Stability-controlled hybrid adaptive feedback cancellation scheme for hearing aids

Sven Nordholm[a)]
Department of Electrical and Computer Engineering, Curtin University, Perth, WA, 6102, Australia

Henning Schepker
Signal Processing Group, Department of Medical Physics and Acoustics and Cluster of Excellence "Hearing4All," University of Oldenburg, Oldenburg, Germany

Linh T. T. Tran
Department of Electrical and Computer Engineering, Curtin University, Perth, WA, 6102, Australia

Simon Doclo
Signal Processing Group, Department of Medical Physics and Acoustics and Cluster of Excellence "Hearing4All," University of Oldenburg, Oldenburg, Germany

(Received 31 May 2017; revised 6 November 2017; accepted 12 December 2017; published online 12 January 2018)

Adaptive feedback cancellation (AFC) techniques are common in modern hearing aid devices (HADs) since these techniques have been successful in increasing the stable gain. Accordingly, there has been a significant effort to improve AFC technology, especially for open-fitting and in-ear HADs, for which howling is more prevalent due to the large acoustic coupling between the loudspeaker and the microphone. In this paper, the authors propose a hybrid AFC (H-AFC) scheme that is able to shorten the time it takes to recover from howling. The proposed H-AFC scheme consists of a switched combination adaptive filter, which is controlled by a soft-clipping-based stability detector to select either the standard normalized least mean squares (NLMS) algorithm or the prediction-error-method (PEM) NLMS algorithm to update the adaptive filter. The standard NLMS algorithm is used to obtain fast convergence, while the PEM-NLMS algorithm is used to provide a low bias solution. This stability-controlled adaptation is hence the means to improve performance in terms of both convergence rate as well as misalignment, while only slightly increasing computational complexity. The proposed H-AFC scheme has been evaluated for both speech and music signals, resulting in a significantly improved convergence and re-convergence rate, i.e., a shorter howling period, as well as a lower average misalignment and a larger added stable gain compared to using either the NLMS or the PEM-NLMS algorithm alone. An objective evaluation using the perceptual evaluation of speech quality and the perceptual evaluation of audio quality measures shows that the proposed H-AFC scheme provides very high-quality speech and music signals. This has also been verified through a subjective listening experiment with $N = 15$ normal-hearing subjects using a multi-stimulus test with hidden reference and anchor, showing that the proposed H-AFC scheme results in a better perceptual quality than the state-of-the-art PEM-NLMS algorithm. https://doi.org/10.1121/1.5020269

[MAH-J] Pages: 150–166

I. INTRODUCTION

In many acoustical systems, such as hearing aid devices (HADs), assisted listening devices, and public address systems, the microphone signals are amplified and played back through loudspeakers.[1–3] Due to the acoustic coupling between the loudspeakers and the microphones, acoustic feedback loops are generated. Here, we only consider the single-microphone and single-loudspeaker case, also called single-input single-output system. Depending on the acoustic transfer function between the loudspeaker and the microphone and the feedforward gain, annoying audible artifacts such as reverberation echoes and howling (indicating instability[4]) may occur. The maximum stable gain (MSG) of the system is determined by the acoustic feedback path between the loudspeaker and the microphone. For example, for HAD users the available maximum gain is an important parameter, since it determines if their hearing loss can be compensated or not. At the current time, the feedback problem in HADs has become more prevalent due to the use of small-sized HADs as well as open-fitting HADs (e.g., to avoid occlusion effects),[5] for which the acoustic coupling between the hearing aid receiver and the microphone is large.

In the last decades, several algorithms to manage feedback have been proposed, which can be broadly classified into feedforward suppression and feedback cancellation algorithms.[1,2] In feedforward suppression algorithms the signal is modified prior to being played back through the loudspeaker, such that undesired processing artifacts may arise. As a typical example, in notch-filtering-based howling suppression the feedforward gain function is modified using notch filters after howling has been detected.[6] In adaptive

[a)]Electronic mail: s.nordholm@curtin.edu.au

feedback cancellation (AFC) an adaptive filter, e.g., the normalized least mean squares (NLMS) algorithm, is used to identify the time-varying acoustic feedback path.[7–9] The latter approach is the most promising, allowing—at least in theory—for a perfect cancellation of the feedback signal.[10] However, in order to achieve perfect cancellation, knowledge of the correlation properties of the incoming signal is required.[1,2,11] In general, this correlation cannot be predetermined but needs to be estimated directly from the available microphone signal. Mismatches in this estimate typically lead to a biased solution and, in the worst-case scenario, it causes the feedback cancellation system to fail. Different techniques have been proposed to reduce this bias, thereby increasing the MSG while avoiding howling. Some of these techniques include phase modification and frequency shifting,[2,12] non-linear processing,[13] probe noise injection,[11,14,15] or the use of multiple microphones to estimate the incoming signal and whiten it prior to adapting the feedback canceler.[16–18] Furthermore, the use of adaptive decorrelating pre-filters has been proposed in the framework of the so-called prediction-error-method AFC (PEM-AFC).[1,19,20] In PEM-AFC, it is assumed that the incoming signal is white Gaussian noise filtered by a time-varying all-pole filter. The coefficients of the all-pole filter can be estimated from the microphone signal or from the error signal after the feedback canceler. The inverse of the estimated all-pole filter is then used to pre-filter the loudspeaker signal and the microphone signal prior to adapting the feedback canceler.

A significant practical challenge is to improve the AFC performance when there are fast changes in the acoustic feedback path or when howling occurs. In Ref. 11 recovery from howling was discussed without explicitly investigating this topic. In Ref. 13 techniques were presented to improve the AFC performance during clipping, resulting in a larger MSG, but no results on recovery from howling were provided. In addition, an earlier study showed that an affine combination of adaptive filters can be used to trade off the convergence rate and misalignment (MIS),[21] but again no results on recovery from howling were presented.

In this paper, we study the recovery of a hearing aid from a howling period using AFC. Since the standard NLMS algorithm is known to provide fast re-convergence from a howling period,[1,11] while the PEM-NLMS algorithm provides low MIS and a low bias solution,[1,19,20] we propose a hybrid switched combination of both adaptive filter algorithms. The proposed hybrid NLMS (H-NLMS) algorithm is controlled by a stability detector utilizing a soft clipper (SC) to detect whether the error signal after the feedback canceler is outside a certain bound, indicating instability. When instability is detected, the standard NLMS algorithm is used, whereas otherwise the PEM-NLMS algorithm is used. The parameters of the stability detector have been designed to provide both a high sensitivity to detecting instability (in order to re-converge fast after howling) and to provide low signal distortion after the AFC has converged. Another feature of the proposed H-NLMS algorithm is the improved AFC performance for music signals, where for certain sounds pre-filtering with the inverse of the estimated all-pole filter may yield long correlation tails. Since the stability detector will automatically detect instability occurring in sections of music with long correlation tails, it dynamically switches the hybrid combination of adaptive filters to aid the AFC operation in music. The proposed H-NLMS algorithm has been evaluated using both speech and music signals. The evaluation is based on instrumental performance measures, such as added stable gain (ASG) and MIS, as well as a formal listening test. First, the performance of the H-NLMS algorithm is investigated for different parameters of the stability detector. This instrumental evaluation shows that the H-NLMS algorithm is robust toward different parameter choices of the stability detector and yields clear improvements compared to the state-of-the-art PEM-NLMS algorithm in terms of convergence rate and recovery from howling. Next, a listening test has been conducted with $N = 15$ normal-hearing subjects, evaluating the perceptual signal quality for the proposed H-NLMS algorithm and the PEM-NLMS algorithm. Results show that the H-NLMS algorithm outperforms the PEM-NLMS algorithm both during the initial convergence as well as during re-convergence after a feedback path change.

The paper is structured as follows. Section II presents the system equations for acoustic feedback cancellation and the proposed hybrid AFC scheme controlled using a soft-clipping-based stability detector. In Sec. III the simulation results using instrumental performance measures and the formal subjective evaluations are presented.

II. HEARING AID SYSTEM DESCRIPTION

A HAD consists of one or more microphones, DSP technology (e.g., for AFC, noise suppression, equalization, and compression), an amplifier, and a loudspeaker. Figure 1 depicts the block diagram of a generic AFC scheme[1] for a HAD with a single microphone, where k denotes the discrete-time index. The feedback path $H(q, k)$ between the loudspeaker and the microphone is assumed to be a polynomial transfer function with L_h time-varying coefficients, i.e., $H(q,k) = \mathbf{h}^T(k)\mathbf{q}$, where the filter coefficients are given by $\mathbf{h}(k) = [h_0(k), h_1(k), \ldots h_{L_h-1}(k)]^T$ and $\mathbf{q} = [1 \quad q^{-1} \ldots \quad q^{-L_h+1}]^T$, with q^{-1} denoting the delay operator. The microphone signal $m(k)$ consists of the loudspeaker output signal $y(k)$ filtered by the feedback path $H(q, k)$ and the incoming signal $x(k)$, i.e.,

$$m(k) = H(q,k)y(k) + x(k). \tag{1}$$

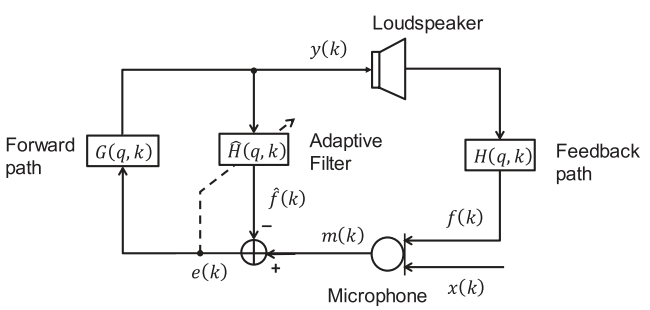

FIG. 1. Block diagram of the generic AFC scheme.

The microphone signal hence consists of the desired component $x(k)$ and the undesired feedback contribution $f(k) = H(q,k)y(k)$. When an estimate $\hat{H}(q,k)$ of the feedback path $H(q, k)$ is available, the feedback contribution can be estimated as $\hat{f}(k) = \hat{H}(q,k)y(k)$ and subtracted from the microphone signal, forming the error signal

$$\begin{aligned} e(k) &= x(k) + f(k) - \hat{f}(k) \\ &= \left[H(q,k) - \hat{H}(q,k)\right]y(k) + x(k). \end{aligned} \tag{2}$$

The error signal $e(k)$ is then filtered by the so-called feedforward path $G(q, k)$, representing the typical hearing aid processing (e.g., amplification, compression, noise suppression), i.e.,

$$y(k) = G(q,k)e(k). \tag{3}$$

By substituting Eq. (3) into Eq. (2), the following relationships between the incoming signal $x(k)$, the loudspeaker output signal $y(k)$, and the error signal $e(k)$ are obtained:

$$y(k) = G(q,k)S(q,k)x(k), \tag{4}$$

$$e(k) = S(q,k)x(k), \tag{5}$$

with the closed-loop transfer function $S(q, k)$ defined as

$$S(q,k) = \frac{1}{1 - G(q,k)\left[H(q,k) - \hat{H}(q,k)\right]}. \tag{6}$$

Thus, if the incoming signal $x(k)$ is bounded, i.e., $|x(k)| < M_x$, and the closed-loop system $S(q, k)$ is stable, the error signal $e(k)$ and the loudspeaker signal $y(k)$ will be bounded, i.e., $|y(k)| < M_y$. On the other hand, if the closed-loop system is unstable, the error signal $e(k)$ and the loudspeaker signal $y(k)$ will not be bounded. Note that in practice the microphone signal $m(k)$ will always be bounded due to the limited range of the analog-to-digital convertor and the loudspeaker signal $y(k)$ will always be bounded due to the loudspeaker circuitry and the non-linear loudspeaker characteristics. Nevertheless, if the (linear) closed-loop system $S(q, k)$ is unstable, there will be a growth in $e(k)$ which is not related to the incoming signal $x(k)$ but is determined by the instability of the linear system. The limiting amplitude output by the circuitry and the loudspeaker will result in different degrees of howling depending on the gain and the amplitude limits. Therefore, the error signal $e(k)$ can be used to monitor the stability of the system, i.e., to design a simple stability detector (cf. Sec. II C).

A. AFC

In the AFC scheme in Fig. 1, the aim of the filter $\hat{H}(q,k)$ is to estimate the (time-varying) feedback path $H(q, k)$. For this purpose, the error signal $e(k)$ in Eq. (2) is used to form the mean square error (MSE) cost function

$$J_{\text{MSE}} = E\{e^2(k)\}, \tag{7}$$

where $e(k) = m(k) - \hat{\mathbf{h}}^T(k)\mathbf{y}(k)$, $\hat{\mathbf{h}}(k)$ are the $L_{\hat{h}}$ filter coefficients of the feedback canceler at time instant k, $\mathbf{y}(k) = [y(k), y(k-1), \ldots, y(k - L_{\hat{h}} + 1)]^T$ is the filter input vector, and $E\{\cdot\}$ is the expectation operator. The signals $m(k)$ and $y(k)$ are considered realizations of the underlying stochastic processes. Minimizing Eq. (7) with respect to $\hat{\mathbf{h}}(k)$ yields the optimal solution

$$\hat{\mathbf{h}}_o(k) = \mathbf{R}_{\mathbf{yy}}^{-1}(k)\mathbf{r}_{\mathbf{y}m}(k), \tag{8}$$

with $\mathbf{R}_{\mathbf{yy}}(k) = E\{\mathbf{y}(k)\mathbf{y}^T(k)\}$ the correlation matrix of the loudspeaker signal and $\mathbf{r}_{\mathbf{y}m}(k) = E\{\mathbf{y}(k)m(k)\}$ the correlation vector between the loudspeaker signal and the microphone signal.

By substituting $m(k) = x(k) + \mathbf{y}^T(k)\mathbf{h}(k)$ into Eq. (8), we obtain

$$\hat{\mathbf{h}}_o(k) = \mathbf{h}(k) + \underbrace{\mathbf{R}_{\mathbf{yy}}^{-1}(k)\mathbf{r}_{\mathbf{y}x}(k)}_{\text{bias term}}, \tag{9}$$

which contains a term that depends on the correlation between the loudspeaker signal $y(k)$ and the incoming signal $x(k)$. If the correlation $\mathbf{r}_{\mathbf{y}x}(k) = 0$, then the feedback path estimate $\hat{\mathbf{h}}_o(k)$ is unbiased. However, since $y(k) = G(q,k)S(q,k)x(k)$, it follows that $y(k)$ and $x(k)$ are correlated, i.e., $\mathbf{r}_{\mathbf{y}x}(k) \neq \mathbf{0}$, giving rise to a biased estimate.[1,11,19]

To achieve an unbiased solution, different methods have been proposed. On the one hand, aiming at reducing the correlation between the loudspeaker signal and the incoming signal, methods based on phase modification and frequency shifting,[2,12] non-linear processing,[13] and probe noise injection[11,14,15] have been proposed, which as a disadvantage typically affect the sound quality of the loudspeaker output signal. On the other hand, the PEM-AFC method has been proposed, which aims at reducing the bias by pre-whitening the incoming signal. This method either assumes the correlation properties of the incoming signal to be known,[10,24] or in the case of an unknown incoming signal, estimates the correlation properties.[1,19,20] In many cases, the incoming signal can be adequately modelled using an auto-regressive (AR) model, which is known to hold well, e.g., for speech signals. The incoming signal $x(k)$ can then be modelled as

$$x(k) = A^{-1}(q,k)w(k), \tag{10}$$

with $A(q, k)$ a monic and stable polynomial transfer function and $w(k)$ a white Gaussian noise sequence. Usually, the error signal $e(k)$ after the feedback canceller is utilized to estimate the AR coefficients $\hat{A}(q,k)$, e.g., using the block-based Levinson-Durbin method[1] or the sample-based Burg Lattice method.[19]

The block diagram of the PEM-AFC scheme is shown in Fig. 2. Note that this block diagram also includes a SC, which will be described in detail in Sec. II C, so for the time being we neglect this SC, i.e., $e_L(k) = e(k)$. PEM-AFC can be viewed as a pre-filtered error method which minimizes the MSE of the error signal, pre-filtered with the estimated AR coefficients $\hat{A}(q,k)$, i.e.,

$$J_{\text{MSE,PEM}} = E\{\tilde{e}_L^2(k)\} = E\{(\hat{A}(q,k)e_L(k))^2\}. \tag{11}$$

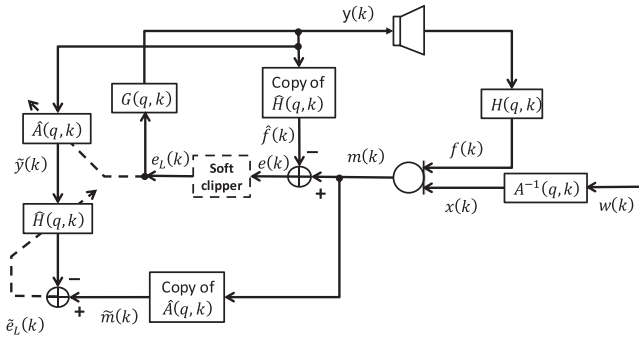

FIG. 2. Block diagram of the PEM-AFC scheme with/without a SC.

As can be seen from Fig. 2, the pre-filtered error signal is computed as

$$\tilde{e}_L(k) = \tilde{m}(k) - \hat{\mathbf{h}}^T(k)\tilde{\mathbf{y}}(k), \tag{12}$$

with the pre-filtered loudspeaker and microphone signals defined as $\tilde{y}(k) = \hat{A}(q,k)y(k)$ and $\tilde{m}(k) = \hat{A}(q,k)m(k)$, respectively. If $\hat{A}(q,k)$ whitens the incoming signal $x(k)$, i.e., $\hat{A}(q,k) = A(q,k)$, it has been shown that the bias term in Eq. (9) is equal to zero when appropriate delays are used.[19,25] Since this may be difficult to achieve, particularly for non-speech signals for which the assumed AR model does not hold well, PEM-AFC has been combined with frequency shifting.[26]

In practice, since the feedback path $H(q, k)$ has dynamic variations (e.g., when a phone moves toward the HAD), the optimal filter in Eq. (8) is replaced by an adaptive filter. The most commonly used adaptive filter in this application is the NLMS algorithm, i.e.,

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mu(k)\mathbf{y}(k)e_L(k), \tag{13}$$

with time-varying step-size $\mu(k) = \mu/(||\mathbf{y}(k)||^2 + \delta)$ and δ a small positive number to avoid division by zero. Similar to the NLMS algorithm, the corresponding PEM-NLMS algorithm using pre-filtered signals is given by

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mu(k)\tilde{\mathbf{y}}(k)\tilde{e}_L(k), \tag{14}$$

with time-varying step-size $\mu(k) = \mu/(||\tilde{\mathbf{y}}(k)||^2 + \delta)$. After updating the filter coefficients using Eq. (14), they are then copied to the feedback canceller, cf. Fig. 2.

B. H-NLMS algorithm

When the AFC is not able to estimate the feedback path $H(q, k)$ accurately enough, the closed-loop system $S(q, k)$ in Eq. (6) may become unstable. Even though the loudspeaker signal $y(k)$ is always bounded in practice, annoying tonal sounds referred to as howling will be generated in this case.[1,2,13] In addition to maximizing the available gain of the HAD, another objective is hence to make such howling periods as short as possible. Moreover, since the MSG depends on how well the AFC is able to estimate the time-varying feedback path, the AFC should provide both fast tracking as well as low bias and MIS, which are conflicting requirements for an adaptive filter algorithm.

When howling occurs, the feedback contribution will be dominating over the incoming signal $x(k)$, i.e., $|(H(q,k) - \hat{H}(q,k))y(k)| \gg |x(k)|$. In Refs. 1 and 11 it was noted that the standard NLMS algorithm is able to stabilize the system when it becomes unstable, i.e., it quickly recovers from howling. On the other hand, from our research we found that the PEM-NLMS algorithm takes a long time to recover from howling or in the worst case may not recover at all when the feedback contribution is not limited. This seems to be related to the fact that the estimated AR coefficients for an unstable system tend to model the howling instead of the incoming signal $x(k)$. To limit the feedback contribution, a SC is hence used which resembles the maximum loudspeaker level (cf. Fig. 2). The PEM-NLMS algorithm using a SC will be referred to as the PEMSC-NLMS algorithm.

Hence, since the standard NLMS algorithm is able to quickly recover from howling and the PEMSC-NLMS algorithm may provide an unbiased estimate, we propose to combine both in a H-NLMS algorithm, merging the best properties of both algorithms. The block diagram of the H-NLMS algorithm is depicted in Fig. 3. As a switched combination of the NLMS update in Eq. (13) and the PEMSC-NLMS update in Eq. (14), the proposed H-NLMS update is given by

$$\begin{aligned}\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) &+ [1-\alpha(k)]\mu_1(k)\tilde{\mathbf{y}}(k)\tilde{e}_L(k) \\ &+ \alpha(k)\mu_2(k)\mathbf{y}(k)e_L(k),\end{aligned} \tag{15}$$

with step-sizes $\mu_1(k) = \mu_1/(||\tilde{\mathbf{y}}(k)||^2 + \delta)$ and $\mu_2(k) = \mu_2/(||\mathbf{y}(k)||^2 + \delta)$, and $\alpha(k)$ a binary control signal. This control signal should be equal to 1 during instability, leading to the NLMS update using the loudspeaker signal $y(k)$ and the error signal $e_L(k)$, and equal to 0 during stable operation, leading to the PEMSC-NLMS update using the pre-filtered loudspeaker signal $\tilde{y}(k)$ and the pre-filtered error signal $\tilde{e}_L(k)$. Since the control signal $\alpha(k)$ is binary, the NLMS update and the PEMSC-NLMS update do not need to run in parallel, such that the computational complexity is only slightly increased due to the stability detector. The computation of the control signal $\alpha(k)$ using a soft-clipping-based stability detector will be explained in Sec. II C.

For comparison purposes, we have also used a combination of two PEMSC-NLMS updates using two different step-

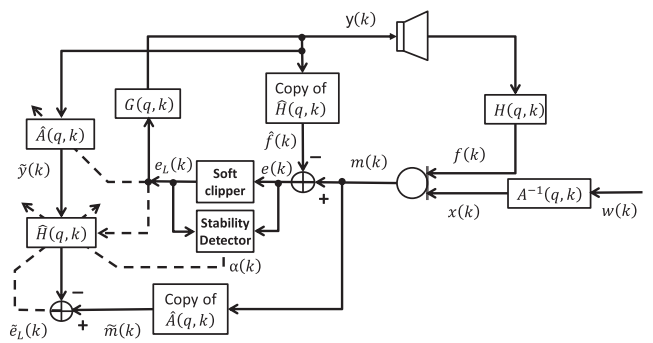

FIG. 3. Block diagram of the proposed H-NLMS algorithm with a soft-clipping-based stability detector.

sizes to a hybrid PEMSC-NLMS algorithm, H-PEMSC-NLMS, given by

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + [1-\alpha(k)]\mu_1(k)\tilde{\mathbf{y}}(k)\tilde{e}_L(k) + \alpha(k)\mu_2(k)\tilde{\mathbf{y}}(k)\tilde{e}_L(k). \quad (16)$$

We call it a hybrid method to be consistent with H-NLMS, in fact this idea is closer to an affine combination method such as presented in Refs. 21–23.

To estimate the AR coefficients $\hat{A}(q,k)$ required for the PEM-NLMS update, we will use the error signal $e_L(k)$ after the SC, since $e_L(k) \simeq x(k)$ when the feedback canceller has converged (cf. design of SC in Sec. II C). Please note that the H-NLMS update in Eq. (15) and the H-PEMSC-NLMS update in Eq. (16) also allow for a non-binary (continuous) control signal, but the investigation of that case is outside the scope of this paper.

C. Soft-clipping-based stability detector

This section discusses the stability detector in Fig. 3, yielding the control signal $\alpha(k)$ which is required for the H-NLMS update in Eq. (15) and the H-PEMSC-NLMS update in Eq. (16). The proposed stability detector is based on a SC on the error signal and a decision threshold. This SC is instantaneous and operates directly on the level of the error signal.

Instead of having an unknown non-linearity due to the loudspeaker characteristics in the feedback loop, it has been shown in Ref. 13 that for identification purposes it is better to use a controlled non-linearity, e.g., a SC, on the loudspeaker output signal, i.e.,

$$y_L(k) = \beta_y \tanh\left(\frac{y(k)}{\beta_y}\right), \quad (17)$$

with β_y a scaling parameter. Since the non-linearity is known, the AFC can be kept linear, improving the feedback cancellation performance.[13] Instead of applying soft clipping to the loudspeaker output signal, it is also possible to apply soft clipping to the error signal (cf. Figs. 2 and 3), i.e.,

$$e_L(k) = \beta \tanh\left(\frac{e(k)}{\beta}\right). \quad (18)$$

The scaling parameter β determines how the error signal $e(k)$ is scaled to the linear range of the tanh-function [cf. Fig. 4(a)]. From this figure, we can observe that different values of β provide a different range for the linear mapping between $e(k)$ and $e_L(k)$ and hence the amount of non-linear distortion. Please note that under the assumption that $G(q, k)$ is a broadband gain G_0, Eqs. (17) and (18) are equivalent with $\beta_y = G_0\ \beta$. However, since the scaling parameter β_y depends on the hearing aid processing $G(q, k)$, which is not the case for β, we will use the SC on the error signal in the remainder of this paper.

Note that the error signal is given by $e(k) = x(k) + \Delta f(k)$, with $\Delta f(k) = f(k) - \hat{f}(k)$. On the one hand, when the AFC has converged, $\Delta f(k)$ is small, such that the error signal $e(k)$ is mainly determined by the incoming signal $x(k)$. On the other hand, when the AFC has not converged or the system is (conditionally) unstable, $\Delta f(k)$ will be dominant or at least not small. Hence, when the scaling parameter β is chosen such that the most likely range of the incoming signal $x(k)$ lies in the linear range of the tanh-function, i.e., $x(k) \simeq \beta\tanh(x(k)/\beta)$, then $e(k) - e_L(k)$ may be used to detect instability or misconvergence of the AFC. Thus, we propose to compute the control signal $\alpha(k)$ in Eq. (15) using the following soft-clipping-based stability detector:

$$\alpha(k) = I\{|e(k) - e_L(k)| < \gamma\}, \quad (19)$$

where I is a function returning 0 or 1 depending on whether the inequality holds. γ is a decision threshold determining the sensitivity of the detector. Figure 4(b) shows a plot of Eq. (19) for different values of β and γ. It can be seen that β has an impact on the linear region of the function and γ determines the sensitivity of detection.

The choice of the parameters β and γ depends on multiple considerations. First, the scaling parameter β determines the linear range of the tanh-function and hence the amount of non-linear distortion of the incoming signal. Second, the decision threshold γ determines the sensitivity of the detector, selecting either the NLMS update or the PEMSC-NLMS update in Eq. (15). It should however be realized that the choice of the decision threshold is not very critical. False detection of stability, i.e., $\alpha(k) = 0$ when the system is unstable, will lead to usage of the PEMSC-NLMS update, slowing down recovery from howling. False detection of instability, i.e., $\alpha(k) = 1$ when the system is stable, will lead to usage of

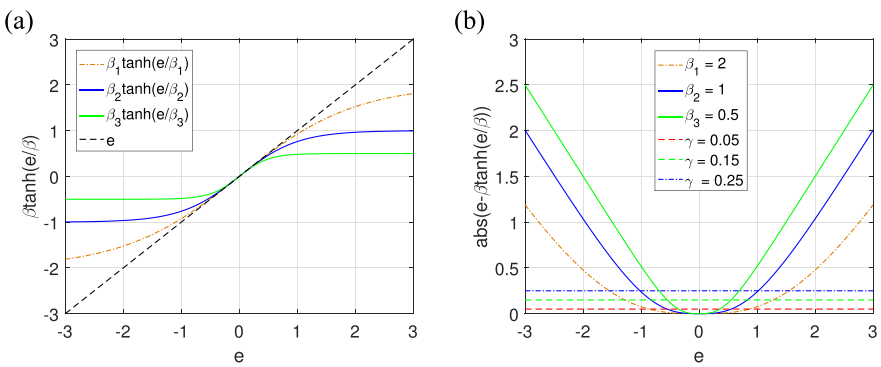

FIG. 4. (Color online) (a) SC with different values of scaling parameter β ($\beta_1 = 2$, $\beta_2 = 1$, $\beta_3 = 0.5$), and (b) detection equation with different values of scaling parameter β and decision threshold γ.

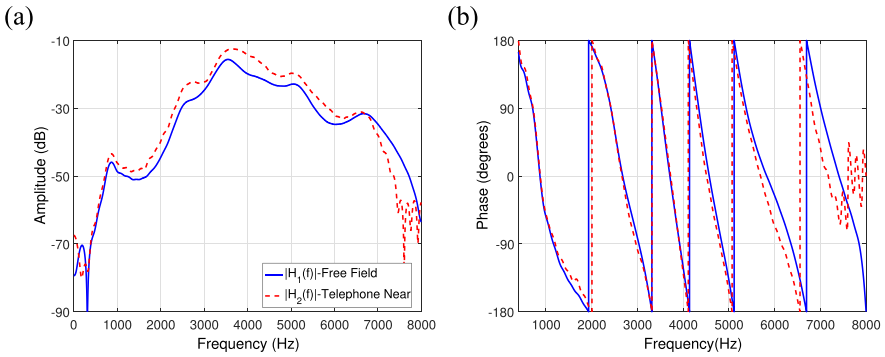

FIG. 5. (Color online) HAD acoustic feedback paths: (a) amplitude response and (b) phase response.

the NLMS update, possibly increasing the bias. The influence of different parameter choices on the performance of the H-NLMS algorithm and a comparison with the H-PEMSC-NLMS algorithm will be studied in Sec. III B. In particular, we will compare the performance for different levels of the incoming signal $x[k]$ (both weak as well as outside of the linear range) and forward path gains.

III. EXPERIMENTS AND EVALUATIONS

In this section, the performance of the proposed stability-controlled H-NLMS algorithm is evaluated and compared to the PEMSC-NLMS and H-PEMSC-NLMS algorithms. Section III A presents the used acoustic setup, the considered performance measures, and the algorithmic settings. Section III B discusses the evaluation using instrumental performance measures, while Sec. III C discusses the formal listening test.

A. Acoustic setup, performance measures, and algorithmic settings

Acoustic feedback paths were measured on a dummy head with adjustable ear canals[28] using a two-microphone behind-the-ear hearing aid and open-fitting earmolds.[29] The measured impulse responses (IRs) were sampled at $f_s = 16\,\text{kHz}$ and truncated to length $L_h = 100$. Figure 5 depicts the amplitude and phase responses of the IRs used in the evaluation: $\mathbf{h}_1$ measured in free-field, and $\mathbf{h}_2$ measured with a telephone receiver in close distance to the HAD. As incoming signal $x(k)$, we used a speech signal consisting of female and male speech from the NOIZEUS database[30] and a music signal ("Imagine" by John Lennon), which has been normalized such that the maximum absolute value is equal to 1. The normalized incoming signal has been scaled to different values to, e.g., reflect a speaker that is closer or further away, respectively. Both speech and music signals were recorded using a microphone placed in the right ear of a dummy head in an anechoic chamber for two different sound source positions (in front and on the right side of the dummy head), resulting in four audio signals (speech 1, speech 2, music 1, music 2). The reason for using both speech and music signals is that speech signals can generally be well modeled using the AR model in Eq. (10), whereas for music signals there is a model mismatch. All signals were 50 s long and an instant change of the acoustic feedback path was simulated after 25 s by switching from $\mathbf{h}_1$ to $\mathbf{h}_2$.

As instrumental performance measures to evaluate the performance of the AFC algorithms, we used the normalized MIS and the ASG. The normalized MIS is defined as the normalized Euclidean distance between the measured and the estimated IR, i.e.,

$$\text{MIS} = 10\log_{10}\left(\frac{||\mathbf{h}-\hat{\mathbf{h}}||_2^2}{||\mathbf{h}||_2^2}\right), \tag{20}$$

while the ASG is defined as[1,11]

$$\text{ASG} = 10\log_{10}\frac{1}{\max_\Omega|H(\Omega)-\hat{H}(\Omega)|^2} - 10\log_{10}\frac{1}{\max_\Omega|H(\Omega)|^2}, \tag{21}$$

with $H(\Omega)$ and $\hat{H}(\Omega)$ the frequency response of the measured and estimated acoustic feedback paths at the normalized frequency Ω, respectively. The average MIS, ($\overline{MIS}$), and the average ASG, ($\overline{ASG}$), are computed by averaging the

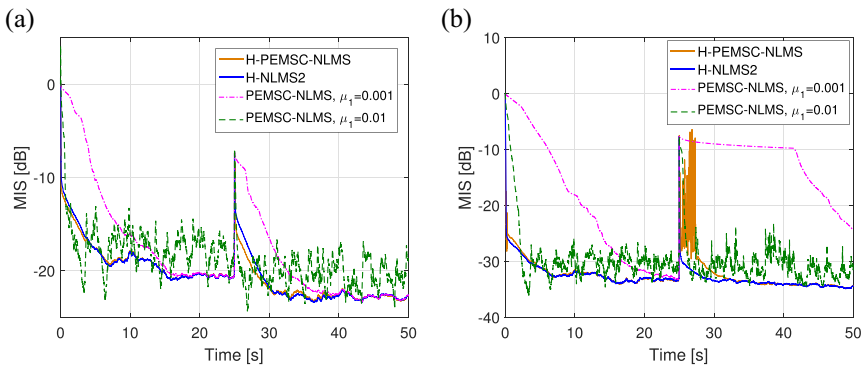

FIG. 6. (Color online) MIS of the PEMSC-NLMS algorithm for two step-sizes, the H-NLMS2 algorithm and the H-PEMSC-NLMS algorithm with speech 1 as the incoming signal with (a) $G_0 = 30\,\text{dB}$ and (b) $G_0 = 45\,\text{dB}$.

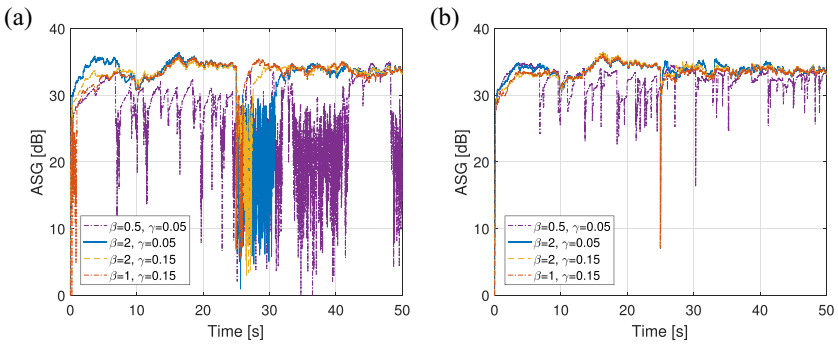

FIG. 7. (Color online) ASG of the H-PEMSC-NLMS algorithm in (a) and the H-NLMS2 algorithm in (b), for different values of β and γ with speech 1 as the incoming signal and $G_0 = 45$ dB.

respective performance measures over the complete signal. In addition, as instrumental performance measures for speech quality and audio quality, we used the perceptual evaluation of speech quality (PESQ) measure from Ref. 30 and the perceptual evaluation of audio quality (PEAQ) measure.[31] The reference signal for both PESQ and PEAQ measures was the incoming signal $x(k)$, while the test signal was the error signal after soft clipping $e_L(k)$.

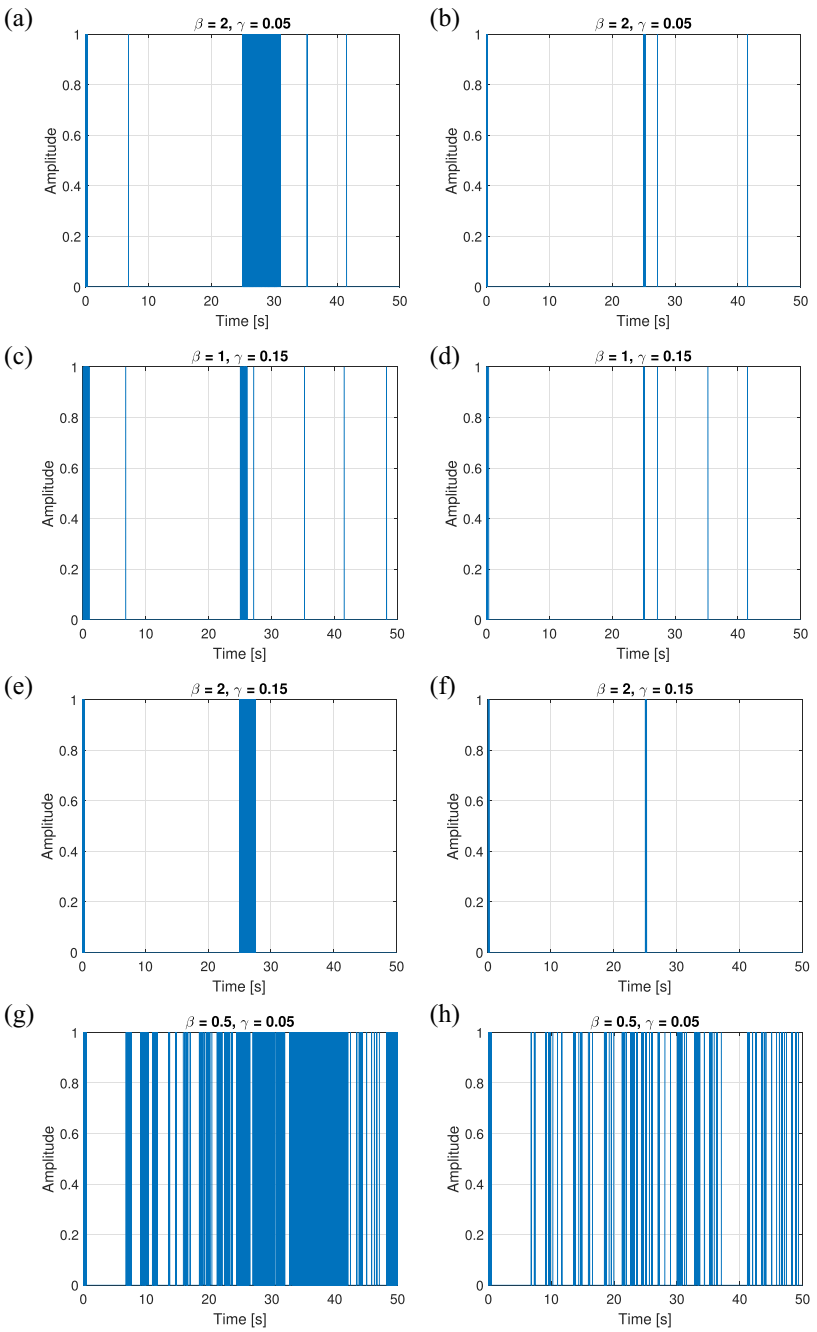

FIG. 8. (Color online) Control signal $\alpha(k)$ corresponding to different values of β and γ for the H-PEMSC-algorithm (left column) and the H-NLMS2 algorithm (right column) with speech 1 as the incoming signal and $G_0 = 45$ dB.

The following algorithmic settings were used for all evaluations. The prediction-error filter $\hat{A}(q,k)$ was of order 20 and was updated using the Levinson-Durbin method with blocks of 10 ms. The feedforward path of the HAD was set to $G(q,k) = G_0 \exp(-j\Omega d_G)$ with d_G corresponding to a delay of 6 ms and G_0 set to 30 or 45 dB. The feedback canceller path was delayed by one sample. The length of the adaptive filter was $L_{\hat{h}} = 64$.

To ensure a fair comparison, the proposed H-NLMS algorithm is compared to the PEMSC-NLMS algorithm, i.e., the PEMSC-NLMS algorithm can be considered a special case of the H-NLMS algorithm with $\mu_2 = 0$ in Eq. (15) or $\gamma = \infty$ in Eq. (19). In addition, the proposed H-NLMS algorithm is compared with the H-PEMSC-NLMS algorithm in Eq. (16).

The following step-sizes are used to update the adaptive filters:

- PEMSC-NLMS algorithm: step-sizes $\mu_1 \in [0.001, 0.01]$;
- H-NLMS algorithm: three different pairs of step-sizes, i.e.,
  - H-NLMS1: $\mu_1 = 0.001$, $\mu_2 = 0.1$;
  - H-NLMS2: $\mu_1 = 0.001$, $\mu_2 = 0.2$;
  - H-NLMS3: $\mu_1 = 0.001$, $\mu_2 = 0.5$;
- H-PEMSC-NLMS algorithm: step-sizes $\mu_1 = 0.001$, $\mu_2 = 0.2$.

B. Instrumental evaluation

1. Speech signals

The instrumental evaluation is designed to illustrate differences in performance between different algorithms under parameter and feedback path changes. We have studied the following parameters: algorithm step-size, forward path gain, volume of the incoming signal, and SC parameters.

In the first experiment, we compare the performance of the PEMSC-NLMS algorithm with the H-NLMS2 and H-PEMSC-NLMS algorithms with normalized speech 1 as the incoming signal for two different gains in the forward path, namely $G_0 = 30$ dB and $G_0 = 45$ dB. Figure 6 shows the MIS of the PEMSC-NLMS algorithm for two different step-sizes $\mu_1 \in [0.01, 0.001]$ as well as the H-NLMS2 and H-PEMSC-NLMS with $\beta = 2$ and $\gamma = 0.15$. Note that the step-sizes for the H-NLMS2 and H-PEMSC-NLMS algorithms are equal ($\mu_1 = 0.001$ and $\mu_2 = 0.2$). We first consider the lower gain $G_0 = 30$ dB in Fig. 6(a). As expected for the PEMSC-NLMS algorithm, there is a trade-off between convergence rate (both initially as well as when the feedback path changes after 25 s) and steady-state MIS, i.e., fast convergence but high steady-state MIS are obtained for a large step-size, whereas slow convergence but low steady-state MIS are obtained for a small step-size. In order to obtain a low steady-state MIS, a step-size $\mu_1 = 0.001$ will be used in the following for the PEMSC-NLMS algorithm. It can also be seen that both the H-NLMS2 algorithm and the H-PEMSC-NLMS algorithm have a very similar performance. However, when the gain is increased to $G_0 = 45$ dB [Fig. 6(b)], it can be seen that the PEMSC-NLMS algorithm with the small step-size has difficulty to converge and re-converge, most likely also leading to the H-PEMSC-NLMS algorithm having difficulty to re-converge after a feedback path change. The only algorithm which shows both good convergence, re-convergence, and low MIS for both gains is the proposed H-NLMS2 algorithm.

In the second experiment, we compare the performance of the H-PEMSC-NLMS algorithm and the proposed H-NLMS2 algorithm for different values of the scaling parameter β and the decision threshold γ of the stability detector (cf. Sec. II C). Figure 7 shows the ASG for the H-PEMSC-

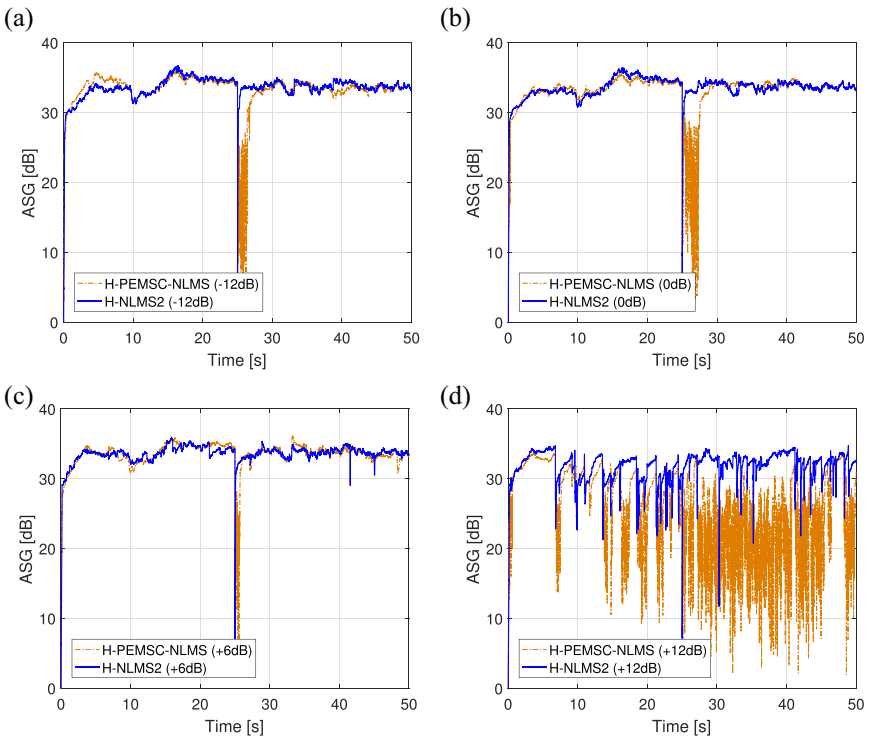

FIG. 9. (Color online) ASG for the H-NLMS2 and H-PEMSC-NLMS algorithms ($\mu_1 = 0.001$, $\mu_2 = 0.2$, $\beta = 2$, $\gamma = 0.15$), for different volume levels $[-12, 0, +6, +12]$ dB with speech 1 as the incoming signal and $G_0 = 45$ dB.

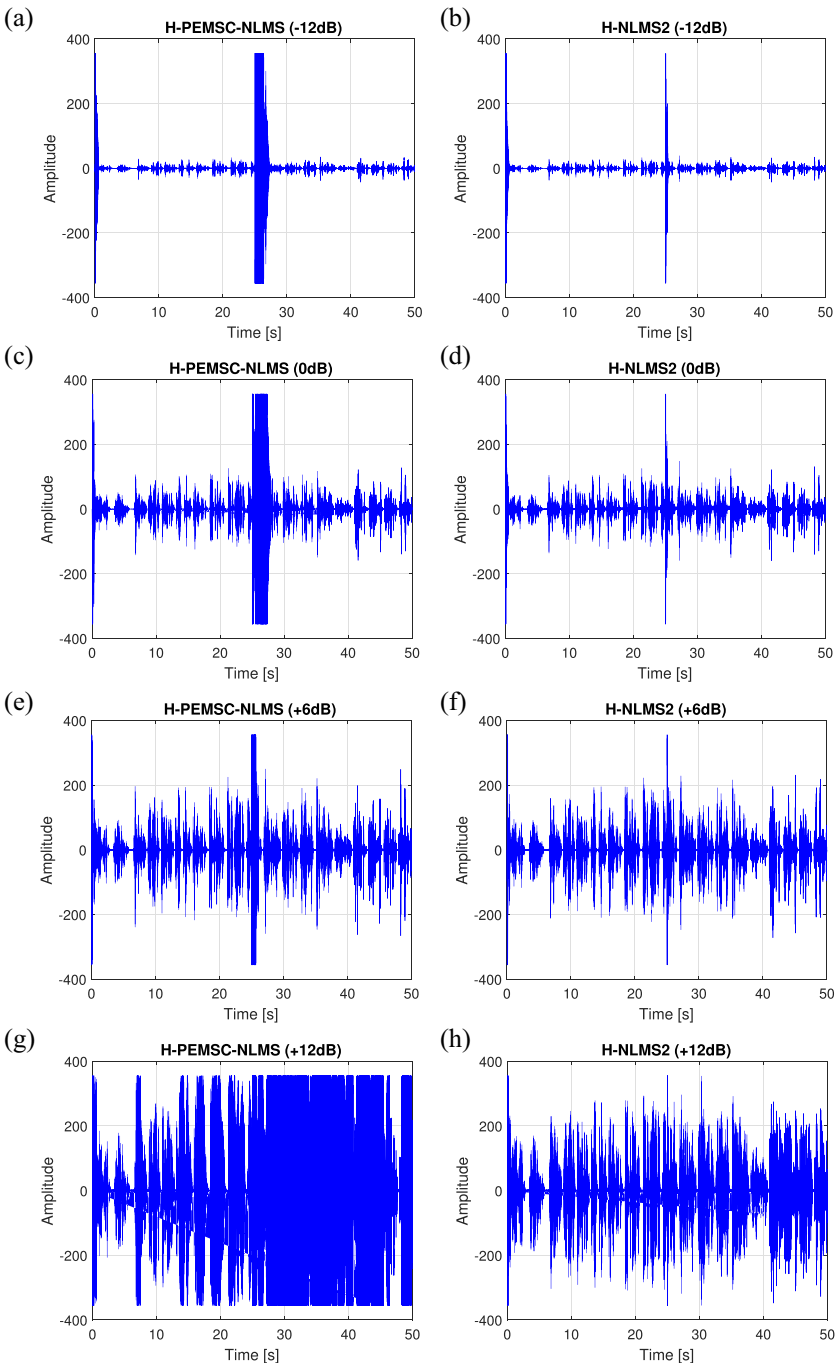

FIG. 10. (Color online) Loudspeaker output signals for different volume levels [−12, 0, +6, +12] dB with speech 1 as the incoming signal and $G_0 = 45$ dB for the H-PEMSC-NLMS algorithm (left column) and the H-NLMS2 algorithm (right column).

NLMS [Fig. 7(a)] and H-NLMS2 [Fig. 7(b)] algorithm ($\mu_1 = 0.001$, $\mu_2 = 0.2$) with speech 1 as the incoming signal and forward path gain $G_0 = 45$ dB. It can be clearly observed that β and γ have a large impact on the ASG. This is especially prominent for the H-PEMSC-NLMS algorithm, whereas the H-NLMS2 algorithm is more robust toward changes in these parameters. Since the linear input range of the SC depends on the scaling parameter β, a different

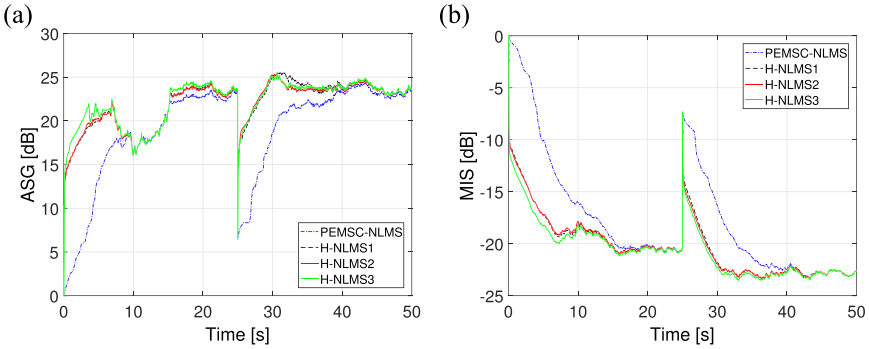

FIG. 11. (Color online) ASG and MIS for the PEMSC-NLMS and H-NLMS algorithms with speech 1 as the incoming signal, $\beta = 2$, $\gamma = 0.15$, $\mu_1 = 0.001$, and $\mu_2 \in [0.1, 0.2, 0.5]$.

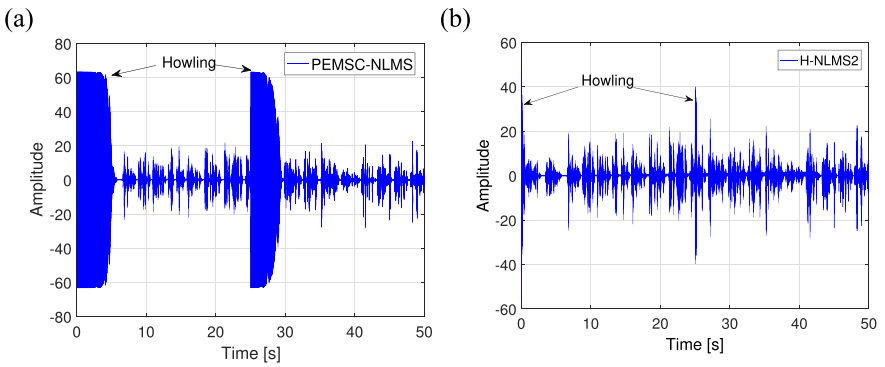

FIG. 12. (Color online) Loudspeaker output signals with speech 1 as the incoming signal: (a) PEMSC-NLMS with $\mu_1 = 0.001$; (b) H-NLMS2.

choice of β requires a different choice of the decision threshold γ. Figure 8 presents the corresponding control signals $\alpha(k)$ for the H-PEMSC-NLMS algorithm in the left column and the H-NLMS algorithm in the right column, respectively. When working properly, the stability detector should give a small number of false detections, both in terms of stability and instability. On the one hand, when γ is too sensitive, i.e., too small, many false detections of instability occur [cf. Fig. 8(h)], such that the NLMS update with $\mu_2 = 0.2$ is selected inadvertently in the H-NLMS update in Eq. (15) and the ASG suffers [cf. Fig. 8(b)]. For the same parameters a similar behavior can be found for the H-PEMSC-NLMS [cf. Fig. 8(g)] but even more accentuated due to the PEM estimate starting to model the feedback path and not the speech source. On the other hand, when γ is not sensitive enough, i.e., too large, howling might not be detected, such that the PEMSC-NLMS update is selected in Eq. (15) and recovery from howling takes a long time (or in the worst case does not happen at all). The H-PEMSC-NLMS shows a larger sensitivity to the selection of the parameters β and γ. However, for both algorithms $\beta = 2$ and $\gamma = 0.15$ gives overall good performance. Accordingly, for all following simulations, these parameter values have been chosen since they seem to provide no false detections, cf. Figs. 8(e) and 8(f).

In the third experiment we compare the performance of the H-NLMS2 algorithm and the H-PEMSC-NLMS algorithm for different volume levels $[-12, 0, +6, +12]$ dB of the incoming signal. The volume level is given relative to the normalized original signal. The incoming signal is speech 1 and the forward gain is $G_0 = 45$ dB. From the ASG in Fig. 9 it can be seen that the H-NLMS2 algorithm outperforms the H-PEMSC-NLMS algorithms for all considered volume levels, particularly when the feedback path changes. For high levels of the incoming signal [e.g., + 12 dB in Fig. 9(d)] false detection of instability due to the soft-clipping will happen more often than for low levels of the incoming signal. As can be observed, the ASG of the H-NLMS algorithm for the high-level incoming signal in Fig. 9(d) is worse (and more peaky) than the ASGs for lower-level incoming signals, cf. Figs. 9(a)–9(c). Nevertheless, the obtained ASG is still very high and the proposed H-NLMS algorithm outperforms the H-PEMSC-NLMS algorithm (which is a hybrid combination of two PEMSC-NLMS updates), especially when the feedback path changes. The corresponding loudspeaker signals are shown in Fig. 10. It can be observed that the H-NLMS algorithm recovers fast from howling, even when the incoming signal is 12 dB larger than the nominally designed value, see Fig. 10(h), whereas the H-PEMSC-NLMS, see Fig. 10(g), does not re-converge after the feedback path changes. This can be explained by the fact that the H-PEMSC-NLMS only uses PEMSC-NLMS updates with two different step sizes and when there is a strong feedback the PEM starts to model the howling instead of the incoming signal $x[k]$. Therefore, the H-PEMSC-NLMS algorithm will be slow to re-converge, whereas for the proposed H-NLMS2 algorithm the initial adaptation using the NLMS update will directly suppress the howling and then switch to the PEMSC-NLMS update.

In the following simulations we have only compared the H-NLMS and PEMSC-NLMS algorithms since it was established that the H-NLMS algorithm outperforms or at least performs equally well as the H-PEMSC-NLMS algorithm for all considered scenarios. Figure 11 compares the ASG and the MIS of the PEMSC-NLMS algorithm ($\mu_1 = 0.001$) and the H-NLMS algorithm ($\mu_1 = 0.001$) for three different values of the step-size μ_2 with speech 1 as the incoming signal. It can be observed that the H-NLMS algorithm converges much faster than the PEMSC-NLMS algorithm, while maintaining a similar steady-state MIS. It also provides a faster tracking rate when the feedback path changes. After convergence, the H-NLMS algorithm provides a

TABLE I. Evaluation for different speech and music incoming signals.

AFC methods	Incoming signals	PESQ	$\overline{ASG}$	$\overline{MIS}$
PEMSC-NLMS	Speech 1	1.66	18.88	−17.52
H-NLMS1		4.14	21.93	−20.22
H-NLMS2		4.09	21.85	−20.20
H-NLMS3		4.15	22.14	−20.53
PEMSC-NLMS	Speech 2	1.90	20.48	−18.66
H-NLMS1		3.97	22.32	−20.59
H-NLMS2		3.6	22.29	−20.65
H-NLMS3		3.78	22.40	−20.82
PEMSC-NLMS	Music 1	N/A	15.62	−13.57
H-NLMS1		N/A	17.36	−15.14
H-NLMS2		N/A	17.39	−15.24
H-NLMS3		N/A	17.37	−15.14
PEMSC-NLMS	Music 2	N/A	15.49	−13.31
H-NLMS1		N/A	17.71	−14.77
H-NLMS2		N/A	17.77	−14.79
H-NLMS3		N/A	17.83	−14.87

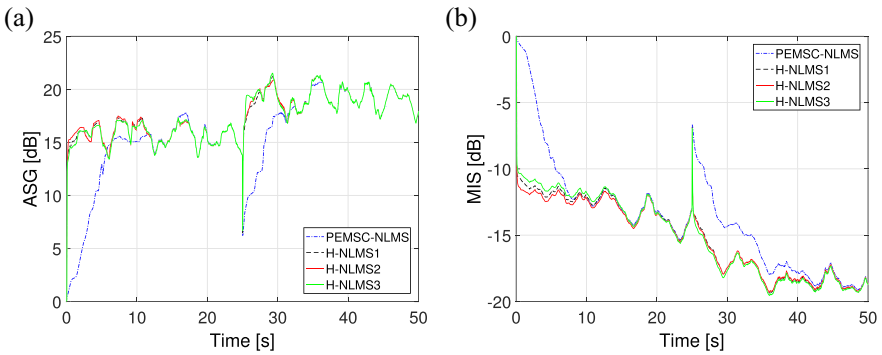

FIG. 13. (Color online) ASG and MIS for the PEMSC-NLMS and H-NLMS algorithms with music 1 as the incoming signal, $\beta = 2$, $\gamma = 0.15$, $\mu_1 = 0.001$, and $\mu_2 \in [0.1, 0.2, 0.5]$.

similar ASG as the PEMSC-NLMS algorithm. Note that the minor differences can be related to the fact that only a single realization is used and the weights do not have the same initial conditions after switching in the H-NLMS. Furthermore, the usage of a different step-size μ_2 in the NLMS update has almost no impact on the results.

Figure 12 illustrates the loudspeaker output signals for the PEMSC-NLMS algorithm and the H-NLMS2 algorithm with speech 1 as the incoming signal. It can be observed that for the PEMSC-NLMS algorithm a long howling period occurs in the initial phase and when the feedback path changes. By using the proposed switched combination adaptive filter with stability detector, the H-NLMS algorithm can dramatically shorten the howling period, resulting in a significant improvement in speech signal quality. Moreover, Fig. 12(b) shows that the howling occurrence in the loudspeaker signal matches well with the control signal $\alpha(k)$ for this case as presented in Fig. 8(f).

For both speech signals (speech 1, speech 2) Table I shows the PESQ scores, the average ASGs, $\overline{ASG}$, and the average MISs, $\overline{MIS}$, for the PEMSC-NLMS algorithm ($\mu_1 = 0.001$) and the H-NLMS algorithm (different pairs of step-sizes). The results show that by selecting an adequate pair of step-sizes the H-NLMS algorithm is able to achieve a considerable improvement in PESQ (about 2 points), ASG (about 2–3 dB) and MIS (about 2–3 dB) compared to the PEMSC-NLMS algorithm. Among the three considered pairs of step-sizes, $\mu_1 = 0.001$ and $\mu_2 = 0.5$ (H-NLMS3) seem to yield the best performance.

2. Music signals

We have also performed similar evaluations using music as the incoming signal. But due to lack of space no comparisons between the H-NLMS algorithm and the H-PEMSC-NLMS algorithm have been presented. However, from our extensive simulations using music 1 and music 2 as incoming signals very similar results have been obtained as for speech. Figure 13 compares the ASG and the MIS of the PEMSC-NLMS algorithm and the H-NLMS algorithm (different pairs of step-sizes) when music 1 is used as the incoming signal. All algorithmic parameters are set to the same values as for the speech signal. It can be clearly observed that the H-NLMS algorithm outperforms the PEMSC-NLMS algorithm. Especially when the feedback path changes from free-field to a telephone receiver in close distance, the H-NLMS algorithm is able to track the feedback path change much faster than the PEMSC-NLMS algorithm.

Figure 14 illustrates the loudspeaker output signals with music 1 as the incoming signal. The howling periods for the H-NLMS algorithm are clearly much shorter than for the PEMSC-NLMS algorithm. Similarly as for the speech signal, Fig. 14(c) shows that the control signal $\alpha(k)$ matches well with the howling occurrence in the loudspeaker signal presented in Fig. 14(b).

For both music signals (music 1, music 2), the results in Table I show that the proposed H-NLMS algorithm yields a much better average ASG, $\overline{ASG}$, and average MIS, $\overline{MIS}$, than the PEMSC-NLMS algorithm. To evaluate the perceptual quality for music signals, we have used the PEAQ measure, where a larger PEAQ score indicates a better signal integrity. Six different signal segments, covering the initial convergence phase (0–1 s, 1–10 s), the re-convergence phase (25–26.5 s, 26.5–32 s), and the steady-state phase (10–25 s, 32–50 s) have been evaluated and the results are given in Table II. It can be observed that the PEAQ score is approximately equal to -4 when howling occurs, but it increases when the system begins to converge. The results also show that for segments 1–10 s and 26.5–32 s the proposed H-NLMS algorithm achieves a significant improvement in PEAQ compared to the PEMSC-NLMS algorithm, corresponding to a shorter howling duration. When the system has converged (corresponding to segments 10–25 s and 3250 s), both algorithms reach a similar steady-state error,

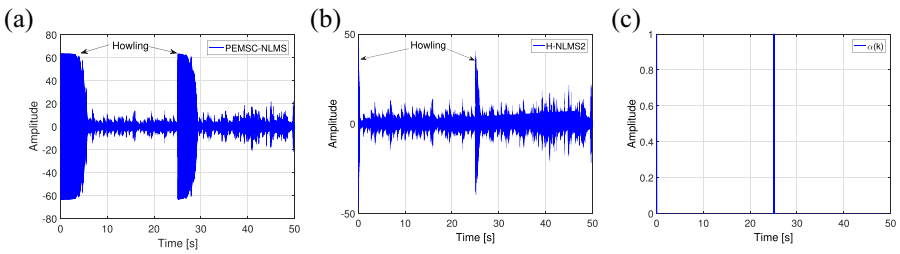

FIG. 14. (Color online) Loudspeaker output signals with music 1 as the incoming signal: (a) PEMSC-NLMS with $\mu_1 = 0.001$; (b) H-NLMS2; (c) control signal $\alpha(k)$.

TABLE II. EAQ score for music signals.

AFC methods	Incoming signals	0–1 s	1–10 s	10–25 s	25–26.5 s	26.5–32 s	32–50 s
PEMSC-NLMS	Music 1	−3.91	−3.90	−2.13	−3.91	−3.90	−2.39
H-NLMS1		−3.91	−2.16	−2.14	−3.91	−2.38	−2.32
H-NLMS2		−3.90	−2.14	−2.13	−3.90	−2.38	−2.33
H-NLMS3		−3.91	−2.17	−2.13	−3.90	−2.34	−2.34
PEMSC-NLMS	Music 2	−3.91	−3.90	−2.27	−3.90	−3.90	−2.53
H-NLMS1		−3.86	−2.17	−2.21	−3.89	−2.48	−2.49
H-NLMS2		−3.86	−2.19	−2.24	−3.89	−2.49	−2.48
H-NLMS3		−3.91	−2.19	−2.20	−3.88	−2.41	−2.48

resulting in similar PEAQ scores. As will be shown in Sec. III C, these scores match with the subjective listening test results.

In conclusion, the proposed stability-controlled hybrid algorithm provides a stable and well-functioning solution also for the considered music signals. However, since the instrumental performance measures do not necessarily reflect all distortions introduced by the algorithms, it is particularly important to also perform a perceptual evaluation.

C. Perceptual evaluation

In this section, the perceptual quality of signals processed by the proposed H-NLMS algorithm is evaluated and compared to the quality of signals processed by the PEMSC-NLMS algorithm using a formal listening test. Previously, the perceptual quality of AFC algorithms has been evaluated in different studies, e.g., Refs. 7, 9, 27, and 32, where different scales have been used to assess the perceived quality. Similar to the study in Ref. 27, the Multi-Stimulus Test with Hidden Reference and Anchor (MUSHRA)[33] is employed here.

1. Method

a. Subjects. $N = 15$ self-reported normal-hearing subjects participated in the listening test. The mean age of the subjects was 28.1 yrs. The subjects participated voluntarily and were paid a small compensation for their time.

b. Equipment. The evaluations were conducted using a personal computer and MATLAB software. The stimuli were pre-computed and stored on a hard-drive prior to the listening test. An (RME Audio, Haimhausen, Germany, Fireface Babyface Soundcard) was used and the signals were presented via (Sennheiser electronic GmbH & Co., KG Wedemark, Germany, HDA200) headphones in a quiet office room. All stimuli were sampled at 16 kHz. The sound signals were calibrated to a level of 60 dB sound pressure level (SPL) for the reference signals. Note that none of the presented signals exceeded a level of 80 dB SPL.

c. Procedure. In order to assess the quality of the processed speech and music signals, the MUSHRA framework is used. The task of the subjects was to judge the quality of the processed signals with respect to a reference signal. The reference signal was chosen to be the signal processed with an

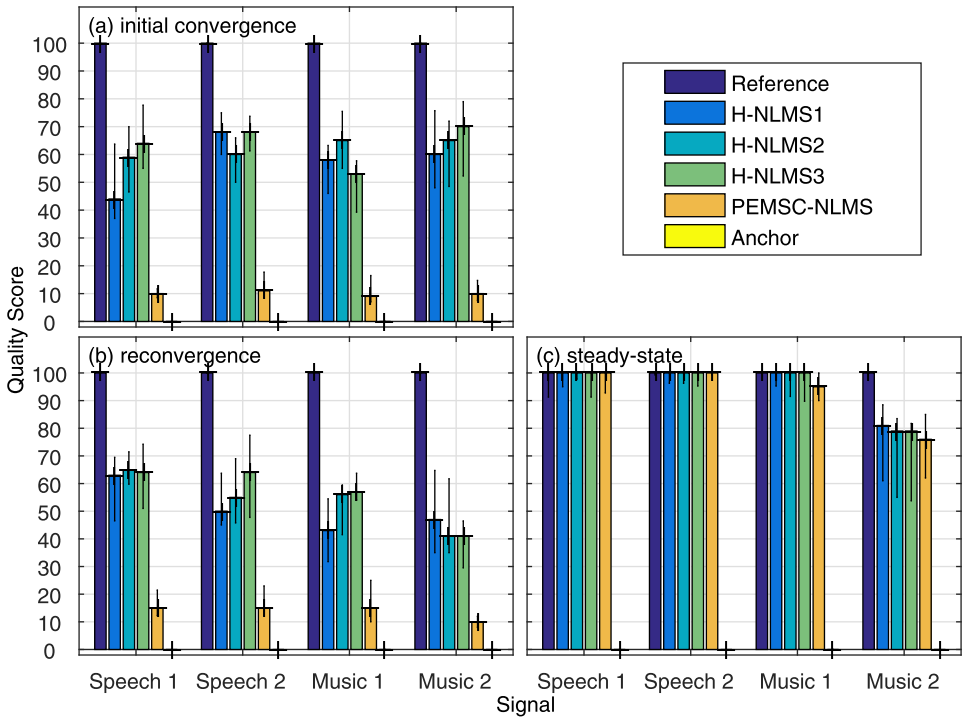

FIG. 15. (Color online) Median quality scores and interquartile ranges of the formal listening test for different audio signals and algorithms for (a) the initial convergence, (b) the reconvergence, and (c) steady-state performance.

ideal feedback cancellation algorithm, which was also hidden as a test signal. From each of the four considered signals (speech 1, speech 2, music 1, music 2) 3 segments with a duration of 10 s were selected that included (a) the initial convergence phase (0–10 s), (b) the re-convergence phase (22–32 s), and (c) the steady-state phase (35–45 s). This led to a total of 12 different audio files the subjects compared the quality of. In total six different algorithmic settings were judged for each audio file, where in addition to the algorithmic settings presented in Sec. III A the following settings were considered:

- Reference: ideal feedback cancellation, i.e., $\hat{\mathbf{h}}(k) = \mathbf{h}(k)$;
- Anchor: no feedback cancellation, i.e., $\hat{\mathbf{h}}(k) = \mathbf{0}$.

In order to allow for a similar use of the rating scale, the subjects were instructed to rate at least one of the settings with a score of 100 and at least one other setting with a score of 0. Since the sound level of the presented signals is strongly affected by the amount of howling, i.e., the sound level may be much larger than the calibrated 60 dB SPL, all signals were limited in level prior to presentation. The limit level, i.e., the maximum amplitude of the played back signal, was chosen such that the amplitude limited signals could not be distinguished from the same signals without limiting as determined in a preliminary listening test with five subjects.

d. Statistical analysis. A statistical analysis was conducted using R statistics software. Since Shapiro-Wilk tests showed that not all data could be assumed to be normally distributed, an aligned rank transform (ART)[34] was employed before using standard analysis of variance (ANOVA) procedures. Separate two-way ANOVAs were conducted for each of the three segments with factors algorithm and signal. For a two-factor data-set the ART produces three different data-sets, i.e., one for each factor and one for their two-way interaction. Each data-set is then assumed to

TABLE III. Paired comparisons for the initial convergence for (a) speech 1, (b) speech 2, (c) music 1, and (d) music 2. Asterisks indicate significant differences, i.e., $p < 0.05$, after Bonferroni correction.

(a)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		*	*	*	*
H-NLMS2	*	*		n.s.	*	*
H-NLMS3	*	*	n.s.		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

(b)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		*	n.s.	*	*
H-NLMS2	*	*		n.s.	*	*
H-NLMS3	*	n.s.	n.s.		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

(c)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		*	*	*	*
H-NLMS2	*	*		*	*	*
H-NLMS3	*	*	*		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

(d)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		n.s.	n.s.	*	*
H-NLMS2	*	n.s.		n.s.	*	*
H-NLMS3	*	n.s.	n.s.		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

only depend on either one of the main factors or the interaction. For each of these data-sets, a two-way ANOVA is carried out while only the results for the dependent factor may be interpreted.[34] *Post hoc* analyses were carried out (if appropriate) using the Wilcoxon-sign-rank test on the original data. Differences were assumed to be significant for p-values smaller than 0.05 and the level of significance was adjusted using Bonferroni correction when multiple comparisons were conducted.

2. Results

The results of the listening test for the different audio signals, algorithms, and signal segments are shown in Fig. 15. Figure 15(a) depicts the median quality scores and interquartile ranges for the evaluation of the initial convergence. All subjects were able to identify both the hidden reference and the anchor signal. It can be observed that the proposed H-NLMS algorithm in general leads to an improved quality compared to the PEMSC-NLMS algorithm. Depending on the audio signal, different step-size settings of the H-NLMS algorithm show the highest median quality score. A two-way ANOVA after ART was computed to assess the statistical significance of the results, showing significant effects of the factors algorithm and signal and their interaction (algorithm: $F^{5,70} = 252.93$, $p < 0.05$; signal: $F^{3,42} = 16.53$, $p < 0.05$; algorithm $\times$ signal: $F^{15,210} = 11.11$, $p < 0.05$). Due to the significant interactions, *post hoc* analyses were carried for each signal separately. Table III shows the results of the *post hoc* analyses, where asterisks indicate significant differences, i.e., $p < 0.05$, after Bonferroni correction.

Figure 15(b) depicts the median quality scores and interquartile ranges for the evaluation of the re-convergence. All subjects were able to identify both the hidden reference and the anchor signal. It can be observed that the proposed H-NLMS algorithm in general leads to an improved quality

TABLE IV. Paired comparisons for the re-convergence for (a) speech 1, (b) speech 2, (c) music 1, and (d) music 2. Asterisks indicate significant differences, i.e., $p < 0.05$, after Bonferroni correction.

(a)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		n.s.	n.s.	*	*
H-NLMS2	*	n.s.		n.s.	*	*
H-NLMS3	*	n.s.	n.s.		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

(b)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		n.s.	n.s.	*	*
H-NLMS2	*	n.s.		n.s.	*	*
H-NLMS3	*	n.s.	n.s.		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

(c)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		*	n.s.	*	*
H-NLMS2	*	*		n.s.	*	*
H-NLMS3	*	n.s.	n.s.		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

(d)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		n.s.	n.s.	*	*
H-NLMS2	*	n.s.		n.s.	*	*
H-NLMS3	*	n.s.	n.s.		*	*
PEMSC-NLMS	*	*	*	*		*
ANCHOR	*	*	*	*	*	

compared to the PEMSC-NLMS algorithm. Depending on the audio signal, different step-size settings of the H-NLMS algorithm show the highest median quality score. A two-way ANOVA after ART was computed to assess the statistical significance of the results, showing significant effects of the factors algorithm and signal and their interaction (algorithm: $F^{5,70} = 285.97$, $p < 0.05$; signal: $F^{3,42} = 29.17$, $p < 0.05$; algorithm $\times$ signal: $F^{15,210} = 12.01$, $p < 0.05$). Similarly as before, due to the significant interactions, *post hoc* analyses were carried for each signal separately. Table IV shows the results of the *post hoc* analyses.

Figure 15(c) depicts the median quality scores and interquartile ranges for the evaluation of the steady-state error. Most subjects were able to identify the hidden reference and all subjects were able to identify the anchor signal. It can be observed that for all signals all algorithms in general lead to a similar quality rating. A two-way ANOVA after ART was computed to assess the statistical significance of the results, showing significant effects of the factors algorithm and signal and their interaction (algorithm: $F^{5,70} = 131.52$, $p < 0.05$; signal: $F^{3,42} = 29.92$, $p < 0.05$; algorithm $\times$ signal: $F^{15,210} = 29.23$, $p < 0.05$). Similarly as before, due to the significant interactions, *post hoc* analyses were carried for each signal separately. Table V shows the results of the *post hoc* analyses.

3. Discussion

In this study, the perceived quality for different feedback cancellation algorithms, namely the PEMSC-NLMS algorithm and the H-NLMS algorithm for three different step-sizes, was evaluated using the MUSHRA framework. The listening test was conducted for four different audio signals. While in a previous study[27] the complete signal was rated by the listeners, in this study three different segments of interest were evaluated separately, i.e., the initial convergence, the

TABLE V. Paired comparisons for the steady-state for (a) speech 1, (b) speech 2, (c) music 1, and (d) music 2. Asterisks indicate significant differences, i.e., $p < 0.05$, after Bonferroni correction.

(a)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		n.s.	n.s.	n.s.	n.s.	*
H-NLMS1	n.s.		n.s.	n.s.	n.s.	*
H-NLMS2	n.s.	n.s.		n.s.	n.s.	*
H-NLMS3	n.s.	n.s.	n.s.		n.s.	*
PEMSC-NLMS	n.s.	n.s.	n.s.	n.s.		*
ANCHOR	*	*	*	*	*	

(b)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		n.s.	n.s.	n.s.	n.s.	*
H-NLMS1	n.s.		n.s.	n.s.	n.s.	*
H-NLMS2	n.s.	n.s.		n.s.	n.s.	*
H-NLMS3	n.s.	n.s.	n.s.		n.s.	*
PEMSC-NLMS	n.s.	n.s.	n.s.	n.s.		*
ANCHOR	*	*	*	*	*	

(c)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		n.s.	n.s.	n.s.	n.s.	*
H-NLMS1	n.s.		n.s.	n.s.	n.s.	*
H-NLMS2	n.s.	n.s.		n.s.	n.s.	*
H-NLMS3	n.s.	n.s.	n.s.		n.s.	*
PEMSC-NLMS	n.s.	n.s.	n.s.	n.s.		*
ANCHOR	*	*	*	*	*	

(d)

	REFERENCE	H-NLMS1	H-NLMS2	H-NLMS3	PEMSC-NLMS	ANCHOR
REFERENCE		*	*	*	*	*
H-NLMS1	*		n.s.	n.s.	n.s.	*
H-NLMS2	*	n.s.		n.s.	n.s.	*
H-NLMS3	*	n.s.	n.s.		n.s.	*
PEMSC-NLMS	*	n.s.	n.s.	n.s.		*
ANCHOR	*	*	*	*	*	

re-convergence after a feedback path change, and the steady-state performance, allowing for a more detailed analysis of the different algorithms.

The quality ratings indicate that the proposed H-NLMS algorithm leads to a significantly improved performance in the initial convergence and re-convergence than the state-of-the-art PEMSC-NLMS algorithm. This is in line with the instrumental evaluation in Sec. III B where an increase in convergence rate was observed. Furthermore, the H-NLMS algorithm and PEMSC-NLMS algorithm perform similarly during the steady-state segment of the signal, which is expected as well, since the H-NLMS update rule is the same as for the PEMSC-NLMS algorithm during stable steady-state segments of the signal.

As indicated by the significant interaction between the main factors algorithm and signal, there is a dependency of the step-size μ_2 on the performance for different scenarios. This is expected since, in general, the optimal step-size of an adaptive algorithm depends on the signal characteristics. The results indicate that a step-size of $\mu_2 = 0.2$ or $\mu_2 = 0.5$ provides a better performance than $\mu_2 = 0.1$, with $\mu_2 = 0.5$ resulting in a higher perceived quality in most evaluated conditions.

IV. CONCLUSION

In this paper a hybrid AFC algorithm for HADs has been proposed, using either the NLMS update rule or the PEMSC-NLMS update rule based on a stability detector. The NLMS update rule is used to obtain fast convergence, while the PEMSC-NLMS update rule is used to provide low bias and MIS. The parameters of the soft-clipping-based stability detector have been designed to provide both a high sensitivity to detecting instability and to provide low signal distortion after the AFC has converged. The proposed H-NLMS algorithm has been evaluated for both speech and music signals for a changing acoustic feedback path. The results in terms of ASG and MIS show that the proposed algorithm converges and re-converges quickly while providing a very stable solution over time. In addition, based on the PESQ and PEAQ scores and the subjective listening test it is clear that there is almost no deterioration of the signal integrity and a significant improvement in perceived quality can be achieved compared to the state-of-the-art PEMSC-NLMS algorithm in the initial convergence and after changes of the acoustic feedback path. In conclusion, these results show that by utilizing the combined properties of two adaptive filters and using a simple stability detector, significant performance improvements can be achieved with only a minor increase in computational complexity.

ACKNOWLEDGMENTS

This research was supported in part by DAAD-ATN through the grant "Individualized acoustic feedback cancellation for hearing aids and assistive listening devices," by the Research Unit FOR 1732 "Individualized Hearing Acoustics" and the Cluster of Excellence 1077 "Hearing4All," funded by the German Research Foundation (DFG) and Project No. 57142981 "Individualized acoustic feedback cancellation" funded by the German Academic Exchange Service (DAAD).

[1]A. Spriet, S. Doclo, M. Moonen, and J. Wouters, "Feedback control in hearing aids," in *Springer Handbook of Speech Processing* (Springer-Verlag, Berlin Heidelberg, 2008), pp. 979–1000.
[2]T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," Proc. IEEE **99**, 1–40 (2011).
[3]S. Doclo, W. Kellermann, S. Makino, and S. E. Nordholm, "Multichannel signal enhancement algorithms for assisted listening devices: Exploiting spatial diversity using multiple microphones," IEEE Signal Process. Mag. **32**(2), 18–30 (2015).
[4]H. Nyquist, "Regeneration theory," Bell Syst. Tech. J. **111**, 126–147 (1932).
[5]A. Winkler, M. Latzel, and I. Holube, "Open versus closed hearing-aid fittings: A literature review of both fitting approaches," Trends Hear. **20**, 1–13 (2016).
[6]T. van Waterschoot and M. Moonen, "Comparative evaluation of howling detection criteria in notch-filter-based howling suppression," J. Audio Eng. Soc. **58**(11), 923–940 (2010); available at http://www.aes.org/e-lib/browse.cfm?elib=15738.
[7]J. A. Maxwell and P. M. Zurek, "Reducing acoustic feedback in hearing aids," IEEE Trans. Speech Audio Process. **3**(4), 304–313 (1995).
[8]J. M. Kates, "Constrained adaptation for feedback cancellation in hearing aids," J. Acoust. Soc. Am. **106**(2), 1010–1019 (1999).
[9]J. E. Greenberg, P. M. Zurek, and M. Brantley, "Evaluation of feedback-reduction algorithms for hearing aids," J. Acoust. Soc. Am. **108**(5), 2366–2376 (2000).
[10]U. Forssell and L. Ljung, "Closed-loop identification revisited," Automatica **35**(7), 1215–1241 (1999).
[11]J. Hellgren and U. Forsell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," IEEE Trans. Speech Audio Process. **9**(8), 906–913 (2001).
[12]M. Guo and B. Kuenzle, "On the periodically time-varying bias in adaptive feedback cancellation systems with frequency shifting," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (March, 2016), pp. 539–543.
[13]D. J. Freed, "Adaptive feedback cancellation in hearing aids with clipping in the feedback path," J. Acoust. Soc. Am. **123**(3), 1618–1626 (2008).
[14]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. **19**(5), 283–286 (2012).
[15]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. **20**(9), 2549–2563 (2012).
[16]C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Dual microphone solution for acoustic feedback cancellation for assistive listening," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto (March, 2012), pp. 149–152.
[17]C. R. C. Nakagawa, S. Nordholm, and W. Y. Yan, "Analysis of two microphone method for feedback cancellation," IEEE Signal Process. Lett. **22**(1), 35–39 (2015).
[18]L. T. T. Tran, H. H. Dam, H. Schepker, S. Doclo, and S. E. Nordholm, "Evaluation of two-microphone acoustic feedback cancellation using uniform and non-uniform sub-bands in hearing aids," in *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Hong Kong (December, 2015), pp. 308–313.
[19]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. **53**(10), 3749–3763 (2005).
[20]A. Spriet, G. Rombouts, and M. Moonen, "Adaptive feedback cancellation in hearing aids," J. Franklin Inst. **343**, 545–573 (2006).
[21]H. Schepker, L. T. T. Tran, S. Nordholm, and S. Doclo, "Improving adaptive feedback cancellation in hearing aids using an affine combination of filters," in *Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, New Orleans (March, 2016), pp. 231–235.
[22]J. Arenas-Garcia, A. Figueiras-Vidal, and A. Sayed, "Mean-square performance of a convex combination of two adaptive filters," IEEE Trans. Signal Process. **54**(3), 1078–1090 (2006).

[23]N. Bershad, J. Bermudez, and J.-Y. Tourneret, "An affine combination of two LMS adaptive filters transient mean-square analysis," IEEE Trans. Signal Process. **56**(5), 1853–1864 (2008).
[24]L. Ljung, *System Identification: Theory for the User*, 2nd ed., (Prentice-Hall, Upper Saddle River, NJ, 1999), Chap. 7, pp. 169–197.
[25]C. R. C. Nakagawa, S. Nordholm, and Y. Wei-Yong, "New insights into optimal acoustic feedback cancellation," IEEE Signal Process. Lett. **20**(9), 869–872 (2013).
[26]F. Strasser and H. Puder, "Correlation detection for adaptive feedback cancellation in hearing aids," IEEE Signal Process. Lett. **23**(7), 979–983 (2016).
[27]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. **61**(3), 125–137 (2013); available at http://www.aes.org/e-lib/browse.cfm?elib=16673.
[28]M. Hiipakka, M. Tikander, and M. Karjalainen, "Modeling the external ear acoustics for insert headphone usage," J. Audio Eng. Soc. **58**(4), 269–281 (2010); available at http://www.aes.org/e-lib/browse.cfm?elib=15253.
[29]T. Sankowsky-Rothe, M. Blau, H. Schepker, and S. Doclo, "Reciprocal measurement of acoustic feedback paths in hearing aids," J. Acoust. Soc. Am. **138**(4), 399–404 (2015).
[30]P. C. Loizou, *Speech Enhancement: Theory and Practice* (CRC Press, Taylor Francis Group, Boca Raton, London, New York, 2007), Chap. 11, pp. 479–571.
[31]ITU-R BS-1387, "Method for objective measurements of perceived audio quality," *International Telecommunications Union*, Geneva, Switzerland, 100 pp. (1998).
[32]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. **39**, 147–161 (2003).
[33]ITU-R BS. 1534-1, "Method for the subjective assessment of intermediate quality level of coding systems," International Telecommunications Union, Geneva, Switzerland, 18 pp. (2003).
[34]J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only Anova procedures," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, Vancouver (May, 2016), pp. 143–146.