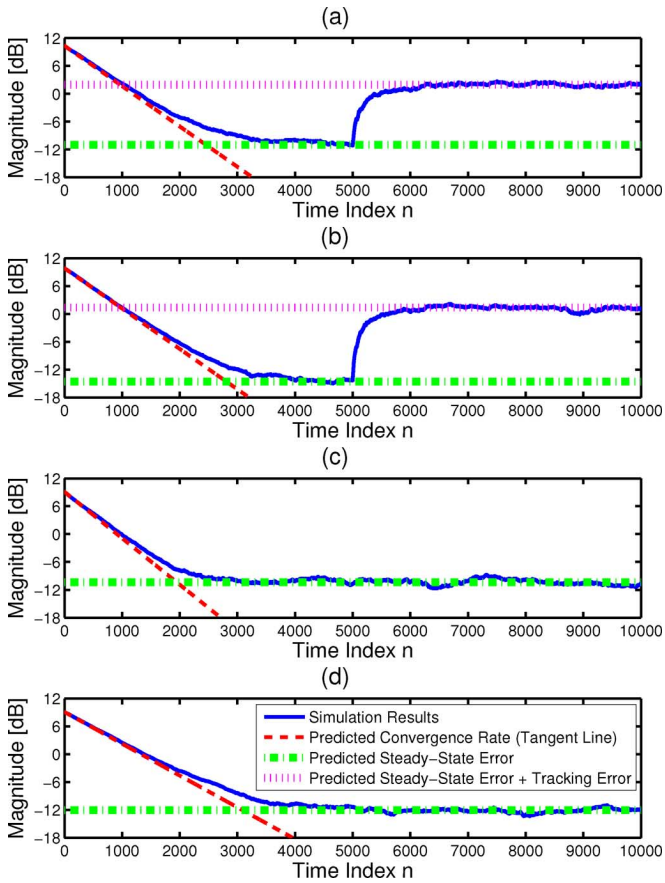

Fig. 5. RLS algorithm: the simulation results based on 100 simulation runs and the predicted values (a)-(b) Using (59) and (61) at frequency bin $l = 7$ and $l = 11$. (c)-(d) Using (62)–(63) at frequency bin $l = 7$. (a) PTF at frequency bin l $= 7$. (b) PTF at frequency bin l $= 11$. (c) Desired Convergence Rate $= -0.01\ \frac{\text{dB}}{\text{iteration}}$. (d) Desired Steady-State Error $= -12$ dB.

Figs. 3(c) and (d) and 4(c) and (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.

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 (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 (62)–(63), to obtain a desired convergence rate of $-0.01\ \frac{\text{dB}}{\text{iteration}}$ 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) and (d) and 5(c) and (d) shows that identical behaviors for the NLMS and RLS algorithms are obtained at one specific frequency. 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.

B. Simulation Experiment for Acoustic Feedback Cancellation

As discussed in Section III, 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.

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.

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.

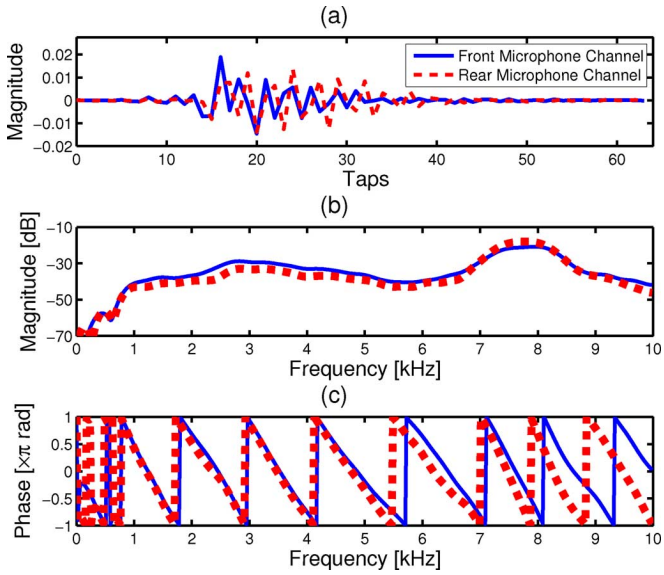

Fig. 6. Measured acoustic feedback paths from a behind-the-ear hearing aid with a sampling frequency $f_s = 20$ kHz. (a) Impulse response. (b) Magnitude response. (c) Phase response.

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 (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 (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 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.

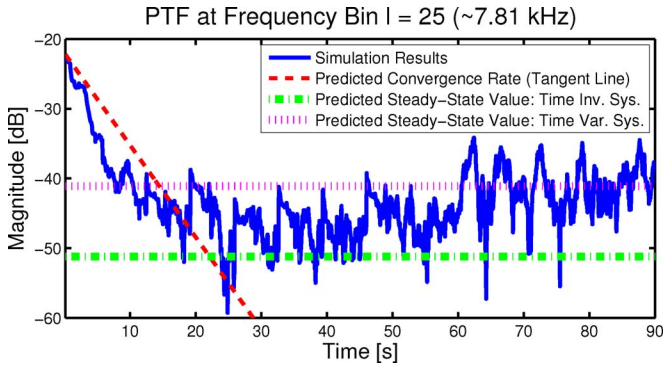

Fig. 7. Simulation results based on one simulation run in a closed-loop hearing aid acoustic feedback cancellation system with a female speech as the incoming signal. The results are given at the frequency bin 25 (~ 7.8 kHz), closest to the most critical frequency of 7.7 kHz.

In practice, the incoming signals $x_i(n)$ are unknown, and it is therefore not possible to compute $S_{x_{ij}}(\omega)$ directly. However, from the derived PTF expressions (49) and (51) for the LMS algorithm, (54) and (56) for the NLMS algorithm, and (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.

VII. 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.

Appendix
Estimation Error Correlation Matrix

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

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

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 (68) in the second last term of (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

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

It can be shown that (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 (28) and considering the case where $n-1 \geq k_0$, (69) can be further expressed by

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

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

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

Inserting (71) in (70), it can be seen, that all terms besides $\mathbf{I}$ in (71) result in higher order terms involving $\mu(n)\mu(n+k)\mathbf{u}(n)\mathbf{u}^T(n+k)$ in (70) and are thereby neglected. Hence, (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 (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 (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 (72)–(73) in (27), we get (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, 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. Amer.*, vol. 36, no. 9, pp. 1718–1724, 1964.

[4] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2001.

[5] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley-IEEE Press, 2003.

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

[7] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio applications," *Signal Process.*, vol. 89, no. 11, pp. 2185–2201, 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, 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, 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, 2011.

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

[12] B. Widrow and M. E. Hoff, "Adaptive switching circuits," in *IRE WESCON Conv. Rec., Part 4*, 1960, vol. 2, pp. 96–104.

[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, 1989.

[14] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification—A survey," *Automatica*, vol. 26, no. 1, pp. 7–21, 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, 1994.

[16] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. New York: Springer-Verlag, 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, 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, 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, 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.*, 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.*, 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, 2011.

[23] B. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. Bingley, U.K.: Emerald Group Publishing, 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. Amer.*, vol. 128, no. 3, pp. 1245–1261, 2010.

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

[26] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*. New York: Springer-Verlag, 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, 1989.

[28] S. Gunnarsson, "On the quality of recursively identified fir models," *IEEE Trans. Signal Process.*, vol. 40, no. 3, pp. 679–682, 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.*, pp. 433–436.

[30] 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 Eur. Signal Process. Conf.*, 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. Adv. Signal Process.*, vol. 2009, pp. 1–15, 2009.
[32] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory*. Cambridge, MA: MIT Press, 1984.
[33] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. Hanover, MA: Now Publishers Inc., 2006.
[34] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, 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, 1995.
[36] J. Huo, S. Nordholm, and Z. Zang, "New weight transform schemes for delayless subband adaptive filtering," in *IEEE Global Telecommun. Conf.*, 2001, vol. 1, 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," *Signal Process.*, vol. 64, no. 1, pp. 71–85, 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, 2011.
[39] J. Hellgren, T. Lunner, and S. Arlinger, "Variations in the feedback of hearing aids," *J. Acoust. Soc. Amer.*, vol. 106, no. 5, pp. 2821–2833, 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," *Speech Commun.*, vol. 39, no. 1-2, pp. 147–161, 2003.
[41] J. M. Kate, "Feedback cancellation in a hearing aid with reduced sensitivity to low-frequency tonal inputs," U.S. Patent 6 831 986 B2, Dec. 14, 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, Apr. 2, 2007.

Meng Guo (S'10) received the M.Sc. degree in applied mathematics from the Technical University of Denmark, Lyngby, Denmark, in 2006.

From 2007 to 2010, he was with Oticon A/S, Smørum, Denmark, as a Research and Development Engineer in the area of acoustic signal processing for hearing aid applications, especially in algorithm design of acoustic feedback cancellation. Currently, he is an industrial Ph.D. Fellow with Aalborg University, Aalborg, Denmark, and Oticon A/S. His main research interests are in the area of acoustical signal processing, including acoustic feedback cancellation, acoustic echo cancellation, adaptive filtering techniques and auditory signal processing.

Thomas Bo Elmedyb received the M.Sc. degree in electrical engineering from the Technical University of Denmark, Lyngby, Denmark, in 1999.

Currently, he is with Oticon A/S, Smørum, Denmark. His main research interests are in the area of signal processing for hearing aids including speech enhancement, beamforming, acoustic feedback cancellation, and perceptual aspects of signal processing for hearing impaired people.

Søren Holdt Jensen (S'87–M'88–SM'00) received the M.Sc. degree in electrical engineering from Aalborg University, Aalborg, Denmark, in 1988 and the Ph.D. degree in signal processing from the Technical University of Denmark, Lyngby, Denmark, in 1995.

Before joining the Department of Electronic Systems of Aalborg University, he was with the Telecommunications Laboratory of Telecom Denmark, Ltd., Copenhagen, Denmark; the Electronics Institute of the Technical University of Denmark; the Scientific Computing Group of Danish Computing Center for Research and Education (UNI•C), Lyngby, Denmark; the Electrical Engineering Department of Katholieke Universiteit Leuven, Leuven, Belgium; and the Center for PersonKommunikation (CPK) of Aalborg University. He is a Full Professor and is currently heading a research team working in the area of numerical algorithms and signal processing for speech and audio processing, image and video processing, multimedia technologies, and digital communications.

Prof. Jensen was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and Elsevier *Signal Processing*, and is currently Member of the Editorial Board of the *EURASIP Journal on Advances in Signal Processing* and Area Editor of the *Elsevier Digital Signal Processing*. He is a recipient of an European Community Marie Curie Fellowship, former Chairman of the IEEE Denmark Section, and Founder and Chairman of the IEEE Denmark Section's Signal Processing Chapter. He is member of the Danish Academy of Technical Sciences and was in January 2011 appointed as member of the Danish Council for Independent Research—Technology and Production Sciences by the Danish Minister for Science, Technology and Innovation.

Jesper Jensen received the M.Sc. degree in electrical engineering and the Ph.D. degree in signal processing from Aalborg University, Aalborg, Denmark, in 1996 and 2000, respectively.

From 1996 to 2000, he was with the Center for Person Kommunikation (CPK), Aalborg University, as a Ph.D. student and Assistant Research Professor. From 2000 to 2007., he was a Postdoctoral Researcher and Assistant Professor with Delft University of Technology, The Netherlands, and an external Associate Professor with Aalborg University, Denmark. Currently, he is with Oticon A/S, Smørum, Denmark. His main research interests are in the area of acoustical signal processing, including signal retrieval from noisy observations, coding, speech and audio synthesis, intelligibility enhancement of speech signals, and perceptual aspects of signal processing.