Fifty Years of Acoustic Feedback Control: State of the Art and Future Challenges

The authors evaluate current methods available for dealing with the problem of acoustic feedback and identify the challenges facing future research in developing reliable and affordable solutions to the problem of controlling acoustic feedback.

By Toon van Waterschoot and Marc Moonen

ABSTRACT | The acoustic feedback problem has intrigued researchers over the past five decades, and a multitude of solutions has been proposed. In this survey paper, we aim to provide an overview of the state of the art in acoustic feedback control, to report results of a comparative evaluation with a selection of existing methods, and to cast a glance at the challenges for future research.

I. INTRODUCTION

Ever since sound reinforcement and public address (PA) systems have been in use, their performance has been troubled by the problem of acoustic feedback. Whenever a microphone captures a desired sound signal which is then processed (e.g., amplified) and played back by a loudspeaker in the same environment, as is the case in a PA system, the loudspeaker signal is unavoidably fed back into the microphone. In this way, a closed signal loop is created which affects the system performance, deteriorating the sound quality and limiting the achievable amplification. Among the different artifacts that are produced by this acoustic coupling between loudspeaker and microphone, the howling effect is without any doubt the most characteristic one.

Historically, some ambiguity has arisen in the terminology associated with the acoustic feedback problem. The term *acoustic feedback* has been used to refer to the undesired acoustic coupling between a loudspeaker and a microphone as well as to the howling effect that results from the coupling. We will use the term acoustic feedback in the first sense. Both the acoustic coupling and the howling effect are sometimes also referred to as the *Larsen effect*, after the Danish physicist Søren Larsen, who is said to have been one of the first researchers to investigate the acoustic feedback problem [1].

Acoustic feedback control refers to the process of attempting to solve the acoustic feedback problem either completely (i.e., to remove the acoustic coupling) or partially (e.g., to remove the howling artifacts from the loudspeaker signal). This paper only deals with *automatic* methods for acoustic feedback control, i.e., methods that do not require the interaction of an operator. Moreover, the emphasis will be on discrete-time methods that can be implemented on a digital signal processor. Surprisingly enough, despite 50 years of research on automatic acoustic feedback control, many PA system technicians still prefer to prosecute *manual* control of acoustic feedback. The main reason for this is the lack of reliability in the available automatic acoustic feedback control solutions, i.e., howling may still occur and may even take more time to be eliminated than in case of manual control.

T. van Waterschoot is on leave from the Department of Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, B-3001 Leuven, Belgium. He is now with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: toon.vanwaterschoot@esat.kuleuven.be).

M. Moonen is with the Department of Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, B-3001 Leuven, Belgium (e-mail: marc.moonen@esat.kuleuven.be).

The goal of this paper is threefold. First, we aim to provide an overview of the state of the art in acoustic feedback control, by reviewing relevant scientific papers and patents that have been published over the past five decades. The focus of the literature review is on acoustic feedback control for room acoustic sound reinforcement systems, a PA system being the most exemplary application. We should point out, however, that during the past two decades a considerable amount of research results has been published related to feedback control in hearing aids (HAs), where the feedback is due to a combined acoustic and mechanical coupling. Some of these publications will also be referenced here as they have provided solutions that have successfully been extrapolated to room acoustic applications. Our second goal is to report results of a comparative evaluation of the three most widely used methods for acoustic feedback control, namely phase-modulating feedback control (PFC), notch-filter-based howling suppression (NHS), and adaptive feedback cancellation (AFC). To our knowledge, such a comparative evaluation has not been reported earlier. This is presumably due to the fact that these three methods in fact attempt to solve different problems (i.e., smoothing the system loop gain versus howling suppression versus removal of the acoustic loudspeaker–microphone coupling) and hence different measures have been used previously to quantify the performance of each of these methods individually. We observe, however, that even though the PFC, NHS, and AFC problem formulations are different, the ultimate objectives of these methods are the same: to improve sound quality, to increase the amount of achievable amplification, and to operate in a reliable way. Hence the reported comparative evaluation is carried out with these three objectives in mind. Finally, out third goal is to formulate the challenges that we believe are most prevalent to steer future research in acoustic feedback control towards the development of reliable and affordable solutions.

The paper is organized as follows. In Section II, the acoustic feedback problem is formulated in a formal way, where the key result is the Nyquist stability criterion, based on which all the acoustic feedback control methods can be derived. Section III deals with the state of the art in acoustic feedback control: the existing feedback control solutions are divided into four categories [phase-modulation (PM) methods, gain reduction methods, spatial filtering methods, and room modeling methods], and a detailed literature review of each category is given. In the next sections, the three most popular acoustic feedback control methods are outlined in more detail: Section IV deals with the PFC method, Section V with the NHS method, and Section VI with the AFC method. In Section VII, these three methods are then evaluated in terms of the resulting sound quality, the achievable amplification, as well as their reliability. The evaluation is based on computer simulation results using realistic room acoustic models and for both speech and audio signals. Finally, in the concluding Section VIII, we summarize the results obtained with the state-of-the-art methods and formulate future research challenges in acoustic feedback control.

II. THE ACOUSTIC FEEDBACK PROBLEM

A typical PA system scenario is shown in Fig. 1: a number of microphones are positioned such as to pick up the sound of
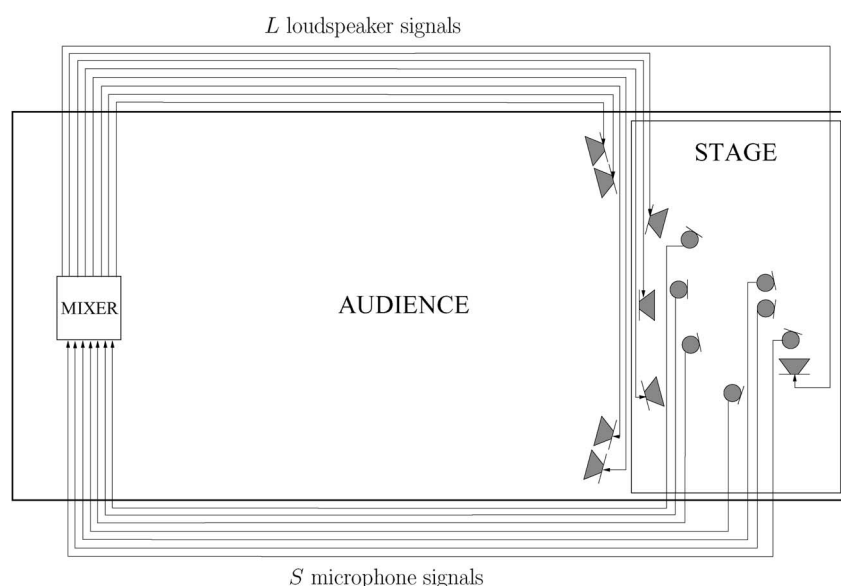

Fig. 1. A typical public address (PA) system scenario, featuring seven microphones, four onstage loudspeakers, four loudspeakers directed towards the audience, and a mixing/signal processing/amplification console.
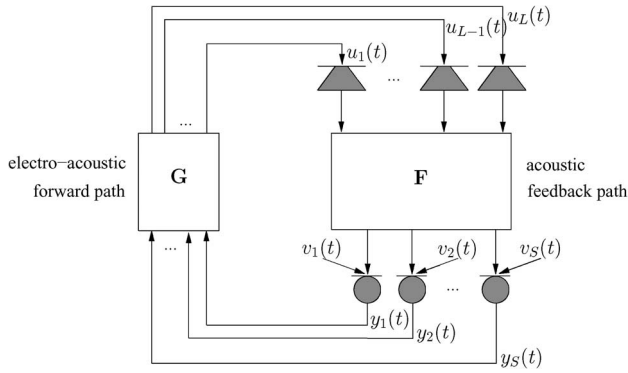
Fig. 2. *Discrete-time model of a PA system with S microphones and L loudspeakers.*

possibly multiple sound sources that are of interest. The microphone signals are then routed to the mixing console and may be subject to additional processing, such as dynamic range processing, artificial reverberation, etc., which is usually performed in the digital domain. The mixed signals are then amplified and sent to the loudspeakers, which are often arranged in a group-wise fashion (i.e., all the loudspeakers in the same loudspeaker group broadcast the same signal). Usually, the microphones and loudspeakers are positioned in such a way that, taking into account their directivity, the loudspeaker sound does not directly hit the microphones, i.e., no direct acoustic coupling between the loudspeakers and the microphones exists. However, in nearly every sound reinforcement application it is unavoidable that the loudspeaker sound is reflected by the boundaries (walls, floor, and ceiling) of the acoustic environment (denoted as the "room") and by subjects and objects within the environment. These reflections constitute an indirect acoustic coupling between the loudspeakers and the microphones of the system.

The PA scenario can be modeled in a discrete-time context as shown in Fig. 2. All continuous-time signals involved are assumed to be bandlimited in such a way that they can be sampled at a standard sampling frequency (e.g., $f_s = 16$ kHz in speech applications, $f_s = 44.1$ kHz in audio applications) and represented by their discrete-time counterparts.[1] If we represent the S source signals by $v_i(t)$, $i = 1, \ldots, S$, the corresponding S microphone signals as $y_i(t)$, $i = 1, \ldots, S$, and the L loudspeaker signals as $u_j(t)$, $j = 1, \ldots, L$, then the discrete-time

closed-loop system model in Fig. 2 can be described by the following relations:

$$\bar{\mathbf{y}}(t) = \mathbf{F}(q, t)\bar{\mathbf{u}}(t) + \bar{\mathbf{v}}(t) \tag{1}$$

$$\bar{\mathbf{u}}(t) = \mathbf{G}[\bar{\mathbf{y}}(t), t]. \tag{2}$$

Here, the source signal, microphone signal, and loudspeaker signal vectors are defined as

$$\bar{\mathbf{v}}(t) = [v_1(t) \quad \ldots \quad v_S(t)]^T \tag{3}$$

$$\bar{\mathbf{y}}(t) = [y_1(t) \quad \ldots \quad y_S(t)]^T \tag{4}$$

$$\bar{\mathbf{u}}(t) = [u_1(t) \quad \ldots \quad u_L(t)]^T \tag{5}$$

and the multichannel acoustic feedback path $\mathbf{F}(q, t)$ and electroacoustic forward path characteristics $\mathbf{G}[\cdot, t]$ are defined below.

Between each loudspeaker–microphone pair (j, i), $j = 1, \ldots, L, i = 1, \ldots, S$, there exists an acoustic coupling, which can be modeled by the acoustic feedback path transfer function

$$F_{ij}(q, t) = f_{ij}^{(0)}(t) + f_{ij}^{(1)}(t)q^{-1} + \ldots + f_{ij}^{(n_F)}(t)q^{-n_F} \tag{6}$$

where q denotes the discrete-time shift operator, i.e., $q^{-k}u_j(t) = u_j(t - k)$. The multichannel feedback path matrix in (1) is then defined as an $S \times L$ polynomial matrix

$$\mathbf{F}(q, t) = \begin{bmatrix} F_{11}(q, t) & \ldots & F_{1L}(q, t) \\ \vdots & \ddots & \vdots \\ F_{S1}(q, t) & \ldots & F_{SL}(q, t) \end{bmatrix}. \tag{7}$$

The acoustic feedback path model is linear, time varying, and of finite order[2] n_F. The linearity assumption is generally considered to be a reasonable one, since the effects of sound propagation and reflections in the acoustic environment (i.e., signal attenuations and time delays) are quasi level independent. The finite-order assumption, which contrasts with the infinite impulse response (IIR) nature of room acoustics, can be justified by the observation that a typical room impulse response (RIR) has an exponentially decaying envelope such that it can be truncated to have $n_F + 1 < \infty$ coefficients.

---

[1] In our notation, we discriminate between continuous-time and discrete-time signals by using curly brackets for the former and round brackets for the latter. For example, $x\{\tau\}$ is a continuous-time signal and $x(t)$ is the corresponding discrete-time signal. The discrete-time index t is related to the continuous-time index τ as $t = \tau/T_s$ with the sampling interval defined as $T_s = 1/f_s$.

[2] For ease of notation, we assume that all the acoustic feedback path transfer functions $F_{ij}(q, t)$, $i = 1, \ldots, S, j = 1, \ldots, L$, have the same order n_F. This is also reasonable from a physical point of view, since the reverberation time in a room does not depend on the loudspeaker and microphone positions. We will make the same assumption further on for the electroacoustic forward path transfer functions $G_{ji}(q, t)$.
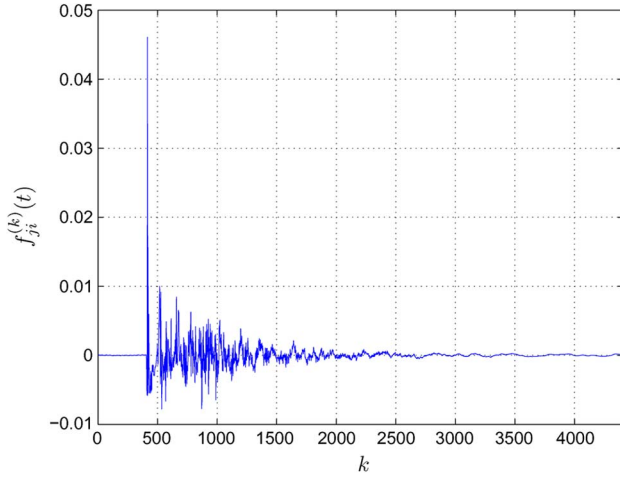
Fig. 3. *A typical RIR, measured at $f_s = 44.1$ kHz and truncated at a length of $n_F + 1 = 4410$ coefficients.*

An example RIR, which was measured at $f_s = 44.1$ kHz and truncated at a length of $n_F + 1 = 4410$ coefficients (corresponding to 100 ms), is shown in Fig. 3. The frequency response of this RIR is displayed in Fig. 4. It can be seen that the magnitude response has an overall low-pass behavior as well as many local magnitude peaks and dips. This irregular behavior was explained and quantified by Schroeder [2], under the assumption that the acoustic coupling is mainly due to reflections and not due to a direct acoustic path between the loudspeaker and the microphone. The average frequency distance between two magnitude peaks is then about 10 Hz, and the peak magnitude can be up to 10 dB larger than the average magnitude in the frequency response [2].

In the electroacoustic forward path, the S microphone signals are mixed and amplified to obtain L loudspeaker signals, and moreover, some additional signal processing is performed. Since usually nonlinear dynamics processing (e.g., compression, limiting, etc.) is involved here, the forward path mapping $G_{ji}[\cdot, t]$ between the (i, j)th microphone–loudspeaker pair should be modeled as a nonlinear, time-varying filter. However, to be able to perform a stability analysis of the closed-loop system, we will mostly assume that the forward path can be modeled by a linear, time-varying transfer function, $\forall i, j$

$$G_{ji}[\cdot, t] = G_{ji}(q, t)$$
$$= g_{ji}^{(0)}(t) + g_{ji}^{(1)}(t)q^{-1} + \ldots + g_{ji}^{(n_G)}(t)q^{-n_G} \quad (8)$$

and

$$\mathbf{G}[\cdot, t] = \mathbf{G}(q, t) = \begin{bmatrix} G_{11}(q, t) & \ldots & G_{1S}(q, t) \\ \vdots & \ddots & \vdots \\ G_{L1}(q, t) & \ldots & G_{LS}(q, t) \end{bmatrix}. \quad (9)$$

If the forward path includes IIR components, such as IIR equalization filters, we have that $n_G = \infty$. We further assume that the sound sources have sufficient directivity and are close enough to the respective microphones, such that the acoustic transfer function matrix from the sources to the microphones is an identity matrix. These assumptions can be justified since these do not relate directly to the feedback problem.
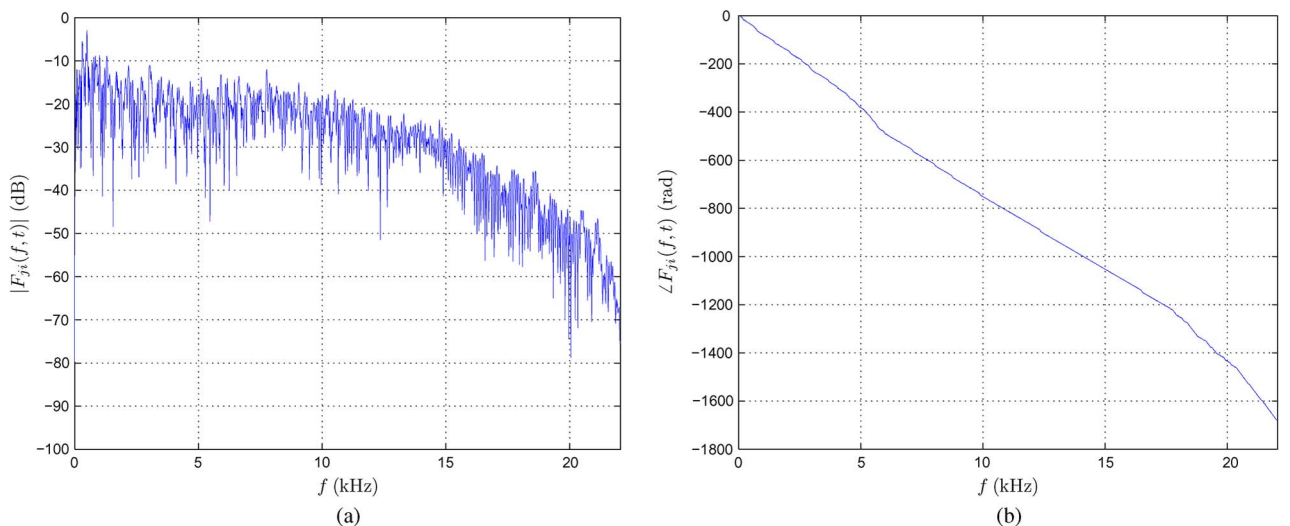

Fig. 4. *Frequency response of the RIR shown in Fig. 3. (a) Magnitude response. (b) Phase response.*

While many sound reinforcement systems comprise multiple loudspeakers and microphones, most acoustic feedback control methods have been proposed in a single-channel context (i.e., for one loudspeaker and one microphone), without a framework for an extension to multichannel systems being explicitly provided. For this reason, we will analyze the acoustic feedback problem and explain the acoustic feedback control methods in a single-channel context, and drop the subscripts i and j. We will however comment on the implications of extending a particular method to a multichannel system whenever appropriate.

In a single-channel sound reinforcement system, the closed-loop frequency response from the source signal to the loudspeaker signal can be expressed as follows:

$$\frac{U(\omega, t)}{V(\omega, t)} = \frac{G(\omega, t)}{1 - G(\omega, t)F(\omega, t)}. \quad (10)$$

Here, $\omega \in [0, 2\pi]$ represents the radial frequency variable, $U(\omega, t)$ and $V(\omega, t)$ denote the short-term frequency spectra of the loudspeaker and source signal, and $G(\omega, t)$ and $F(\omega, t)$ are the short-term frequency responses of the forward and feedback path, which can be calculated using the short-time discrete Fourier transform (DFT). The frequency function $G(\omega, t)F(\omega, t)$ appearing in the denominator of (10) is often referred to as the "loop response" of the system, and plays a crucial role in acoustic feedback control [the corresponding magnitude response $|G(\omega, t)F(\omega, t)|$ is then referred to as the "loop gain" and the phase response $\angle G(\omega, t)F(\omega, t)$ as the "loop phase"]. It is well known that a closed-loop system can exhibit instability, which may lead to oscillations that, in an acoustic system, are perceived as howling. Stability analysis of linear closed-loop systems is by now a well-understood topic in control systems theory, which originated from early studies on feedback amplifiers. The current approach to closed-loop system stability analysis is based on a classical paper by Nyquist [3]. The *Nyquist stability criterion* can be formulated as follows[3]: if there exists a radial frequency $\omega = 2\pi(f/f_s)$ for which

$$\begin{cases} |G(\omega, t)F(\omega, t)| \geq 1 & (11) \\ \angle G(\omega, t)F(\omega, t) = n2\pi, \quad n \in \mathbb{Z} & (12) \end{cases}$$

then the closed-loop system is unstable. If the unstable system is moreover excited at the critical frequency f, i.e., if the source signal contains a nonzero frequency

---

[3] We should note that the Nyquist stability criterion is defined in [3] for linear time-invariant systems. The stability of linear time-varying systems should be analyzed by the so-called circle criterion instead [4, Ch. 5]. However, to achieve consistency with the literature on acoustic feedback control, we will still define stability using the Nyquist criterion, under the assumption that the electroacoustic forward path and the feedback path characteristics are slowly time varying.

component at f, then an oscillation at this frequency will occur. The criterion in (11) and (12) is essential in the remainder of this paper, since any acoustic feedback control method effectively attempts at preventing either one or both of these conditions from being met.

With the aim of quantifying the achievable amplification in a sound reinforcement system with and without acoustic feedback control, it is customary to define a broadband gain factor $K(t)$ as the average magnitude of the forward path frequency response $G(\omega, t)$ and extract it from the forward path transfer function $G(q, t)$, i.e.,

$$G(q, t) = K(t)J(q, t) \quad (13)$$

with

$$K(t) = \frac{1}{2\pi} \int_0^{2\pi} |G(\omega, t)| \, d\omega. \quad (14)$$

Assuming now that $J(q, t)$ is given, and that $K(t)$ can be varied, the maximum stable gain (MSG) can be defined as follows:

$$\text{MSG}(t)[\text{dB}] \stackrel{\Delta}{=} 20 \log_{10} K(t) \quad \text{such that}$$
$$\max_{\omega \in \mathcal{P}} |G(\omega, t)F(\omega, t)| = 1 \quad (15)$$
$$= -20 \log_{10} \left[\max_{\omega \in \mathcal{P}} |J(\omega, t)F(\omega, t)| \right] \quad (16)$$

where $\mathcal{P}$ denotes the set of frequencies at which the phase condition (12) is fulfilled, i.e.,

$$\mathcal{P} = \{\omega | \angle G(\omega, t)F(\omega, t) = n2\pi\}. \quad (17)$$

From a statistical analysis of room acoustics, assuming a flat forward path magnitude response and a unity average feedback path magnitude response, Schroeder concluded that in a sound reinforcement system without feedback control and having a reverberation time of T_{60} s and a bandwidth of B Hz, the expected MSG can be calculated as [2]

$$\text{MSG}(t)[\text{dB}] = -10 \log_{10}[\log_{10}(BT_{60}/22)] - 3.8. \quad (18)$$

The gain margin is defined as the difference between the MSG and the actual gain of the system. From a sound quality point of view, a gain margin of 2–3 dB is recommended to avoid audible ringing effects [2], [5].

III. STATE OF THE ART IN ACOUSTIC FEEDBACK CONTROL

As already mentioned, we will only deal with automatic methods for acoustic feedback control. A review of manual feedback control methods is given in [6]. These methods are based on a proper microphone and loudspeaker selection and positioning, suppression of discrete room modes using notch filters, and equalization of the entire room response using 1/3 octave graphic equalizer filters, and may result in an MSG increase of 5–8 dB [6].

Automatic feedback control methods may be categorized into four classes: PM methods, gain reduction methods, spatial filtering methods, and room modeling methods.

A. Phase-Modulation Methods

One of the earliest approaches to acoustic feedback control consists in frequency shifting (FS) the microphone signals before these are amplified and sent to the loudspeakers. The FS approach can largely be attributed to Schroeder, who published a number of papers on this topic in the early 1960s [2], [7]–[9]. By applying FS, the loop gain can be smoothed, such that ideally, the MSG is determined by the average magnitude response rather than the peak magnitude response [9]. Since the average frequency distance between two magnitude peaks in a room response was found to be around 10 Hz, the optimal FS value is expected to be around 5 Hz [2]. An MSG increase up to 14 dB was reported [7], however, the subjectively acceptable MSG increase is limited to 6 dB if audible beating effects due to the FS operation are to be avoided [2], [9]. It is claimed in [9] that a frequency shift of 5 Hz is inaudible both for speech and music signals. The earliest FS implementations were based on analog single-sideband modulation [10] or phase modulation [11]. More recently, a digital FS implementation using a truncated FIR Hilbert filter has been proposed [12]. A drawback of the FS approach is that it does not preserve the harmonic relations between tonal components in voiced speech and music signals. It was shown in [13] that a bandwidth compression does preserve harmonic relations and results in a feedback stability improvement similar to the FS approach.

Another early feedback control method employs phase modulation (PM) in the electroacoustic forward path, with the aim of bypassing the phase condition (12) in the Nyquist criterion. In 1958, Mishin [14], [170] described a sinusoidal PM approach in which the choice of the modulation parameter relates to the zeros of Bessel functions of the first kind. In a 1968 paper by Nishinomiya [15], an MSG increase up to 7 dB is reported using sinusoidal frequency modulation (FM), which is conceptually equivalent to sinusoidal PM. Guelke and Broadhurst [5] applied the sinusoidal PM technique in the context of reverberation enhancement (RE) systems, using a very low modulation frequency (∼1 Hz), and resulting in a 4-dB MSG increase. The apparent suitability of PM, FM, and other periodic modulations for feedback control in digital RE systems resulted in a renewed interest in these methods in the 1990s. Svensson [16] and Nielsen and Svensson [17] provided a unifying approach to PFC in which the modulators, including sinusoidal PM, FM, amplitude modulation (AM), and delay modulation (DM), are viewed as linear periodically time-varying filters. Moreover, they showed that the FS approach also fits into this framework, hence labeling FS-based feedback control as a special case of PFC. Svensson [16] reported an average 4-dB MSG increase with a synthetic acoustic feedback path, while Nielsen and Svensson [17] obtained MSG increases up to 8 dB in real room acoustic feedback scenarios. Poletti [18] was the first to study the performance of PFC (in particular using an FS approach) in multichannel sound systems. His somewhat discouraging conclusion was that the stability improvement due to FS reduces as the number of channels increases. Finally, while the impact of the PFC approach on sound quality may be considerable, to our knowledge only a single study has been devoted to its perceptual evaluation. The results of a perceptual study by Svensson [19] indicate that the PFC approach (and in particular the FS approach) may be well suited for transient signals like speech but is less appropriate for sustained tones often occurring in audio signals.

In Section IV, a more extensive treatment of the PFC approach is provided.

B. Gain Reduction Methods

The most straightforward approach to acoustic feedback control is to automate the actions that a human operator would undertake for preventing or eliminating howling in a sound reinforcement system. These actions usually consist in reducing the electroacoustic forward path gain, such that the system moves away from magnitude condition (11) in the Nyquist criterion. Depending on the width of the frequency band in which the gain is actually reduced, we can discriminate between three gain reduction methods:

1) in automatic gain control (AGC) methods [20]–[22], the gain is reduced equally in the entire frequency range by decreasing the broadband gain factor $K(t)$ defined in (14);
2) in automatic equalization (AEQ) [22]–[30], the gain reduction is applied in critical subbands of the entire frequency range, namely those subbands in which the loop gain is close to unity;
3) in NHS [31]–[58], the gain is reduced in narrow frequency bands around critical frequencies, i.e., frequencies at which the loop gain is close to unity.

Every gain reduction method has to be activated in some way, when a closed-loop instability or a tendency towards instability is detected. Only a few gain reduction

methods have been proposed which are based on a proactive instability detection: these are either based on an online measurement of the feedback path magnitude response [52], [59] or on an early detection of the spectral accumulation effect that can be observed at critical frequency components in the microphone signal [28]–[30], [60], [61]. Most gain reduction methods are reactive, in the sense that howling can usually be perceived before it is actually detected. In these methods, howling detection is typically based on a combined spectral and temporal analysis of the microphone signal. Due to the sinusoidal nature of howling, the microphone signal frequency components having the largest magnitude are considered to be candidate howling components. The true howling components within this set of candidates can then be discriminated from the source signal tonal components (originating from voiced speech or musical tones) using several criteria. Spectral criteria for discriminating between howling and tonal components are based on one or more of the following features: the power ratio of the candidate howling component and the entire spectrum [25]–[27], [35]–[38], [45], [49]–[51], [55], [56], the power ratio of the candidate howling component and its (sub)harmonics [32], [34]–[36], [43], and the power ratio of the candidate howling component and its neighboring frequency components [22], [28]–[30]. On the other hand, temporal criteria for howling detection rely on the observation that howling components typically persist for a longer time than tonal components [20], [21], [31], [32], [34], [38], [43]–[45], [49] and exhibit an exponentially increasing magnitude until the sound reinforcement system saturates [28]–[30]. A comparative evaluation of these spectral and temporal howling detection criteria is reported in [62] and [63].

The AGC method is the earliest gain reduction method, which was proposed by Patronis in 1978 [20], [21]. If howling is detected, the broadband gain is immediately reduced, and after a specified time interval the gain is restored to the initial value. Candidate howling frequencies are discriminated from tonal source signal components by assuming that howling components persist for several seconds. A subband implementation of this method was proposed by Ando [22], featuring a spectral approach to howling detection by evaluating power ratios between adjacent subbands. Obviously, AGC methods do not increase the MSG since the spectral shape of the loop gain is not altered. The main strength of AGC methods is their reliability: if the gain is sufficiently reduced, an unstable system is guaranteed to be stabilized. Therefore, many other acoustic feedback control methods include an AGC method as a "rescue procedure" that is activated if all else fails; see, e.g., [22], [35], [36], [55], [56], and [64].

The AEQ method follows directly from the subband approach to AGC, as proposed by Ando [22]. If howling detection is performed in frequency subbands, then the gain reduction can be limited to those subbands in which howling is detected. Hanajima et al. [25], [26] further improved the subband howling detection, by first performing a howling detection in relatively wide subbands, and subsequently dividing the most critical subband in narrower subbands in which the howling detection is then repeated. They use ten logarithmically spaced wide subbands in the 10–10 000-Hz range, which are then divided into ten linearly spaced narrower subbands to obtain a more accurate howling detection. An even more advanced howling detection can be found in the AEQ method of Osmanovic et al. [28]–[30]. The detection criterion consists of a linear combination of two features that are calculated for all candidate howling components: the "slopeness" is a temporal feature that models the exponential buildup of a howling component, while the "peakness" is a spectral feature that estimates the power ratio of a candidate howling component and its neighboring frequency components. For the equalization, Osmanovic et al. use 14 logarithmically spaced eighth-order IIR bandstop filters in the speech range 300–6000 Hz [28]–[30].

The NHS methods can be divided into two categories, i.e., one-stage and two-stage NHS methods, depending on whether the howling detection and notch filtering are performed jointly or separately. The earliest NHS methods are one-stage methods, which are usually implemented using adaptive notch filters (ANFs). In 1989, Foley proposed the adaptive periodic noise canceller [31] for speech applications, which is an FIR-ANF that is able to track and cancel a narrowband component in the microphone signal. Since the FIR-ANF in [31] is adapted using the least mean squares (LMS) algorithm, it is expected to be too slow to cancel tonal speech components, which vary more quickly in time than howling components. Also, the FIR-ANF is preceded by a delay of eight samples such that it cannot cancel the short-term-correlated speech formants. Foley's FIR-ANF was shown to be H_∞-robust for a first-order feedback path (i.e., $n_F = 1$), provided that the LMS stepsize is properly chosen [41], [47]. Staudacher [40] proposed an extension to Foley's FIR-ANF, by using a variable LMS stepsize that increases as the FIR-ANF input signal power increases, such that the convergence is accelerated when howling occurs. To reduce the impact of the ANF on sound quality, the notch filter bandwidth should be as small as possible. A disadvantage of the FIR-ANF implementation is that a large filter order is required to obtain a narrowband notch characteristic, e.g., Foley [31] and Staudacher [40] use 32nd-order filters to cancel a single narrowband component. If multiple howling components are to be canceled, the required FIR-ANF filter order may become unpractically large. Following this observation, several IIR-ANF implementations have been proposed, which only require a biquadratic (i.e., second-order) filter structure to cancel one narrowband component. The main difficulty with IIR-ANF implementations is that the least squares (LS) cost function associated with the howling component frequency estimation is typically nonconvex.

Kuo and Chen [33] proposed a constrained biquadratic IIR-ANF in which the global minimum of the LS cost function can be found with high probability by increasing the notch bandwidth during the howling detection process. Once howling has been detected, the notch filter is activated in the electroacoustic forward path with a reduced bandwidth, to avoid a loss of sound quality. Another approach to bypass local minima in the LS cost function associated with the IIR-ANF implementation is to only adapt the FIR part of the filter, and subsequently copy the numerator coefficients to the denominator [43], perhaps after including some scaling factor [39], [54]. A biquadratic IIR-ANF implementation featuring an advanced howling detection method was proposed by Porayath and Mapes-Riordan [43]: a howling frequency is detected when it has a power that is 30 dB larger than its first harmonic and when this power difference persists for at least 50–100 ms. Since the power spectral density is however hard to estimate when using the ANF approach, a different howling detection method was recently proposed by Gil-Cacho et al. [58], which is based on running multiple regularized biquadratic IIR-ANFs in parallel with different regularization factors. Yet another second-order ANF implementation was proposed by Wei et al. [48], in which the input samples to the ANF consist of phase-shifted instead of time-shifted microphone signal samples.

The two-stage NHS method, which is by now probably the most popular gain reduction method for acoustic feedback control, originates from the work of Lewis et al. [32], [34] and Er et al. [35], [36] in the early 1990s. A nonparametric frequency analysis of the microphone signal is computed using a fast Fourier transform (FFT) algorithm, from which the candidate howling components are determined using a peak picking algorithm. The power of the candidate howling components is then compared to an absolute power threshold [35], [36], to the average signal power [35], [36], and to the (sub)harmonics power [32], [34]–[36] to determine if howling occurs. This spectral criterion is combined with a temporal criterion for howling detection by Lewis et al. [32], [34]. Whenever howling is detected, biquadratic notch filters are inserted in the electroacoustic forward path. Several improvements to the methods by Lewis et al. and Er et al. have been reported. Kawamura et al. [37], [38] propose an online modification of the thresholds used in the spectral and temporal howling detection criteria, steered by estimates of the background noise spectrum, the source signal spectrum, the reverberation time, and the acoustic feedback path response. Lane et al. [42] apply a parametric frequency analysis instead of the nonparametric analysis proposed earlier, using a set of adjustable bandpass filters having relatively wide passbands as compared to the stopbands of the notch filters. An alternative way of determining the set of candidate howling components was proposed by Williams [44], [45]: instead of executing a peak picking algorithm on the FFT magnitude spectrum estimate, a so-called "ballistics procedure" is applied to

model the temporal buildup of narrowband components such that components with an increasing power can be identified. Rocha and Ferreira [49] and Börsch [50], [51] replace the FFT algorithm in the nonparametric frequency analysis by an odd FFT algorithm and a frequency-warped FFT algorithm, respectively. Moreover, the frequency analysis described by Börsch [50], [51] is the only nonparametric method which includes a compensation for the estimation errors due to the limited FFT resolution. In [52], Rombouts et al. propose a proactive howling detection method applied to NHS, based on the estimation of critical closed-loop system frequencies from an adaptive estimate of the feedback path response. Abe [53] was the first to consider NHS in a multichannel sound reinforcement system, and succeeded in reducing the computational and memory requirements by frequency-analyzing the individual microphone signals with a low-resolution FFT algorithm and the mixed signal with a high-resolution FFT algorithm. Finally, Somasundaram [55], [56] proposes an advanced spectral howling detection criterion, in which the power of the candidate howling component is compared to a threshold that is calculated using the mean and standard deviation of the entire FFT spectrum estimate. Furthermore, the notch filters used in [55] and [56] are gradually enabled and disabled using a so-called leaky integrator, to avoid artifacts in the loudspeaker signal.

Since the majority of the available gain reduction methods are described in patents, not many experimental results are available and no MSG increase values have been reported. However, from Schroeder's statistical analysis of a feedback path frequency response [2], it can be expected that if the loop gain could be perfectly smoothed using an AEQ or an NHS approach, a maximal MSG increase of about 10 dB may be achieved. The two-stage NHS method, being the most popular of all gain reduction methods, will be described in more detail in Section V.

C. Spatial Filtering Methods

Spatial filtering methods for acoustic feedback control aim at altering the loop response $G(\omega, t)F(\omega, t)$ of the closed-loop system by using microphone and/or loudspeaker arrays of which the received/transmitted signals are processed by beamforming filters. The general objective is then to design a microphone array beamformer that has its main lobe (i.e., its maximal spatial response) in the direction of the source while having a null (i.e., zero spatial response) in the direction of the loudspeaker, and/or a loudspeaker array with the main lobe directed towards the audience and a null in the direction of the microphone. The first spatial filtering approach to acoustic feedback control was proposed by Duong et al. in 1984 for hands-free telephony applications [65], focusing on the combined use of a microphone and loudspeaker array for a single-channel scenario with fixed microphone and loudspeaker positions. The stringent spatial constraints (i.e.,

the microphone and loudspeaker arrays are to have the same center and lie orthogonal to each other) make this method rather impractical for many sound reinforcement applications. A more flexible approach, which allows for scenarios with arbitrary microphone and loudspeaker array positions, consists in adapting the beamformer coefficients based on the available sound signals. Obviously, an adaptive microphone array is more straightforward to implement than an adaptive loudspeaker array, since the latter does not collect any information on the acoustic environment. A fundamental problem that occurs when computing the coefficients of an adaptive microphone array beamformer in a closed-loop system is the fact that the source signal is correlated with the loudspeaker signal [i.e., the loudspeaker signal can be calculated by filtering the source signal with the closed-loop response; see (10)]. Due to this correlation, a conventional adaptive beamforming algorithm will not converge to the desired solution, and consequently, part of the source signal will eventually be attenuated while part of the feedback signal will still appear in the output of the microphone array. Several solutions to this correlation problem have been proposed. Janse and Belt [66] propose the combined use of an adaptive feedback canceller (AFC) and a microphone array beamformer. By feeding the feedback-compensated signal from the AFC to the microphone array, the influence of the feedback signal on the beamforming algorithm can be decreased. In this case, however, it is not possible to create a beamformer null directed towards the loudspeaker, since the feedback-compensated signal (ideally) does not provide any information on the loudspeaker position. Another solution was proposed by Kobayashi et al. [67], [68], in which the coefficients of an adaptive microphone array beamformer outside the closed signal loop are computed by canceling the source signal using a null beamformer (NBF) and inserting an artificial source signal. The adaptive beamformer coefficients are then copied to a microphone array beamformer in the closed signal loop, resulting in an MSG increase up to 15 dB [67], [68]. Due to the source signal cancellation, the adaptive beamformer can unambiguously identify the loudspeaker direction, however, the direction of the source with respect to (w.r.t.) the microphone array needs to be known a priori [67] or estimated by an adaptive NBF [68]. The artificial source signal, of which the design is not specified in [67] and [68], serves to constrain the adaptive beamformer response to unity in the source direction. A more recent solution to the correlation problem in adaptive microphone array beamforming was proposed by Rombouts et al. [69], [70], and consists in prewhitening the source signal component in the adaptive beamformer desired signal using an adaptive decorrelation filter that is estimated concurrently with the beamformer coefficients. This approach was shown to result in an MSG increase between 7 and 14 dB (depending on the reverberation time of the room), while it does not require a priori information on the source position and is

considerably cheaper than the approach in which an AFC is also used. Finally, a fundamentally different approach to spatial filtering for acoustic feedback control was proposed by Goodwin and Elko [71], [72]. In the so-called "beam dithering" approach, a loudspeaker array is steered by a beamformer of which the coefficients are varied periodically with time, by time stepping through a discrete sequence of approximate Chebyshev coefficients. In this way, a spatial modulation is obtained that provides a smoothing of the loop gain, comparable to the smoothing effect obtained with the PM methods for acoustic feedback control. An MSG increase up to 6 dB has been obtained [72], however, the spatial constraints of the beam dithering approach are rather stringent (in that the audience should always be in the main beamformer lobe, while the microphones should be in the sidelobes) and a perceptual calibration of the system is required [71].

D. Room Modeling Methods

In room modeling methods for acoustic feedback control, a model of the acoustic feedback path is identified either offline (during the initialization of the sound reinforcement system) or online (during the operation of the sound reinforcement system). We can distinguish between two room modeling methods, depending on how the model is subsequently applied for acoustic feedback control. In AFC, the acoustic feedback path model is used to predict the feedback signal component in the microphone signal (i.e., the part of the microphone signal that stems from the loudspeaker signal through the acoustic coupling). The predicted feedback signal is then subtracted from the microphone signal, hence resulting in a feedback-compensated signal, which is in fact an estimate of the source signal component in the microphone signal. If an accurate model of the acoustic feedback path can be identified, then the AFC method achieves a nearly complete elimination of the acoustic coupling (i.e., the loop gain comes close to zero for all frequencies), and consequently very large MSG increases may be obtained. Alternatively, the inverse of the acoustic feedback path can be modeled and identified, and this inverse model can then be inserted in the closed signal loop to optimally equalize the microphone signal. This approach is referred to as adaptive inverse filtering (AIF), and ideally results in a perfect smoothing of the loop gain, for which the MSG increase can be expected to be around 10 dB [2].

The AIF approach has received only little attention in the context of acoustic feedback control. In 1994, Ushiyama et al. [73] proposed an inverse filtering approach in which an inverse model of the minimum-phase components in the acoustic feedback path is identified offline. It is observed that a smoothing of the inverse model frequency response increases the robustness of the (time-invariant) inverse model w.r.t. time variations in the acoustic feedback path response. Another offline approach to inverse filtering was

proposed by Nagata *et al.* [23], [24], and consists in automatically adjusting a large number of equalizers in the electroacoustic forward path, based on an offline measurement of the acoustic feedback path response using a noise probe signal. Finally, a hybrid AIF-AFC approach was proposed by Janse and Belt [66] and Schmidt and Haulick [74], in which the inverse model coefficients are adjusted based on the acoustic feedback path model that is identified in the AFC algorithm. More results on the AIF approach can be found in the literature on acoustic dereverberation and equalization; see, e.g., [75]–[78].

In the AFC approach, which is conceptually similar to the well-known acoustic echo cancellation (AEC) approach, an adaptive filter is used to model, identify, and track the impulse response of the acoustic feedback path. Analogously to the correlation problem found in adaptive microphone array beamforming (see Section III-C), the fundamental problem encountered in AFC lies in the fact that, unlike in the AEC case, the adaptive filter's input signal (i.e., the loudspeaker signal) and disturbance signal (i.e., the source signal) are now correlated; see (10). Applying a standard adaptive filtering algorithm to the AFC problem hence results in a biased estimate of the acoustic feedback path impulse response [79]–[81], and consequently, the source signal component in the microphone signal ends up being partially canceled. For this reason, a decorrelation method is generally incorporated in the AFC scheme which is either included in the closed signal loop or in the adaptive filtering circuit [81]; see [82] for an overview and comparative evaluation.

Decorrelation in the closed signal loop can be accomplished by injecting a noise signal, by including a nonlinear or time-varying signal operation, or by inserting a processing delay in the electroacoustic forward path. The earliest AFC reference appears to be a 1988 patent by Ibaraki *et al.* [83], in which a white noise signal is injected in the closed signal loop noncontinuously (e.g., during source signal pauses) to identify the low-frequency response of the acoustic feedback path. Goertz [84] proposes to inject a white noise signal continuously and reports a 5-dB MSG increase in a severely undermodeled AFC scenario (i.e., the adaptive filter length being only 1/15 of the feedback path length). Decorrelation by continuous white noise injection was also applied by Stott and Wells [64], van Waterschoot [85], and Schmidt and Haulick [74]. With the aim of reducing the sound quality deterioration due to noise injection, several attempts have been made to shape the spectrum of the injected noise signal such that it becomes less perceptible. Goertz [84] proposes to use A-weighted noise instead of white noise, while van Waterschoot [85] and Janse and Tchang [86] apply a time-varying noise shaping based on a psychoacoustic model. However, to obtain an AFC performance comparable to the methods using white noise injection, the psychoacoustically shaped noise has to be amplified to a

level at which it is found to be even more disturbing than white noise [85]. Decorrelation in the closed signal loop can also be achieved by including a nonlinear or time-varying signal operation in the electroacoustic forward path. Janse *et al.* [87]–[89] propose to use a frequency shifter or a periodic phase or delay modulator. The AFC robustness can then be increased since these decorrelating operations also have a stabilizing effect on the closed-loop system (see Section III-A). Another nonlinear decorrelation technique, which was adopted from the stereo AEC literature [90] by van Waterschoot *et al.* [91] and Schmidt and Haulick [74], consists in adding a half-wave rectified version of the loudspeaker signal to the original loudspeaker signal, yet was found to improve the AFC performance only marginally [91]. Finally, in the context of HA AFC applications, inserting a processing delay in the electroacoustic forward path has been proposed for reducing the correlation between the source and loudspeaker signals [79], [92]. The motivation for this approach is that the source and loudspeaker signal cross-correlation function is expected to decrease for increasing time lags, which is particularly the case for voiceless speech signals.

While most of the above decorrelation techniques are rather effective when applied in the closed signal loop, their effect on the sound quality may be detrimental. For this reason, there has been an increased interest in the application of decorrelating signal operations in the adaptive filtering circuit, such that the closed-loop signals remain unaffected. A first approach, which was proposed by Ortega *et al.* [93], [94], consists in having the adaptive filter preceded by a processing delay. The resulting decorrelation effect is similar to when a processing delay is inserted in the electroacoustic forward path. However, the delay length in the adaptive filtering circuit should not exceed the initial delay (i.e., the "dead time") in the acoustic feedback path impulse response (e.g., with the acoustic feedback path impulse response shown in Fig. 3, the maximum allowable processing delay would be 405 samples). A second approach consists in the use of decorrelating prefilters, that are designed to whiten the source signal component in the microphone signal. This approach was adopted from HA AFC research [80], [95], [96], and was applied to PA systems by van Waterschoot *et al.* [81], [91] and to in-car communication systems by Ortega *et al.* [97]. A fundamental difficulty lies in the concurrent identification of the optimal prefilter and the acoustic feedback path model from the closed-loop signals. This identification problem was tackled following a prediction-error-method (PEM)-based approach [98, Ch. 3], [99, Ch. 7] by Rombouts *et al.* [100]–[105]. The PEM-based AFC approach developed in [100]–[102] is based on a nonstationary all-pole source signal model, the inverse of which is then used as a time-varying FIR decorrelating prefilter in the AFC scheme. The robustness of the PEM-based AFC approach

was further improved in [103] by including some additional features such as adaptation control and the joint use of a foreground and background adaptive filter. Also, efficient subband and frequency domain implementations of the PEM-based AFC method were proposed in [103]. It was shown by van Waterschoot *et al.* [104], [105] how the convergence of the PEM-based AFC scheme can be improved even further by incorporating prior knowledge on the source signal and the acoustic feedback path through regularization.

In recent years, several remaining issues concerning the AFC approach have been analyzed and further improvements have been reported. The overall performance of the AFC approach may be improved by combining AFC with other acoustic feedback control methods and signal enhancement techniques, leading to so-called hybrid AFC methods. Ortega *et al.* [93], [94] propose the combination of AFC with a residual feedback and noise suppression postfilter, and this hybrid AFC scheme was further expanded by Janse and Belt [66] with an adaptive microphone array beamformer and an AIF. The combination of AFC with an NHS method is of particular interest due to the robustness of the NHS methods to system instability: Schmidt *et al.* [74], [106] and Cifani *et al.* [107], [108] use an ANF that operates on the AFC feedback-compensated signal, while Rombouts *et al.* [52], [103] apply a two-stage NHS method in which the howling detection is based a frequency analysis of the AFC feedback path estimate. The considerable computational complexity of the AFC approach in room acoustic applications is another issue that has recently been addressed. An interesting approach towards AFC complexity reduction was proposed by Okumura and Fujita [109] and consists in applying two or more parallel adaptive filters, preceded by a processing delay in the adaptive filtering circuit, to model a single acoustic feedback path. The first filter (which can be understood to model the late reverberation in the acoustic feedback path impulse response) has many coefficients that are adapted not very frequently using a transform domain approach, while the second filter (which then models the early reflections) is a short filter that is adapted at each instant using a time domain adaptive filtering algorithm. A final issue is related to AFC in audio applications: none of the aforementioned AFC methods has been designed to operate in a high-fidelity audio environment. When applying decorrelation in the closed signal loop, introducing signal distortion is unavoidable, while decorrelation techniques in the adaptive filtering circuit are typically based on the assumption that the source signal is a speech signal. Van Waterschoot and Moonen [110], [111] have recently proposed a novel PEM-based AFC method that is designed particularly for audio signals, but performs equally well in speech applications. The method is based on a cascade of two source signal models, where one models the tonal components in the source signal and the other one models the source signal noise components.
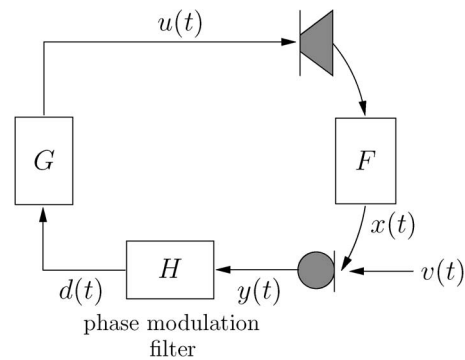

Fig. 5. *PFC by inserting a PM filter in the electroacoustic forward path.*

IV. PHASE-MODULATING FEEDBACK CONTROL

A. Concept

The goal of PFC is to control the phase of the microphone signal in such a way that every frequency component in the feedback signal has a different phase each time it arrives at the microphone after having traveled one cycle around the closed signal loop [17]. In this way, the phase condition in the Nyquist criterion (12) can be guaranteed not to hold for the same frequency at two successive instants, hence the closed-loop system stability can be improved, regardless of the magnitude condition (11). The PFC goal can be achieved by inserting a PM filter in the electroacoustic forward path, which operates directly on the microphone signal $y(t)$ and delivers an output signal $d(t)$ to the forward path processing unit $G(q, t)$; see Fig. 5.

The behavior of a PM filter can be analyzed elegantly using the theory of linear time-varying (LTV) filters [16], [17]. A discrete-time[4] LTV filter can be described in the time domain using the input–output relationship [112]

$$d(t) = \sum_{\tau=-\infty}^{\infty} h(\tau, t) y(t - \tau) \qquad (19)$$

with $h(\tau, t)$ the LTV filter's impulse response, which depends on both the observation instant t and the time difference τ between excitation and observation. If the LTV filter is moreover periodically time-varying (LPTV) with a period T_m that corresponds to an integer number of sampling periods, i.e., $T_m = NT_s$, then the periodic LTV frequency response

$$H(\omega, t) = \sum_{\tau=-\infty}^{\infty} h(\tau, t) e^{-j\omega\tau} \qquad (20)$$

---

[4]Note that in [16] and [17], the theory of LTV filters is described for the continuous-time case.

admits an N-point DFT representation with coefficients

$$\mathcal{H}(\omega, n) = \sum_{t=0}^{N-1} H(\omega, t)e^{-jn(2\pi/N)t} \qquad (21)$$

and the input–output relationship in (19) can be written in the frequency domain as follows [112]:

$$D(\omega) = \frac{1}{N}\sum_{n=0}^{N-1} \mathcal{H}(\omega - n\omega_m, n)Y(\omega - n\omega_m) \qquad (22)$$

with $\omega_m = 2\pi/N$ the LPTV filter fundamental frequency. In other words, the LPTV filter output spectrum is a sum of N frequency-weighted and frequency-shifted versions of the input spectrum. The LPTV filter frequency response DFT coefficients $\mathcal{H}(\omega, n)$ are usually referred to as the carrier response (for $n = 0$) and the sideband responses (for $n \neq 0$).

It can be seen from (22) that the output spectrum also contains a nonfrequency-shifted version of the input spectrum (for $n = 0$), which is undesirable in view of the acoustic feedback control performance [17]. The contribution of the nonfrequency-shifted version of the input spectrum to the total output spectrum is quantified using the so-called carrier suppression[5] [17]

$$CS[dB] = -10\log_{10}\left[\int_0^{2\pi} |\mathcal{H}(\omega, 0)|^2 \, d\omega\right] \qquad (23)$$

and it has been hypothesized that the CS corresponds to an upper bound for the increase in MSG that can be obtained using the PFC approach [17]. Another hypothesis stated in [17] is that a modulation scheme having a larger number of sideband responses with a relatively large power $|\mathcal{H}(\omega, n)|^2$ provides a better acoustic feedback control performance, since in this case, more input signal energy is shifted away from the original (carrier) frequency. However, this hypothesis is based on a continuous-time analysis and may not hold in a discrete-time context, since aliasing will fold all the input signal energy that has been shifted above the Nyquist frequency back to lower frequencies.

The following four PM techniques have been studied in the context of acoustic feedback control [17].
1) Sinusoidal PM [5], [14], [16], [17], [170]: a sinusoidal PM filter has a frequency response

$$H(\omega, t) = e^{j\beta \sin \omega_m t} \qquad (24)$$

which is characterized by frequency-independent carrier and sideband responses $\mathcal{H}(n)$ that correspond to the Bessel functions of the first kind and order n

$$\mathcal{H}(n) = J_n(\beta), \qquad n = 0, \ldots, N - 1. \qquad (25)$$

These functions are plotted as a function of the so-called modulation index β in Fig. 6.
2) Sinusoidal FM [15], [17]: the effect of a sinusoidal FM filter with a modulation frequency $f_m = \omega_m(f_s/2\pi)$ and a modulation depth Δ_f can be shown to be identical to the effect of a sinusoidal PM filter with the same modulation frequency f_m and a modulation index $\beta = \Delta_f/f_m$ [17].
3) FS [2], [7], [11], [12], [17], [18]: an FS device can either be viewed as a nonlinear time-invariant system or as an LPTV system. From the latter interpretation, it can be shown that an FS operation with a frequency shift of $f_m = \omega_m(f_s/2\pi)$ Hz corresponds to a PM operation with a phase function that increases linearly with time [11], [17], i.e.,

$$H(\omega, t) = e^{j\omega_m t} \qquad (26)$$
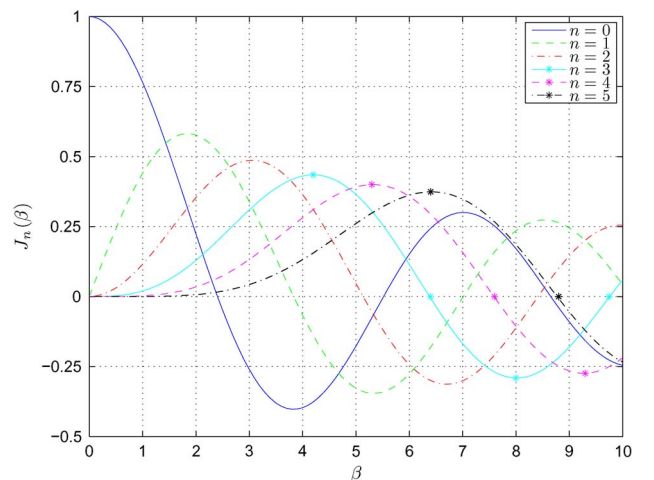
Fig. 6. *Bessel functions of the first kind for different orders* $n = 0, \ldots, 5$.

---

[5]Note that our definition of the carrier suppression as given in (23) is somewhat more general than the definition in [17], because we do not restrict the sideband responses $\mathcal{H}(\omega, n)$ to be independent of ω.

and, as a consequence

$$\mathcal{H}(n) \begin{cases} = 1, & n = 1 \qquad\qquad (27) \\ = 0, & n = 0, 2, \dots, N-1. \quad (28) \end{cases}$$

In other words, an FS device can be described as an LPTV filter with zero carrier response (i.e., $CS = \infty$) and only one nonzero sideband response (for $n = 1$).

4) Sinusoidal DM [14], [16], [17], [170]: a sinusoidal DM filter varies the input signal's time delay sinusoidally around a time delay offset τ_0 with a maximum time delay deviation Δ_τ and a modulation frequency ω_m, as can be seen in its frequency response

$$H(\omega, t) = e^{-j\omega(\tau_0 + \Delta_\tau \sin \omega_m t)}. \qquad (29)$$

This can be interpreted as a sinusoidal PM filter with the same modulation frequency ω_m and a modulation index $\beta = \omega \Delta_\tau$ that is proportional to the original (carrier) frequency ω. As a consequence, the corresponding carrier and sideband responses are frequency selective (as opposed to the frequency-independent PM, FM, and FS responses)

$$\begin{aligned} \mathcal{H}(\omega, n) &= J_n(\omega \Delta_\tau), \\ n &= 0, \dots, N-1. \end{aligned} \qquad (30)$$

From the above expression, it can be understood that a sinusoidal DM filter performs poorly in the low-frequency range since in this case, the carrier response has a much larger magnitude than the sideband responses (see Fig. 6).

B. Realization

The sinusoidal PM, sinusoidal FM, and FS filters are usually realized by operating on the so-called analytical representation of the microphone signal $y(t)$. In continuous time, the analytical signal $y_a\{\tau\}$ is defined as follows [113]:

$$y_a\{\tau\} = y\{\tau\} + j\hat{y}\{\tau\} \qquad (31)$$

where $\hat{y}\{\tau\}$ represents the Hilbert transform of $y\{\tau\}$. The corresponding discrete-time analytical signal $y_a(t)$ can be calculated in several ways. The first approach is to design a FIR filter $L(q)$ approximating the Hilbert transform such that an approximation to $\hat{y}(t)$ can be calculated as $L(q)y(t)$ [114], and then the discrete-time analytical signal can be obtained as $y_a(t) = y(t) + jL(q)y(t)$. Since the so-called

Hilbert filter $L(q)$ is noncausal, a processing delay of half the filter length of $L(q)$ has to be introduced in the signal path [12]. Another drawback is that this approach does not preserve the orthogonality between $y(t)$ and $\hat{y}(t)$ which can be obtained in the continuous-time case [115]. The second approach is to design two complex FIR filters, so-called dual quadrature FIR filters, that produce orthogonal approximations to $y(t)$ and $\hat{y}(t)$, respectively, which are then added according to (31) [116]. Unfortunately, this approach does not preserve the original data since the real part of the discrete-time analytical signal is not exactly equal to $y(t)$ [115]. In the third approach, which combines the desirable properties of original data preservation in the real part and orthogonality between the real and imaginary part of the discrete-time analytical signal, $y_a(t)$ is approximated as the inverse DFT of a one-sided discrete spectrum (with zero negative frequency content) that is calculated using the DFT of the original signal $y(t)$ [115]. This approach is frame based, hence a processing delay equal to the frame size minus the frame overlap is required. We will use this latter approach for the PFC evaluation in Section VII.

Given the discrete-time analytical signal $y_a(t) = y(t) + j\hat{y}(t)$, the output signal of the PM, FM, and FS filters can be calculated by modulating $y_a(t)$ with the LPTV frequency response $H(\omega, t)$, and then taking the real part (denoted with $\mathrm{Re}\{\cdot\}$) [18], i.e.,

$$d(t) = \mathrm{Re}\{y_a(t)H(\omega, t)\}. \qquad (32)$$

Using (24) and (26), this leads to

$$d(t) = y(t)\cos\phi(t) - \hat{y}(t)\sin\phi(t) \text{ with}$$

$$\phi(t) = \begin{cases} \beta \sin \omega_m t, & \text{for sinusoidal PM} \quad (33) \\ \frac{\Delta_f}{f_m} \sin \omega_m t, & \text{for sinusoidal FM} \quad (34) \\ \omega_m t, & \text{for FS.} \qquad\qquad (35) \end{cases}$$

A sinusoidal DM filter can be realized by directly operating on the microphone signal $y(t)$, which is then fed to a variable-length delay line. Such delay lines have also been used for realizing DM-based digital audio effects such as vibrato, flanging, and chorus; see, e.g., [117]–[119]. The sinusoidal DM variable-length delay line has an LPTV transfer function that can be approximated as the cascade of an integer delay of K samples and a fractional delay of l/D samples [117]–[119], where D is denoted as the interpolation ratio and $l = 0, \dots, D-1$ is the fractional phase

$$H(q, t) = q^{-(\tau_0 + \Delta_\tau \sin \omega_m t)} \qquad (36)$$

$$\approx q^{-K}q^{-l/D}. \qquad (37)$$

The fractional part of the transfer function in (37) can be realized using any of the available methods for fractional delay filter design [120], e.g., using linear [117]–[119], all-pass [117]–[119], or spline [118], [119] interpolation filters. We will use a linear FIR interpolation filter that is a Hamming-windowed, truncated (length-$2I$) approximation of the ideal sinc-like interpolation filter [120]

$$H(q, t) = q^{-K} \sum_{i=-I}^{I-1} w_h(i + l/D) \text{sinc}(i + l/D) q^i \qquad (38)$$

where $w_h(t)$ denotes the Hamming window, centered at $t = 0$, and the integer delay and fractional phase are chosen as $K = \lfloor \tau_0 + \Delta_\tau \sin \omega_m t \rfloor$ and $l = [(\tau_0 + \Delta_\tau \sin \omega_m t - K)/D]$, with $\lfloor \cdot \rfloor$ the floor function and $[\cdot]$ the nearest integer function, respectively. Note that τ_0, Δ_τ, and I should be chosen such that $\tau_0 - \Delta_\tau \geq I - 1$ to guarantee causality of the sinusoidal DM filter.

C. Discussion

The main strength of the PFC approach is its simplicity, both conceptually and computationally. The design of a PFC system requires little effort, since only the modulation technique (PM, FM, FS, or DM) and a few parameter values have to be decided on. The main computational load lies in the calculation of the analytical microphone signal (for PM, FM, and FS) and the fractional delay interpolation filtering (for DM), which should not be a barrier for real-time implementation. Moreover, the PFC approach does not involve any form of learning or adaptivity, such that it behaves in a completely deterministic way, which is beneficial in terms of robustness.

The choice of the modulation technique depends on the envisaged application. The FS technique is known to generally deliver a larger MSG increase than the other modulation techniques, but is perceptually inappropriate for music applications [18], [86]. The MSG increase obtained with modulation techniques that have a larger number of sideband responses (PM, FM, and DM, with a sufficiently large β) appears to be more or less independent of the modulation frequency ω_m, such that these techniques can operate at a lower value of ω_m as compared to FS, which is perceptually advantageous [17]. DM is known to perform poorly at low signal frequencies, such that it should preferably be combined with another modulation technique or even with a non-PM-based acoustic feedback control method [16].

For a given modulation technique, the main parameters determining the PFC performance are the modulation frequency ω_m and the modulation index β. It has been theoretically shown and experimentally verified that in the case of FS, an optimal value of the frequency shift $f_m = \omega_m(f_s/2\pi)$ is around $4/T_{60}$ Hz, with T_{60} the room reverberation time in seconds [2]. The optimal value for f_m

is less related to the reverberation time in the case of PM, FM, and DM, and values as low as 0.5 Hz may provide a satisfactory MSG increase, especially at high modulation index values [17]. The influence of the modulation index β in the case of PM, FM, and DM is governed by two effects [17]: as the value of β approaches the zeros of $J_0(\beta)$ (e.g., see Fig. 6), the CS and hence the maximum achievable MSG increase become larger, and on the other hand, a larger value of β leads to a larger number of influential sideband responses which (at least in the continuous-time case) can be expected to improve the acoustic feedback control performance [17]. The former effect provides an explanation for the value of $\beta = 2.4$ having been suggested as an optimal choice in early studies on PFC using sinudoidal PM [5], [14], [170].

Finally, the PFC method has three major drawbacks. First, the achievable MSG increase is limited. An MSG increase of 12 dB has been found to be the theoretical maximum using FS in a typical room acoustic sound reinforcement system, and moreover, to avoid the FS effect to be clearly audible, a system equipped with an FS filter should operate 6 dB below the MSG, reducing the practically realizable MSG increase to 6 dB [2]. Similar MSG increase values (around 6 dB) were found in experiments using the other modulation techniques (PM, FM, and DM), as reported in several studies [14]–[17], [121], [170]. A second drawback is that inserting a PM filter in the electroacoustic forward path unavoidably leads to signal distortion, the perceptual consequences of which may be detrimental, particularly in audio applications [19]. A third disadvantage is the fact that in multichannel systems, the stability improvement obtained with PFC has been shown to decrease as the number of channels increases [18], hence the practical use of PFC in large-scale sound reinforcement systems (e.g., PA or RE systems) is expected to be limited.

V. NOTCH-FILTER-BASED HOWLING SUPPRESSION

A. Concept

The objective of the NHS method can be either to prevent the closed-loop system from becoming unstable by reducing the loop gain $|G(\omega, t)F(\omega, t)|$ in the neighborhood of critical frequencies, or to stabilize the system and suppress howling after oscillations have occurred. The former objective requires a proactive approach to instability detection, while the latter approach is reactive in the sense that notch filters are activated only after the detection of howling. We will mainly focus on the reactive approach to NHS, which is much more widespread than the proactive approach. Also, the emphasis is on two-stage NHS methods, since these are much more popular as compared to the ANF-based one-stage NHS methods. In a two-stage NHS method, the microphone signal $y(t)$ is first
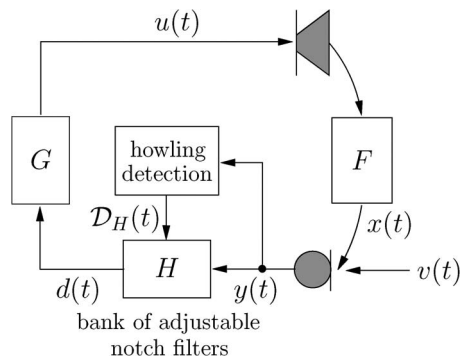
Fig. 7. *Two-stage NHS by feeding the microphone signal to a howling detection algorithm, which forwards a set of design parameters $\mathcal{D}_H(t)$ to a bank of adjustable notch filters $H(q,t)$ that is inserted in the electroacoustic forward path.*

processed by a howling detection algorithm, which forwards a set of design parameters $\mathcal{D}_H(t)$ to a bank of adjustable notch filters $H(q,t)$ that is inserted in the electroacoustic forward path; see Fig. 7.

The howling detection algorithm is the most critical part of the two-stage NHS method. Since howling is known to consist of sinusoidal signal components, the detection of howling is based on a frequency analysis of the microphone signal. It can be understood that howling components can be recognized as signal components having a large magnitude in the frequency domain. However, voiced speech components and tonal music components also have this property, hence it is crucial to discriminate howling

components from tonal source signal components. We will use an example to illustrate the signal attributes that can be used to discriminate between howling and tonal components. Let us consider a single-channel closed-loop system defined by the acoustic feedback path shown in Figs. 3 and 4, and an electroacoustic forward path consisting of a cascade of a unit delay and a broadband gain factor $K = 5.53$ dB. The loop gain of this system is shown in Fig. 8(a) for $f \in [0, 3]$ kHz. It can be observed that the Nyquist magnitude condition (11) is fulfilled for a frequency value just above 500 Hz, such that an oscillation at this frequency can be expected. When an audio signal fragment, more specifically a 10-s excerpt from the Partita No. 2 in D minor (Allemande) for solo violin by J. S. Bach, is applied as a source signal in the closed-loop system, the corresponding microphone signal has a spectrogram as shown in Fig. 8(b) (zooming in on the frequency region $f \in [0, 3]$ kHz). The buildup of a howling component at a frequency slightly above 500 Hz is clearly visible from the spectrogram. Moreover, it can be observed that the howling component has some distinct features that may be used to discriminate it from the tonal source signal components. Spectral features include the fact that the howling component has a relatively large magnitude, and does not have any harmonic or subharmonic frequency components. Temporal features typical to the howling component are its long duration and its increasing magnitude with time.

Apart from detecting howling components in the microphone signal spectrum, the howling detection algorithm in the two-stage NHS method shown in Fig. 7 also calculates some features of the detected howling
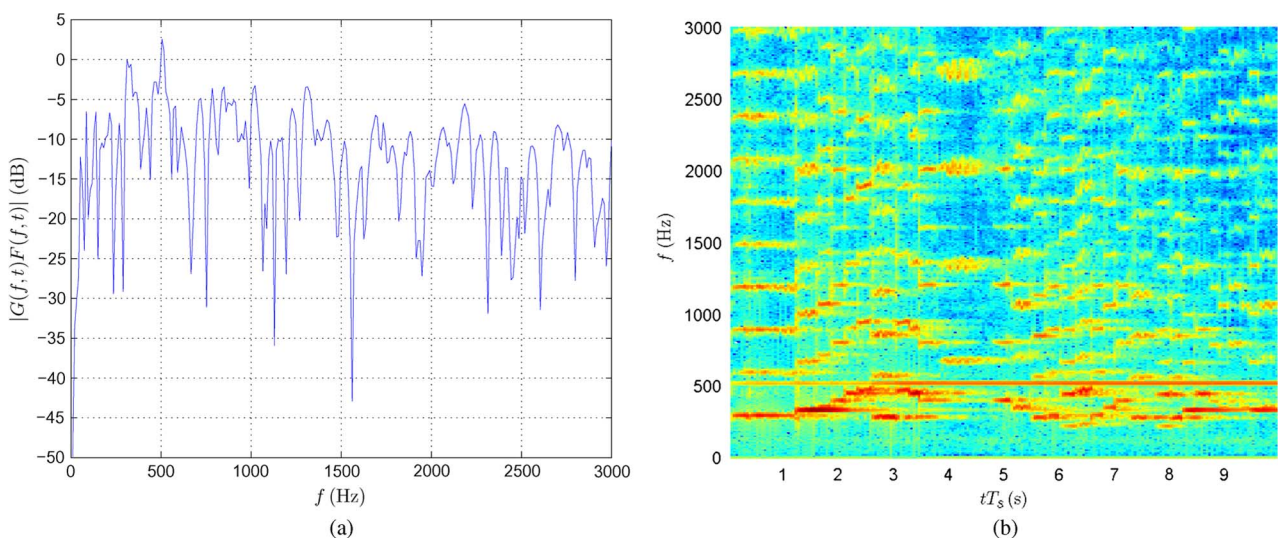
Fig. 8. *Example to recognize discriminating features between howling and tonal components (zooming in on the frequency region $f \in [0, 3]$ kHz): (a) loop gain of the unstable closed-loop system defined by the acoustic feedback path response shown in Fig. 4 and a flat electroacoustic forward path response with gain factor $K = 5.53$ dB, and (b) microphone signal spectrogram after feeding an audio source signal to the unstable closed-loop system.*

components that are subsequently used to design appropriate notch filters. The set of design parameters $\mathcal{D}_{\mathcal{H}}(t)$ typically includes the howling components' frequency and magnitude values. The notch filters are then designed to have center frequencies corresponding to the howling component frequencies and notch depth values depending on the howling component magnitude values. The notch filters' 3-dB bandwidth is usually fixed to a value in the range of 1/10–1/60 octave. A more narrowband notch filter has the advantage of removing less of the desired source signal components, but requires a more accurate howling component frequency estimation.

B. Realization

1) Howling Detection: We assume that the howling detection algorithm operates in a frame-based manner, on microphone signal frames with a frame length of M samples and a frame hop size of P samples (i.e., a frame overlap of $M - P$ samples). At time t, the data in the microphone signal frame can then be represented by the vector [which is not to be confused with the multichannel microphone signal vector $\bar{\mathbf{y}}(t)$ defined in (4)]

$$\mathbf{y}(t) = [y(t + P - M) \quad \ldots \quad y(t + P - 1)]^T \quad (39)$$

and the short-term microphone signal spectrum can be obtained as the DFT of the data in $\mathbf{y}(t)$, i.e.,

$$Y(\omega_k, t) = \sum_{n=0}^{M-1} w(t_n) y(t_n) e^{-j\omega_k t_n}, \quad (40)$$
$$k = 0, \ldots, M - 1$$

with $\omega_k \triangleq 2\pi k/M$ and $t_n \triangleq t + P - M + n$. The microphone signal DFT in (40) is generally calculated using the FFT algorithm, and includes a window function $w(t_n)$ to reduce the spectral leakage [122] (e.g., a Blackman window has successfully been applied to audio signal processing [123][6]). Alternatively, a parametric frequency estimation method may be applied instead of the nonparametric (DFT-based) approach to obtain a good frequency resolution with relatively short signal frames [42]. Also, a frequency-warped DFT [124] may be used to improve the frequency resolution in the low-frequency region [50], [51]. The choice of the signal framing parameters M and P has a rather profound influence on the performance of the howling detection. Small values for the frame length M have been proposed to allow for

very quick howling detection (e.g., $M = 128$, corresponding to 4 ms at $f_s = $ 32 kHz [28]–[30]), such that howling may potentially be detected before it is actually perceived [28]–[30]. On the other hand, larger values for M provide a better frequency resolution in the microphone signal DFT spectrum estimate (e.g., $M = 4096$, corresponding to 92.9 ms at $f_s = 44.1$ kHz [32], [34] or to 85.3 ms at $f_s = 48$ kHz [44], [45]), which is necessary when working with very narrowband notch filters such as the 1/60 octave filters used in [50] and [51]. A large frame hop size P may result in a large time lag between howling detection and notch filtering, unless a P-sample delay is inserted in the electroacoustic forward path. On the other hand, a small value for P leads to an increase in computational complexity since the howling detection algorithm is then executed more often. Generally, a 25%–50% frame overlap ($P = 3M/4, \ldots, M/2$) is found to be a good compromise.

Based on the DFT-based microphone signal spectrum estimation, a predefined number N of spectral peaks is identified from the spectrum estimate, with N typically chosen in the range 1–10. These N frequency components are termed "candidate howling components" and their radial frequency values are collected in the set $\mathcal{D}_{\breve{\omega}}(t) = \{\breve{\omega}_i\}_{i=1}^N$. A spectral peak picking algorithm is usually applied to find the candidate howling components. A more advanced approach consists in selecting frequency components that have a consistently increasing magnitude in successive signal frames. This is possible by applying a so-called "ballistics" procedure [44], [45] before executing the peak picking algorithm. The following spectral and temporal features of the microphone signal have been proposed to determine whether a candidate howling component indeed corresponds to a howling component or rather to a source signal tonal component.

- The peak-to-threshold power ratio (PTPR) [35], [36], [44], [45] is a spectral feature that determines the ratio of the candidate howling component power $|Y(\breve{\omega}_i, t)|^2$ and a fixed absolute power threshold P_0, i.e.,

$$\text{PTPR}(\breve{\omega}_i, t)[\text{dB}] = 10 \log_{10} \frac{|Y(\breve{\omega}_i, t)|^2}{P_0}. \quad (41)$$

Howling is detected at the frequency $\breve{\omega}_i$ if $\text{PTPR}(\breve{\omega}_i, t) \geq 0$ dB. The rationale behind using the PTPR for howling detection is that howling should only be suppressed when it appears with a minimum loudness [44], [45]. The absolute power threshold P_0 depends on the particular sound reinforcement scenario at hand, e.g., a value of $10 \log_{10} P_0 = 85$ dB SPL was suggested in [44]

---

[6]http://ccrma.stanford.edu/~jos/mdft/Use_Blackman_Window.html

and [45] for a loudspeaker–microphone distance of 1 m.

- The peak-to-average power ratio (PAPR) [25], [26], [35]–[38], [45], [49]–[51], [55], [56] is a spectral feature that determines the ratio of the candidate howling component power $|Y(\breve{\omega}_i, t)|^2$ and the average microphone signal power $\hat{P}_y(t)$, i.e.,

$$\text{PAPR}(\breve{\omega}_i, t)[\text{dB}] = 10 \log_{10} \frac{|Y(\breve{\omega}_i, t)|^2}{\hat{P}_y(t)} \quad (42)$$

with

$$\hat{P}_y(t) = \frac{1}{M} \sum_{k=0}^{M-1} |Y(\omega_k, t)|^2. \quad (43)$$

The ith candidate howling component is identified as a howling component if the PAPR exceeds a predetermined threshold, i.e., $\text{PAPR}(\breve{\omega}_i, t) \geq T_{\text{PAPR}}$. The PAPR feature is probably the most widely used feature for howling detection, and different values for the threshold have been proposed, e.g., $T_{\text{PAPR}} = 6$ dB [25], [26], $T_{\text{PAPR}} = 10 \log_{10}(M/150)^2$ dB [44], [45], and $T_{\text{PAPR}} = 10$ dB [49]. Kawamura *et al.* [37], [38] propose the use of a variable threshold $T_{\text{PAPR}}(t)$ that is adapted online, based on estimates of the background noise spectrum, the source signal spectrum, the reverberation time, and the acoustic feedback path response. It is also suggested in [37] and [38] to remove the Q_P largest frequency components from the spectrum $Y(\omega_k, t)$ before estimating the average signal power $\hat{P}_y(t)$ in (43), the value of Q_P depending on the bandwidth of the frequency analysis. Yet another way of estimating the average microphone signal power $\hat{P}_y(t)$ was suggested in [55] and [56], i.e.,

$$\hat{P}_y(t) = \left(\frac{1}{M} \sum_{k=0}^{M-1} |Y(\omega_k, t)|^2 \right)$$
$$+ 2 \sqrt{\frac{1}{M} \sum_{k=0}^{M-1} \left(|Y(\omega_k, t)|^2 - \frac{1}{M} \sum_{m=0}^{M-1} |Y(\omega_m, t)|^2 \right)^2} \quad (44)$$

which should be particularly useful when the source signal has a Gaussian probability density function (pdf).

- The peak-to-harmonic power ratio (PHPR) [32], [34]–[36] is a spectral feature that determines the ratio of the candidate howling component power $|Y(\breve{\omega}_i, t)|^2$ and its mth (sub)harmonic component power $|Y(m\breve{\omega}_i, t)|^2$, i.e.,

$$\text{PHPR}(\breve{\omega}_i, t, m)[\text{dB}] = 10 \log_{10} \frac{|Y(\breve{\omega}_i, t)|^2}{|Y(m\breve{\omega}_i, t)|^2}. \quad (45)$$

In [32] and [34], howling is detected at the frequency $\breve{\omega}_i$ if the PHPR exceeds a predetermined threshold for the second, third, and fourth harmonics and the 0.5th and 1.5th subharmonics, i.e., if

$$\bigcap_{m \in \{0.5, 1.5, 2, 3, 4\}} [\text{PHPR}(\breve{\omega}_i, t, m) \geq T_{\text{PHPR}}] = 1 \quad (46)$$

with $T_{\text{PHPR}} = 33$ dB. In [35] and [36], a simpler howling detection criterion $\text{PHPR}(\breve{\omega}_i, t, 2) \geq T_{\text{PHPR}}$ is used.

- The peak-to-neighboring power ratio (PNPR) [22], [28]–[30] is a spectral feature that determines the ratio of the candidate howling component power $|Y(\breve{\omega}_i, t)|^2$ and its mth neighboring frequency component power $|Y(\breve{\omega}_i + 2\pi m/M, t)|^2$, i.e.,

$$\text{PNPR}(\breve{\omega}_i, t, m)[\text{dB}]$$
$$= 10 \log_{10} \frac{|Y(\breve{\omega}_i, t)|^2}{|Y(\breve{\omega}_i + 2\pi m/M, t)|^2}. \quad (47)$$

In [22], $\breve{\omega}_i$ is determined to be a howling frequency if the PNPR in two adjacent frequency bins on either side of the candidate howling component is consistently above two predetermined thresholds and the PTPR is above 0 dB, i.e., if

$$\left\{ [\text{PTPR}(\breve{\omega}_i, t) \geq 0 \text{ dB}] \wedge \right.$$
$$\left. \bigcap_{m \in \{\pm 1, \pm 2\}} [\text{PNPR}(\breve{\omega}_i, t, m) \geq T_{\text{PNPR}}(|m|)] \right\} = 1. \quad (48)$$

In [28]–[30], howling is detected based on a so-called "peakness" feature, which reflects the time-averaged probability (over eight signal frames) that the PNPR, averaged over six neighboring

frequency bins on either side of $\breve{\omega}_i$ (excluding the closest neighbor on either side), exceeds a 15-dB threshold, i.e.,

$$
\begin{aligned}
\text{peakness}(\breve{\omega}_i, t) \\
= \sum_{j=0}^{7} \frac{1}{16} \Bigg\{ & \left[\frac{1}{6} \sum_{m=2}^{7} \text{PNPR}(\breve{\omega}_i, t-jP, m) \geq 15 \text{ dB} \right] \\
& + \left[\frac{1}{6} \sum_{m=-7}^{-2} \text{PNPR}(\breve{\omega}_i, t-jP, m) \geq 15 \text{ dB} \right] \Bigg\}.
\end{aligned}
\tag{49}
$$

- The interframe peak magnitude persistence (IPMP) [20], [21], [32], [34], [37], [38], [49] is a temporal feature based on counting in how many frames out of Q_M past signal frames the frequency $\breve{\omega}_i$ is in the set of candidate howling frequencies, i.e.,

$$
\text{IPMP}(\breve{\omega}_i, t) = \frac{\sum_{j=0}^{Q_M-1} [\breve{\omega}_i \in \mathcal{D}_{\breve{\omega}}(t-jP)]}{Q_M}. \tag{50}
$$

Howling is usually detected if $\text{IPMP}(\breve{\omega}_i, t) = 1$ [20], [21], [37], [38], [49], with, e.g., $Q_M = 3$ [49]. In [32] and [34], a howling detection criterion $\text{IPMP}(\breve{\omega}_i, t) \geq 3/5$ is proposed with $Q_M = 5$.

- The interframe magnitude slope deviation (IMSD) [28]–[30] is a temporal feature that determines the deviation (over Q_M successive signal frames) of the slope, which is defined by averaging magnitude difference values of a candidate howling component, where the differentiation is carried out between an old signal frame and more recent signal frames, i.e.,

$$
\begin{aligned}
\text{IMSD}(\breve{\omega}_i, t) = \frac{1}{Q_M - 1} \sum_{m=1}^{Q_M-1} \Bigg[& \frac{1}{Q_M} \sum_{j=0}^{Q_M-1} \frac{1}{Q_M - j} \\
& \times (20 \log_{10} |Y(\breve{\omega}_i, t-jP)| \\
& \quad - 20 \log_{10} |Y(\breve{\omega}_i, t-Q_MP)|) \\
& - \frac{1}{m} \sum_{j=0}^{m-1} \frac{1}{m - j} \\
& \times (20 \log_{10} |Y(\breve{\omega}_i, t-jP)| \\
& \quad - 20 \log_{10} |Y(\breve{\omega}_i, t-mP)|) \Bigg]. \tag{51}
\end{aligned}
$$

Small values for the IMSD are characteristic of howling components since these exhibit a nearly linear (decibel-scale) magnitude increase in time, hence a nearly constant slope can be expected. A detection threshold of 0.05 has been proposed in [28], such that howling is detected when $|\text{IMSD}(\breve{\omega}_i, t)| \leq 0.05$, with $Q_M = 7$.

The complete howling detection algorithm is summarized in Fig. 9. Obviously, any combination of the above spectral and temporal features may be used to discriminate between howling and tonal components. In most of the existing NHS methods, at least one spectral and one temporal feature are taken into account for detecting howling.

2) Notch Filtering: When howling has been detected, a notch filter has to be activated to suppress the howling component and stabilize the closed-loop system. The most commonly used notch filter structure in NHS is the second-order IIR (i.e., biquadratic) filter structure

$$
H_l(q, t) = \frac{b_l^{(0)}(t) + b_l^{(1)}(t)q^{-1} + b_l^{(2)}(t)q^{-2}}{1 + a_l^{(1)}(t)q^{-1} + a_l^{(2)}(t)q^{-2}}. \tag{52}
$$

The bank of adjustable notch filters that is inserted in the electroacoustic forward path, as shown in Fig. 7, then consists of a cascade of $n_H/2$ such filters, i.e.,

$$
H(q, t) = \prod_{l=1}^{n_H/2} H_l(q, t) \tag{53}
$$

with n_H the resulting order of the cascade filter.

The notch filter design procedure consists of two parts. First the set of design parameters $\mathcal{D}_H(t)$ delivered by the howling detection algorithm has to be mapped to a set of filter specifications, which are then translated into filter coefficient values. A biquadratic notch filter has five coefficients, which depend on a set of six filter specifications [125]: the (radial) center frequency $\omega_{c,l}$, the (radial) bandwidth B_l, the notch gain $G_{c,l}$, the gain at the band edges $G_{B,l}$, the gain at direct current (dc) $G_{0,l}$, and the gain at the Nyquist frequency $G_{\pi,l}$. If we fix the latter two variables to $G_{0,l} = G_{\pi,l} = 0$ dB and the gain at the band edges to $G_{B,l} = G_{c,l} + 3$ dB in case $G_{c,l} \leq -6$ dB, or to $G_{B,l} = G_{c,l}/2$ in case $G_{c,l} \geq -6$ dB (thereby adopting Moorers bandwidth definition [126]), then only the first three filter specifications remain.

The set of design parameters $\mathcal{D}_H(t)$ should always contain the radial frequencies $\{\breve{\omega}_i\}_{i \in \mathcal{I}_H(t)}$ of the howling components that have been identified in the howling detection algorithm, where $\mathcal{I}_H(t) \subseteq \{1, \ldots, N\}$ denotes the set of indices for which howling has been detected. For each howling component, a notch filter should be activated, with a center frequency corresponding to the
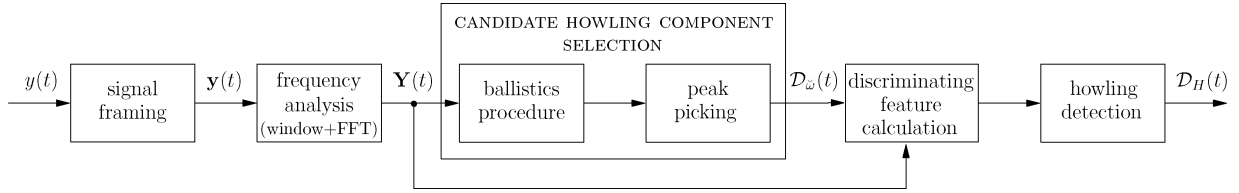
Fig. 9. *Howling detection algorithm for two-stage NHS method: from the microphone signal $y(t)$, a set of notch filter design parameters $D_H(t)$ is calculated.*

howling frequency. It is desirable to compensate for the limited frequency resolution of the microphone signal DFT by linearly interpolating the notch filter center frequency, using the DFT information from frequency bins adjacent to the identified howling component [44], [45], e.g.,

$$\omega_{c,l} = \breve{\omega}_i + \frac{2\pi}{M}\left(\frac{|Y(\breve{\omega}_i + 2\pi/M)| - |Y(\breve{\omega}_i - 2\pi/M)|}{|Y(\breve{\omega}_i - 2\pi/M)| + |Y(\breve{\omega}_i)| + |Y(\breve{\omega}_i + 2\pi/M)|}\right).$$
(54)

In this case, the DFT magnitude values $|Y(\breve{\omega}_i - 2\pi/M)|$, $|Y(\breve{\omega}_i)|$, and $|Y(\breve{\omega}_i + 2\pi/M)|$ should also appear in the set of design parameters $\mathcal{D}_H(t)$. The DFT magnitude information may also be used to determine the notch gain $G_{c,l}$, however, it is common practice to work with fixed notch gain values that are independent of the howling component magnitude. Typically, when a new howling component has been detected (i.e., a howling component at a frequency that has not occurred before), the notch gain is set to an initial value $G_{c,l}^{(0)}$, e.g., $G_{c,l}^{(0)} = -3$ dB [32], [34] or $G_{c,l}^{(0)} = -6$ dB [44], [45]. If howling persists or reoccurs at a frequency close to a previously identified howling frequency, then the gain is decreased with $\Delta G_{c,l}$ dB, e.g., $\Delta G_{c,l} = -3$ dB [32], [34] or $\Delta G_{c,l} = -6$ dB [44], [45]. Finally, the radial notch filter bandwidth B_l is usually chosen proportional to the center frequency, such that the filter has a constant Q factor. The octave bandwidth is then also constant and is typically chosen in the range 1/10–1/60 octave, e.g., 1/10 octave [32], [34], [52], [103], 1/20 octave [52], [103], or 1/60 octave [50], [51].

Finally, the filter specifications $\mathcal{S}_{H_l}(t) = \{\omega_{c,l}, B_l, G_{c,l}\}$ have to be translated to a set of filter coefficients $\mathcal{C}_{H_l}(t) = \{b_l^{(0)}(t), b_l^{(1)}(t), b_l^{(2)}(t), a_l^{(1)}(t), a_l^{(2)}(t)\}$. Most notch filter design methods are based on a bilinear transform of either an analog notch filter transfer function [127]–[133], or a digital notch filter transfer function centered at $\omega_c = \pi/2$ [126]. A novel design procedure for biquadratic notch filters was recently proposed, which operates directly in the digital domain using a technique known as pole-zero placement [125]. This design procedure, which is equally accurate yet more intuitive than the bilinear-transform-based design methods, will be applied in the evaluation of the NHS method in Section VII. The complete notch filter design procedure for the two-stage NHS method is shown schematically in Fig. 10.

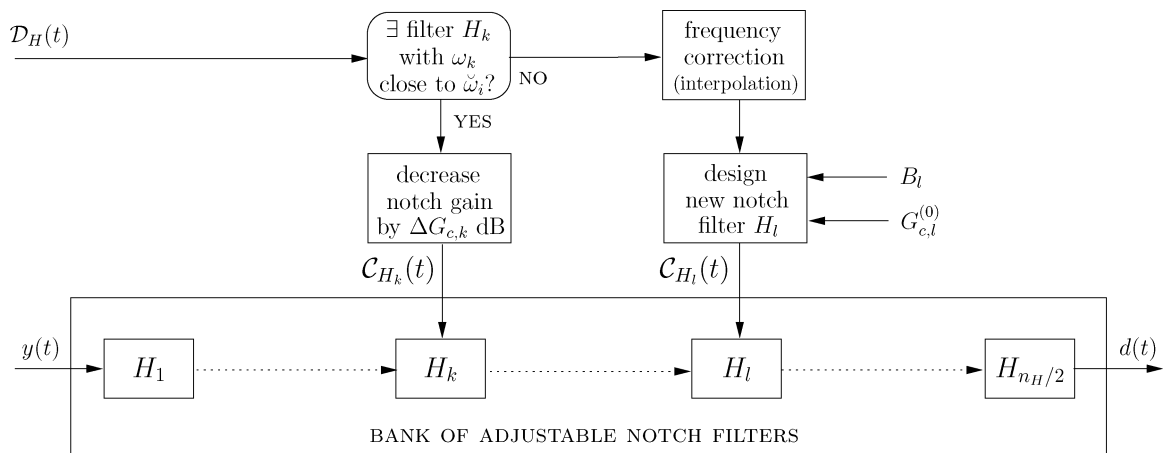

Fig. 10. *Notch filter algorithm for two-stage NHS method: the microphone signal $y(t)$ is filtered in a bank of adjustable notch filters, designed using the design parameters in $D_H(t)$, resulting in the howling-compensated signal $d(t)$.*

C. Initialization

In the PFC method for acoustic feedback control, the optimal values for the algorithm parameters (i.e., the modulation frequency ω_m and modulation index β) were found to be independent of the specific acoustic feedback path characteristics. The optimal notch filter coefficients in the NHS method, however, depend heavily on the spectral properties of the acoustic feedback path. It has long been known that some of the spectral peaks in the acoustic feedback path magnitude response $|F(\omega, t)|$ originate from reflections depending on the room boundaries only and are hence independent of the position of loudspeakers, microphones, and other objects in the room.[7] For this reason, manual equalization and notch filtering is largely performed during initialization (e.g., "ringing out" a PA system during sound check [135]) and fixed filters are applied to compensate for the major room resonances.

Similarly, a number of notch filters in the NHS method may be fixed to the so-called "eigenfrequencies" of the room, while the remaining notch filters can be adjusted to suppress variable-frequency howling components, which are due to, e.g., microphone movements [32], [34], [42]. The fixed notch filter design parameters should then be determined during the initialization of the sound reinforcement system, which is usually accomplished by feeding a white noise signal to the loudspeakers at a relatively high amplifier gain and subsequently identifying persisting spectral components in the microphone signal [23], [24], [32], [34].

The variable notch filters differ from the fixed notch filters in that they can be activated and deactivated during normal operation of the sound reinforcement system. While an extensive part of the NHS literature is devoted to strategies for the activation of these notch filters (i.e., after howling detection), hardly any research results are available dealing with the criteria for notch filter deactivation. One such deactivation criterion was proposed by Terada and Murase [27] in the context of AEQ for HA applications, and consists in deactivating the equalization filters after a time period that is inversely proportional to the time period between two successive occurrences of howling. Finally, we should note that the activation of a notch filter in the electroacoustic forward path leads to transient components in the loudspeaker signal, which may be perceived as short-lived ringing artifacts [35], [36]. This effect can be avoided by gradually activating and deactivating the notch filters, e.g., using a leaky integrator [55], [56].

D. Discussion

The NHS approach has many strengths, the most important one being its robustness. Unlike other acoustic feedback control methods, NHS methods have the powerful property of being able to stabilize an unstable system without having to reduce the broadband gain. For

this reason, it is advisable that a sound reinforcement system that is operated with a different acoustic feedback control method (e.g., PFC or AFC) be supplemented with an NHS method, which should then be activated when the system stability cannot be restored using the PFC or AFC method. As for computational requirements, the NHS approach has a moderate complexity, in between the cheap PFC approach and the expensive AFC approach. The main computational load is in the frequency analysis and can be governed by properly choosing the frame length M and hop size P. Another attractive property is that the extension of the NHS approach to multichannel systems is relatively straightforward. In the multichannel case, it is usually more efficient to have the howling detection and notch filtering algorithms operate on the mixed signals instead of on the microphone signals, since the number of channels is usually reduced after mixing. Alternatively, both the mixed signals and the individual microphone signals can be used for howling detection, where the latter may be analyzed at a lower frequency resolution [53].

A difficulty that arises when applying an NHS method for acoustic feedback control is the multitude of algorithm parameters that have to be set, namely the frame length and hop size, the number of candidate howling components selected in each signal frame, the combination of discriminating features, the thresholds for howling detection, the number of fixed/variable notch filters to use, etc. Unfortunately, few guidelines are available for setting these algorithm parameters. As many NHS methods are described in patents, very few experimental results and no true comparisons between different NHS methods are available. A comparison of three NHS methods with particular choices for the algorithm parameters will be provided in Section VII, but obviously, many more combinations are possible.

The major shortcoming of the NHS approach is that it cannot deliver an MSG increase that is substantially larger than the MSG increase obtained with the PFC approach. At most, i.e., when all the spectral peaks in the loop gain could be removed, an MSG increase of 10 dB could be expected based on the statistical analysis by Schroeder [2]. In practice, this maximum value will never be attained since it is nearly impossible to completely flatten the loop gain and still retain an acceptable degree of sound quality. As an example, if we would increase the gain in the single-channel system associated with the acoustic feedback path shown in Fig. 4(a) to a value that is 10 dB above the MSG without acoustic feedback control, then over 20 frequencies would satisfy the magnitude condition (11) in the Nyquist criterion, most of these lying in the 100–1500-Hz frequency region. Applying the NHS approach would then lead to a broadband attenuation in the 100–1500-Hz band, which would be detrimental for the sound quality (e.g., speech intelligibility). The limited achievable increase in MSG is also observed in manual notch filtering methods, where values of 5–8 dB have been obtained [6]. Finally, in

---

[7]This concept has been formalized in the so-called common acoustical pole and zero (CAPZ) model for room transfer functions [134].

terms of sound quality, the signal distortion due to notch filtering is reasonable if the number of filters that are applied concurrently is small and if the notch filter bandwidths are small. In fact, the main decrease in sound quality is due to the reactive nature of most NHS methods, i.e., howling can usually be perceived before it can be suppressed. From this point of view, proactive NHS methods can be viewed as promising acoustic feedback control solutions (see, e.g., [52] and [59]), however, their current applicability is limited due to their high computational complexity, comparable to the AFC complexity.

VI. ADAPTIVE FEEDBACK CANCELLATION

A. Concept

In a sound reinforcement system, the microphone signal $y(t)$ consists of a source signal component $v(t)$ and a feedback signal component $x(t)$, the latter denoting the entire signal that is fed back from the loudspeaker to the microphone. The AFC approach to acoustic feedback control is aimed at predicting the feedback signal component and then subtracting this prediction from the microphone signal. The predicted feedback signal, denoted as $\hat{y}[t, \hat{\mathbf{f}}(t)]$, is obtained by filtering the loudspeaker signal $u(t)$ with a model $\hat{F}(q, t)$ of the acoustic feedback path; see Fig. 11. This model is calculated using an adaptive filter that is designed to identify the feedback path impulse response $\mathbf{f}(t)$ and track its changes. The feedback path and adaptive filter impulse responses are defined at time t as

$$\mathbf{f}(t) = \begin{bmatrix} f^{(0)}(t) & f^{(1)}(t) & \ldots & f^{(n_F)}(t) \end{bmatrix}^T \quad (55)$$

$$\hat{\mathbf{f}}(t) = \begin{bmatrix} \hat{f}^{(0)}(t) & \hat{f}^{(1)}(t) & \ldots & \hat{f}^{(n_{\hat{F}})}(t) \end{bmatrix}^T \quad (56)$$
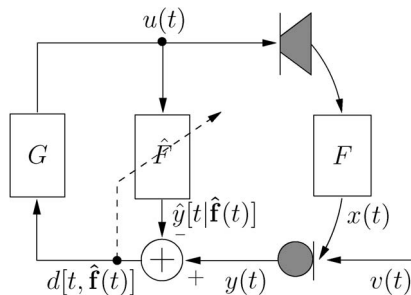
respectively.

Fig. 11. *AFC by predicting the feedback signal component $x(t)$ in the microphone signal, and hence subtracting the prediction $\hat{y}[t, \hat{f}(t)]$ from the microphone signal $y(t)$. The prediction is obtained by filtering the loudspeaker signal with a model $\hat{F}(q, t)$ of the acoustic feedback path, which is calculated using an adaptive filter.*

The closed-loop frequency response of the system shown in Fig. 11, employing an AFC method, is given by

$$\frac{U(\omega, t)}{V(\omega, t)} = \frac{G(\omega, t)}{1 - G(\omega, t)\big[F(\omega, t) - \hat{F}(\omega, t)\big]} \quad (57)$$

and, as a consequence, the Nyquist stability criterion can be rewritten as follows:

$$\begin{cases} \big|G(\omega, t)\big[F(\omega, t) - \hat{F}(\omega, t)\big]\big| \geq 1 & (58) \\ \angle G(\omega, t)\big[F(\omega, t) - \hat{F}(\omega, t)\big] = n2\pi, \quad n \in \mathbb{Z} & (59) \end{cases}$$

which leads to the following expression for the MSG [see also (16)]:

$$\mathrm{MSG}(t)[\mathrm{dB}] = -20 \log_{10}\left[\max_{\omega \in \mathcal{P}_{\hat{F}}} \big|J(\omega, t)\big[F(\omega, t) - \hat{F}(\omega, t)\big]\big|\right] \quad (60)$$

with

$$\mathcal{P}_{\hat{F}} = \big\{\omega \,\big|\, \angle G(\omega, t)\big[F(\omega, t) - \hat{F}(\omega, t)\big] = n2\pi\big\}.$$

From (60), it immediately follows that the better the fit between the estimated and actual feedback path frequency response, particularly at critical frequencies of the closed-loop system, the larger the achievable MSG increase. Theoretically, if $\hat{F}(q, t) \equiv F(q, t)$, the system would no longer exhibit a closed signal loop and hence the MSG would be infinitely large.

While the concept of AFC is relatively simple and similar to the well-known acoustic echo cancellation (AEC) approach, its realization is not straightforward. In the identification of the acoustic feedback path model $\hat{F}(q, t)$, a fundamental problem appears which is due to the closed-loop nature of the system. The LS estimate $\hat{\mathbf{f}}(t)$ of the acoustic feedback path impulse response $\mathbf{f}(t)$ can straightforwardly be calculated as

$$\hat{\mathbf{f}}(t) = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{y} \quad (61)$$

where the data vectors and matrices are defined as follows [and where the loudspeaker signal vector $\mathbf{u}(t)$ is not to be confused with the multichannel loudspeaker signal vector $\bar{\mathbf{u}}(t)$ defined in (5)]:

$$\mathbf{y} = \begin{bmatrix} y(t) & y(t-1) & \ldots & y(1) \end{bmatrix}^T \quad (62)$$

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}(t) & \mathbf{u}(t-1) & \ldots & \mathbf{u}(1) \end{bmatrix}^T \quad (63)$$

$$\mathbf{u}(t) = \begin{bmatrix} u(t) & u(t-1) & \ldots & u(t-n_{\hat{F}}) \end{bmatrix}^T. \quad (64)$$

The LS estimate may be characterized by its bias and variance [136, Ch. 8]. The bias corresponds to the difference between the expected value of the LS estimate and the true feedback path impulse response, i.e.,

$$\text{bias}\{\hat{\mathbf{f}}(t)\} = \begin{bmatrix} E\{\hat{\mathbf{f}}(t)\} \\ \mathbf{0}_{(n_F - n_{\hat{F}}) \times 1} \end{bmatrix} - \mathbf{f}(t) \qquad (65)$$

where $E\{\cdot\}$ denotes the expectation operator. Under a sufficient order assumption (i.e., $n_{\hat{F}} = n_F$), the expected value of the LS estimate can be shown to correspond to [81]

$$E\{\hat{\mathbf{f}}(t)\} = \mathbf{f}(t) + E\{(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{v}\}. \qquad (66)$$

The rightmost term in (66) can be understood to be generally nonzero due to the closed-loop nature of the system, which induces a correlation between the source signal and the loudspeaker signal, and hence

$$\text{bias}\{\hat{\mathbf{f}}(t)\} = E\{(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{v}\} \neq \mathbf{0}. \qquad (67)$$

The resulting effect in AFC is that the adaptive filter does not only predict and cancel the feedback component in the microphone signal, but also (part of) the source signal component. As a consequence, the feedback-compensated signal $d[t,\hat{\mathbf{f}}(t)]$ is a distorted estimate of the source signal $v(t)$. On the other hand, the variance of the LS estimate can be obtained by considering its covariance matrix,[8] which is calculated as [137]

$$\text{cov}\{\hat{\mathbf{f}}(t)\} = E\left\{\left(\hat{\mathbf{f}}(t) - \mathbf{f}(t)\right)\left(\hat{\mathbf{f}}(t) - \mathbf{f}(t)\right)^T\right\} \qquad (68)$$

$$= \left[E\{\mathbf{U}^T\mathbf{R}_\mathbf{v}^{-1}\mathbf{U}\}\right]^{-1} \qquad (69)$$

where the source signal covariance matrix $\mathbf{R}_\mathbf{v}$ is defined as

$$\mathbf{R}_\mathbf{v} = E\{\mathbf{v}\mathbf{v}^T\} \qquad (70)$$

---

[8]Note that the covariance matrix of the estimate $\hat{\mathbf{f}}(t)$ is in fact defined as $\text{cov}\{\hat{\mathbf{f}}(t)\} = E\{(\hat{\mathbf{f}}(t) - E\{\hat{\mathbf{f}}(t)\})(\hat{\mathbf{f}}(t) - E\{\hat{\mathbf{f}}(t)\})^T\}$, which corresponds to $\text{cov}\{\hat{\mathbf{f}}(t)\} = E\{(\hat{\mathbf{f}}(t) - \mathbf{f}(t))(\hat{\mathbf{f}}(t) - \mathbf{f}(t))^T\}$ if $E\{\hat{\mathbf{f}}(t)\} = \mathbf{f}(t)$, i.e., if the estimate is unbiased. However, in the analysis of closed-loop identification methods it has been found more meaningful to work directly with the covariance expression $\text{cov}\{\hat{\mathbf{f}}(t)\} = E\{(\hat{\mathbf{f}}(t) - \mathbf{f}(t))(\hat{\mathbf{f}}(t) - \mathbf{f}(t))^T\}$ even if $E\{\hat{\mathbf{f}}(t)\} \neq \mathbf{f}(t)$; see, e.g., [137].

with

$$\mathbf{v} = [v(t) \quad v(t-1) \quad \ldots \quad v(1)]^T. \qquad (71)$$

The interpretation of (69) can be related to the double-talk problem occurring in AEC [138]. In AEC, when the loudspeaker signal is active while the source signal is not, the covariance matrix of the acoustic echo path LS estimate is relatively small, since $\mathbf{R}_\mathbf{v} \approx \mathbf{0}$. However, when both signals are active at the same time (i.e., in a double-talk situation), the covariance matrix may become large, which may be observed in the adaptive filter performance as a decrease in convergence speed, or even a divergence. This problem becomes more severe as the source signal has a larger degree of coloration, since then the source signal covariance matrix $\mathbf{R}_\mathbf{v}$ exhibits a denser structure [138]. In AFC, the closed signal loop results in a continuous double-talk situation, and then this is made even worse by the correlation between the source and loudspeaker signal.

To prevent the adaptive filter from converging to a biased solution, and to increase its convergence speed despite the inevitable continuous double-talk situation, a decorrelation procedure is typically included in the AFC approach, with the aim of reducing the correlation between the source and loudspeaker signal. We can distinguish between two types of decorrelation [81], namely decorrelation in the closed signal loop and decorrelation in the adaptive filtering circuit. The former approach has the disadvantage of distorting the loudspeaker signal, while the latter approach requires somewhat more computations.

B. Realization

1) Adaptive Filtering: The adaptive calculation of the LS estimate (61) of the acoustic feedback path impulse response, and the subsequent calculation of the feedback-compensated signal can be performed as follows:

$$\varepsilon\left[t,\hat{\mathbf{f}}(t-1)\right] = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t-1) \qquad (72)$$

$$\mathbf{R}(t) = \lambda\mathbf{R}(t-1) + \mathbf{u}(t)\mathbf{u}^T(t) \qquad (73)$$

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mathbf{R}^{-1}(t)\mathbf{u}(t)\varepsilon\left[t,\hat{\mathbf{f}}(t-1)\right] \qquad (74)$$

$$d\left[t,\hat{\mathbf{f}}(t)\right] = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t). \qquad (75)$$

The algorithm in (72)–(74) is known as the recursive least squares (RLS) algorithm [139, Ch. 13], with an exponential forgetting factor λ. Note that the (*a priori*) RLS residual $\varepsilon[t,\hat{\mathbf{f}}(t-1)]$ in (72) differs from the (*a posteriori*) feedback-compensated signal $d[t,\hat{\mathbf{f}}(t)]$, which is subsequently processed in the electroacoustic forward path, in that the former depends on the previous estimate $\hat{\mathbf{f}}(t-1)$. Through the

application of the so-called matrix inversion lemma (MIL), the explicit inversion of the loudspeaker signal correlation matrix $\mathbf{R}(t)$ in (74) can be avoided, and hence the RLS algorithm requires $O(n_{\hat{F}}^2)$ multiplications per time update [139, Ch. 13]. Fast RLS algorithms, requiring only $O(n_{\hat{F}})$ multiplications per time update, have also been derived [140], [141] and successfully applied to, e.g., AEC [142]–[144, Ch. 6]. These fast RLS algorithms, however, are based on the shift invariance property of the loudspeaker signal vector defined in (64), i.e., the difference between $\mathbf{u}(t-1)$ and $\mathbf{u}(t)$ consists in discarding the oldest sample value $u(t-1-n_{\hat{F}})$, shifting the remaining sample values, and inserting the most recent sample value $u(t)$ in the first position. This shift invariance property will generally not hold in the AFC context due to the decorrelation that is applied, either in the closed signal loop or in the adaptive filtering circuit, which often involves a nonlinear or time-varying filtering of the loudspeaker signal (see Section VI-B2).

Alternatively, a computationally cheaper adaptive algorithm can be found in the underdetermined recursive least squares (URLS) family [145]. The affine projection algorithm (APA) provides an estimate of the acoustic feedback path impulse response by using only the M most recent microphone signal samples and loudspeaker signal vectors, i.e.,

$$\varepsilon_M\left[t, \hat{\mathbf{f}}(t-1)\right] = \mathbf{y}_M(t) - \mathbf{U}_M^T(t)\hat{\mathbf{f}}(t-1) \tag{76}$$

$$\begin{aligned}\hat{\mathbf{f}}(t) = &\hat{\mathbf{f}}(t-1) \\ &+ \mu\mathbf{U}_M(t)\left[\mathbf{U}_M^T(t)\mathbf{U}_M(t) + \alpha\mathbf{I}_M\right]^{-1} \\ &\times \varepsilon_M\left[t, \hat{\mathbf{f}}(t-1)\right]\end{aligned} \tag{77}$$

$$d\left[t, \hat{\mathbf{f}}(t)\right] = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t) \tag{78}$$

where μ represents the step size, $\alpha\mathbf{I}_M$ is an $M \times M$ identity regularization matrix scaled with the regularization parameter α, and

$$\mathbf{y}_M(t) = [y(t) \quad \ldots \quad y(t-M+1)]^T \tag{79}$$

$$\mathbf{U}_M(t) = [\mathbf{u}(t) \quad \ldots \quad \mathbf{u}(t-M+1)]. \tag{80}$$

The APA requires $O(Mn_{\hat{F}})$ multiplications per time update, which is obviously much less than the RLS algorithm if the APA projection order $M \ll n_{\hat{F}}$. Since the APA can be shown to decorrelate input signals that admit an Mth-order all-pole representation [146], relatively small values for the projection order have been successfully used in AEC applications with speech signals (e.g., $M \in \{2, \ldots, 50\}$ [147], $M = 8$ [148], $M = 10$ [149]). Fast versions of the APA, requiring only $O(n_{\hat{F}}) + O(M)$ multiplications per time update, have also been derived [147]–[149], but again these algorithms

rely on the shift invariance property of the loudspeaker signal vector, which need not be fulfilled in the AFC context.

Since the required adaptive filter order in room acoustic applications may be very high, a computational load of $O(Mn_{\hat{F}})$ multiplications per time update may still be too heavy to allow for a real-time application. For this reason, in most of the existing AFC schemes, a projection order $M = 1$ is applied, which leads to the well-known normalized least mean squares (NLMS) algorithm

$$\varepsilon\left[t, \hat{\mathbf{f}}(t-1)\right] = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t-1) \tag{81}$$

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mu\frac{\mathbf{u}(t)\varepsilon(t)}{\mathbf{u}^T(t)\mathbf{u}(t) + \alpha} \tag{82}$$

$$d\left[t, \hat{\mathbf{f}}(t)\right] = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t). \tag{83}$$

The required number of multiplications per time update is $O(n_{\hat{F}})$, more specifically $4n_{\hat{F}} + 6$ [if the calculation of the feedback-compensated signal in (83) is also taken into account]. The choice of the NLMS step size μ is crucial to obtain a good compromise between a stable and fast convergence. In our own experience, step size values in the range $\mu = 0.01$–0.05 are best suited for speech applications [101], [103], [105], while values around $\mu = 0.005$ are recommended when working with audio signals [111].

Finally, the choice of the adaptive filter order $n_{\hat{F}}$ is obviously extremely important, regardless of which adaptive filtering algorithm is used. It is clear that the choice of $n_{\hat{F}}$ has a profound influence on the computational requirements of the AFC approach. One could argue that it may be sufficient to choose $n_{\hat{F}}$ such that the largest components in the acoustic feedback path impulse response (originating from the early reflections) can be modeled. Unfortunately, such an approach would be inefficient for two reasons: first, large impulse response components do not necessarily correspond to large frequency response components and hence stability may not be improved by only cancelling the early reflections. Second, if the impulse response is undermodeled (i.e., $n_{\hat{F}} < n_F$) then an additional bias component will appear in the LS estimate (in addition to the bias due to the source and loudspeaker signal correlation) and moreover its variance will increase [150]. The best compromise between computational complexity and feedback control performance probably consists in choosing $n_{\hat{F}}$ just large enough to obtain a satisfying MSG increase, and applying a technique for reducing the bias and variance due to undermodeling [150]–[152]. We should point out that the technique proposed by Rombouts et al. [150] for consistently identifying undermodeled RIRs is particularly interesting in the context of AFC, since it additionally provides a decorrelation in the adaptive filtering circuit.

We should emphasize that the above adaptive algorithms are often not implemented as such, since both the robustness

and the efficiency of these algorithms can be further improved [103]. A robust adaptive filter implementation for AFC may include the following features: an adaptation control that freezes the adaptive filter coefficients during source signal onsets [103], a foreground/background adaptive filter implementation to combine good tracking properties with a small steady-state error [103], and a regularization method that compensates for the coloration of the loudspeaker signal [103], [105]. Moreover, the AFC efficiency in terms of computational load and convergence speed can be improved by considering a subband or frequency domain adaptive filter implementation rather than the time domain implementations shown here [103].

2) Decorrelation: Decorrelation of the source and loud-speaker signals in the closed signal loop can be achieved by inserting a decorrelating signal operation in the electroacoustic forward path; see Fig. 12. We can distinguish between the following approaches.

- Noise injection [64], [74], [83]–[86]: a white noise signal $n(t)$ is added to the feedback-compensated signal after the electroacoustic forward path processing (but usually before the forward path amplification) [see Fig. 12(a)], i.e.,

$$u(t) = K(t)\Big[J(q,t)d\big[t,\hat{\mathbf{f}}(t)\big] + n(t)\Big]. \qquad (84)$$

The acoustic feedback path identification can then proceed in two ways: if the loudspeaker signal $u(t)$ [including the noise signal $n(t)$] is used as the input signal to the adaptive filter, then the effect of the noise injection is that the source and loudspeaker signal cross-correlation is decreased, hence the bias will be reduced but not completely eliminated. A second possibility is to use the noise signal $n(t)$ as the adaptive filter input signal, which leads to an

unbiased estimate since the source signal and the noise signal are uncorrelated. In this case, however, the adaptive filter convergence will be rather slow since not only the source signal, but also the source signal component in the feedback signal acts as a disturbance towards the feedback path identification. With the aim of reducing the influence of the noise injection on sound quality, the noise spectrum can be shaped such as to render the noise less perceptible, e.g., by A-weighting [84] or psychoacoustic noise shaping [85], [86]. Unfortunately, noise shaping decreases the decorrelation effect, making the noise injection less effective in removing the bias.

- Time-varying processing [74], [87]–[89]: each of the LPTV filters described in Section IV can be used as a decorrelating filter in the electroacoustic forward path [see Fig. 12(b)]

$$u(t) = G(q,t)\Big[H(q,t)d\big[t,\hat{\mathbf{f}}(t)\big]\Big]. \qquad (85)$$

An FS decorrelating filter has an LPTV frequency response as given by (26) and has successfully been applied to AFC for speech applications [87]–[89]. While the perceptible signal distortion introduced by the FS operation appears to be acceptable for speech signals [89], the FS decorrelation technique was found to be perceptually inadequate for audio applications [86]. The sinusoidal PM and DM filters, of which the LPTV frequency response is given in (24) and (29), respectively, have also been applied as decorrelation filters in AFC [74], [87], [88]. Note that a beneficial side effect of using LPTV decorrelation filters is that these filters also stabilize the closed-loop system by smoothing the loop gain (see Section IV).

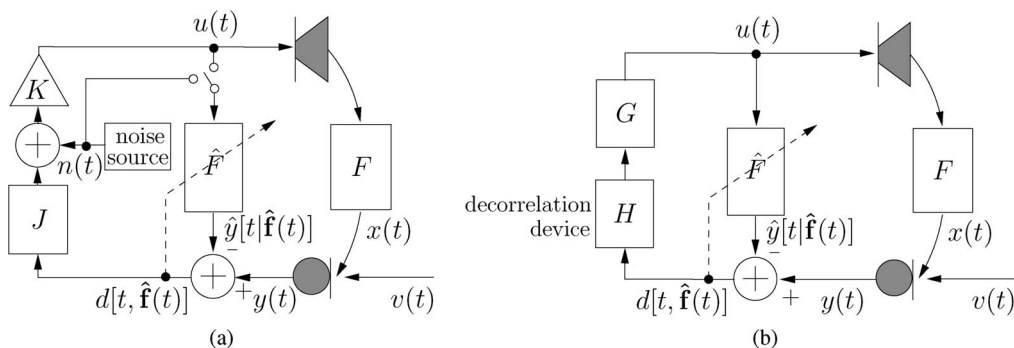**Fig. 12.** *AFC with decorrelation in the closed signal loop. (a) Decorrelation by noise injection: the adaptive filter input signal can be either the loudspeaker signal* $u(t)$ *or the noise signal* $n(t)$. *(b) Decorrelation in the electroacoustic forward path: the decorrelation device corresponds to an LPTV filter* $H(q,t)$, *a nonlinear mapping* $H\{\cdot, t\}$, *or a processing delay* q^{-d_1}.

- Nonlinear processing [74], [91]: in the context of stereo AEC, the correlation between the stereo channels, which leads to an identifiability problem of the acoustic echo path models [153], has been reduced by applying nonlinear decorrelating operations to the loudspeaker signals [90]. These nonlinear operations can also be used to reduce the correlation between the source and loudspeaker signals in an AFC application. In particular, half-wave rectification has been successfully applied to AFC decorrelation [74], [91] [see Fig. 12(b)], i.e.,

$$u(t) = G(q,t)\Big[H\Big(d\big[t,\hat{\mathbf{f}}(t)\big],t\Big)\Big] \qquad (86)$$

$$= G(q,t)\Bigg[d\big[t,\hat{\mathbf{f}}(t)\big]$$
$$+ \alpha\Bigg(\frac{d\big[t,\hat{\mathbf{f}}(t)\big] + \big|d\big[t,\hat{\mathbf{f}}(t)\big]\big|}{2}\Bigg)\Bigg]. \qquad (87)$$

The parameter α can be tuned to obtain the best compromise between decorrelation and perceptible signal distortion.

- Forward path delay: in HA AFC applications, inserting a processing delay of d_1 samples in the electroacoustic forward path has been proposed to decorrelate the source and loudspeaker signals [79], [92] [see Fig. 12(b)], i.e.,

$$u(t) = G(q,t)d\Big[t - d_1, \hat{\mathbf{f}}(t - d_1)\Big]. \qquad (88)$$

This approach is particularly useful for source signals that have an autocorrelation function that decays rapidly, e.g., voiceless speech signals, provided that the delay value d_1 is chosen accordingly.

Note that when applying decorrelation in the closed signal loop, a tradeoff between bias reduction and sound quality should always be sought by properly tuning the decorrelation parameters. Usually, a perceptible signal distortion is unavoidable, either because of the decorrelating signal operation itself (when strong decorrelation is applied), or because of the bias in the acoustic feedback path estimate (when weak decorrelation is applied) [81].

Decorrelation in the adaptive filtering circuit does not require the above tradeoff and generally, the stronger the decorrelation, the better will be the attained sound quality. Two such approaches have been proposed.

- Adaptive filter delay [93], [94]: due to the time needed for the loudspeaker sound to propagate

through a direct coupling to the microphone, the acoustic feedback path impulse response typically exhibits an initial delay (sometimes referred to as the "dead time"; see Fig. 3), the value of which is proportional to the loudspeaker–microphone distance. If this initial delay (or a lower bound for it) is known a priori and corresponds to $d_2 T_s$ s, then the first d_2 coefficients in the acoustic feedback path model can be forced to zero, i.e.,

$$\hat{F}(q,t) = \hat{f}^{(d_2)}(t)q^{-d_2}$$
$$+ \hat{f}^{(d_2+1)}(t)q^{-(d_2+1)} + \ldots + \hat{f}^{(n_{\hat{F}})}q^{-n_{\hat{F}}}. \qquad (89)$$

As a consequence, the first d_2 rows in the expression (61) for the LS estimate of the acoustic feedback path impulse response need not be considered, and likewise for the bias vector in (67). If we now assume that the source and loudspeaker signal cross-correlation function is small for time lags larger than d_2 samples, then the remaining bias can be considered negligible.

- Decorrelating prefilters [81], [91], [97]: from a system identification point of view, the bias in the LS estimate of the acoustic feedback path model can be eliminated by using an appropriate noise model in the identification [137], i.e., a model of the signal that disturbs the identification, more specifically the source signal in the AFC context. If we assume a (time-varying) parametric source signal model $H(q,t)$

$$v(t) = H(q,t)e(t) \qquad (90)$$

and that an estimate $\hat{H}(q,t)$ of $H(q,t)$ is available, then the unbiased identification approach consists in prefiltering the loudspeaker and microphone signals with the inverse source signal model estimate before feeding these signals to the adaptive filtering algorithm. Note that the source signal excitation signal $e(t)$ in (90) is assumed to be an uncorrelated signal (i.e., white noise or a Dirac impulse). This approach is depicted in Fig. 13(a), where the prefiltered loudspeaker and microphone signals are calculated as

$$\tilde{y}\Big[t, \hat{\mathbf{h}}(t)\Big] = \hat{H}^{-1}(q,t)y(t) \qquad (91)$$

$$\tilde{u}\Big[t, \hat{\mathbf{h}}(t)\Big] = \hat{H}^{-1}(q,t)u(t) \qquad (92)$$

and $\hat{\mathbf{h}}(t)$ contains the estimated source signal model parameters. This approach was originally developed for HA AFC applications [80], [95], [96] and later on extended to room acoustic applications [81], [91], [97].
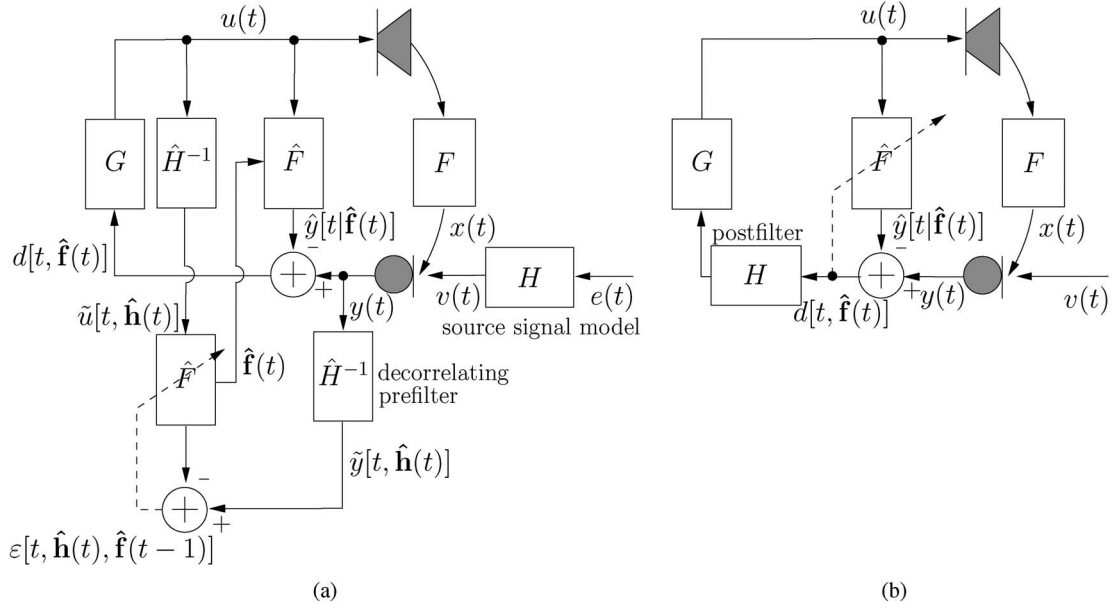
Fig. 13. *(a) AFC with decorrelating prefilters in the adaptive filtering circuit: a linear parametric source signal model $H(q, t)$ is estimated, and subsequently the microphone and loudspeaker signals are prefiltered with the inverse source signal model before being fed to the adaptive filter. (b) AFC with postfiltering: the postfilter $H(q, t)$ can either be a spectral subtraction filter for residual feedback suppression, or a bank of notch filters to avoid closed-loop instability.*

Both approaches to decorrelation in the adaptive filtering circuit rely on additional information that is not necessarily available *a priori* and may moreover be time varying, i.e., the initial delay of the acoustic feedback path and the source signal model. The problem of how to concurrently estimate the initial delay and the model coefficients of the acoustic feedback path impulse response has not yet been treated in the literature. On the other hand, the concurrent estimation of the source signal model and the acoustic feedback path model has been studied extensively by Rombouts *et al.* [100]–[102] for speech applications and by van Waterschoot and Moonen [110], [111] for audio applications. For speech source signals, the parametric source signal model preferably consists of a cascade of two all-pole models [100], [101]

$$H(q, t) = \frac{1}{A(q, t)} \frac{1}{C(q, t)} \qquad (93)$$

with

$$A(q, t) = 1 - \sum_{i=-1}^{1} \alpha^{(i)}(t) q^{-K-(l/D)-i} \qquad (94)$$

$$C(q, t) = 1 + \sum_{i=1}^{n_C} c^{(i)}(t) q^{-i}. \qquad (95)$$

The three-tap fractional pitch prediction model $1/A(q, t)$ is used to model the periodic speech components that stem from the vibration of the vocal chords. Here, $K + l/D$ represents the fractional pitch lag, with K the integer pitch lag, D the interpolation ratio, and $l \in \{0, \dots, D-1\}$ the fractional phase [154]. The all-pole model $1/C(q, t)$ represents the vocal tract response that produces the formant speech components [155]. The cascade model in (93)–(95) can also be used for monophonic audio signals, while for polyphonic audio signals a cascade of a constrained pole-zero model with an all-pole model appears to be better suited [111], [156], i.e.,

$$H(q, t) = \frac{B(q, t)}{A(q, t)} \frac{1}{C(q, t)} \qquad (96)$$

with

$$\frac{A(q, t)}{B(q, t)} = \prod_{i=1}^{n_A/2} \frac{1 - 2\nu_i \cos \theta_i q^{-1} + \nu_i^2 q^{-2}}{1 - 2\rho_i \cos \theta_i q^{-1} + \rho_i^2 q^{-2}}. \qquad (97)$$

The constrained pole-zero model $B(q, t)/A(q, t)$ then models the tonal components in the audio signal, while the all-pole model $1/C(q, t)$ models the "noise-like" components. The constrained pole-zero model is usually parametrized using a

second-order sections structure, as shown in (97), where the θ_i correspond to the pole-zero angles, and ν_i and ρ_i are the zero and pole radii, $i = 1, \ldots, n_A/2$.

The concurrent estimation of the source signal models and the acoustic feedback path model can be performed using a prediction error identification approach [98, Ch. 3], [99, Ch. 7], which then leads to the so-called PEM-based AFC algorithms proposed in [100]–[103], [110], and [111].

3) Postfiltering: Mainly owing to undermodeling and steady-state as well as tracking errors, a misadjustment between the AFC adaptive filter coefficients and the acoustic feedback path impulse response will unavoidably exist. As a result, the feedback signal $x(t)$ will typically not be completely canceled from the microphone signal, and so the feedback-compensated signal contains a residual feedback signal component $r[t, \hat{\mathbf{f}}(t)]$

$$d\left[t, \hat{\mathbf{f}}(t)\right] = v(t) + \underbrace{\left[F(q,t) - \hat{F}(q,t)\right]u(t)}_{\triangleq r[t, \hat{\mathbf{f}}(t)]}. \tag{98}$$

A similar problem was previously encountered in AEC, and residual echo suppression postfilters have successfully been applied in this area [157]–[159]. These postfilters operate on the echo-compensated signal and attempt to suppress the residual echo component using a spectral subtraction approach. Several attempts have been made to apply the AEC postfiltering approach to the AFC scenario [66], [94], resulting in the AFC scheme shown in Fig. 13(b). We should emphasize that, again, the correlation between the loudspeaker and source signal makes the residual feedback suppression problem much harder in the AFC case as compared to the AEC case. Since the postfiltering approach is based on spectral subtraction, the postfilter is usually designed directly in the frequency domain.

Janse and Belt [66] propose the procedure, shown in (99)–(100) at the bottom of the page, to determine the postfilter magnitude response, where $|Y(\omega_k, t)|$, $|\hat{Y}[\omega_k, t, \hat{\mathbf{f}}(t)]|$, and $|D[\omega_k, t, \hat{\mathbf{f}}(t)]|$ denote the short-term DFT magnitude spectra of the microphone signal, the feedback signal estimate, and the feedback-compensated signal, respectively, which are defined similarly to (40). Ideally, the filter in (99) should behave as follows: when the source signal component dominates in the short-term magnitude spectrum of the microphone signal, the amount of spectral subtraction should be small, while if the

feedback signal component dominates, the amount of subtraction should be large [66]. The subtraction factor γ is chosen larger than one in case the estimated maximum loop gain $\max_\omega |G(\omega, t)\hat{F}(\omega, t)| \geq 1$, while $\gamma < 1$ if $\max_\omega |G(\omega, t)\hat{F}(\omega, t)| < 1$. The first-order low-pass filtering operation in (100) is performed to obtain a smoothly time-varying postfilter behavior. Unfortunately, the postfilter response in (99) also depends on an estimate of the short-term residual feedback signal spectrum $|\hat{R}[\omega_k, t, \hat{\mathbf{f}}(t)]|$, yet no details are provided in [66] on how to obtain this estimate.

An alternative postfilter design procedure for residual feedback suppression was proposed by Ortega *et al.* [94], which is based on the observation that an optimal expression for the postfilter [in the sense of forcing the closed-loop frequency response in (10) to be exactly equal to the electroacoustic forward path response $G(\omega, t)$] is given by

$$H(\omega, t) = \frac{1}{1 + G(\omega, t)\left[F(\omega, t) - \hat{F}(\omega, t)\right]} \tag{101}$$

$$= 1 - \sqrt{\frac{S_r\left[\omega, t, \hat{\mathbf{f}}(t)\right]}{S_d\left[\omega, t, \hat{\mathbf{f}}(t)\right]}} \tag{102}$$

where $S_r[\omega, t, \hat{\mathbf{f}}(t)]$ and $S_d[\omega, t, \hat{\mathbf{f}}(t)]$ denote the short-term power spectral density (PSD) of the residual feedback signal and the feedback-compensated signal, respectively. Here, $S_d[\omega, t, \hat{\mathbf{f}}(t)]$ is estimated from the feedback-compensated signal $d[t, \hat{\mathbf{f}}(t)]$ using the periodogram followed by a Mel-scale-based frequency smoothing. Finally, $S_r[\omega, t, \hat{\mathbf{f}}(t)]$ is estimated recursively

$$\hat{S}_r\left[\omega, t, \hat{\mathbf{f}}(t-1)\right]$$
$$= [\delta + 2\lambda(1-\lambda)]\hat{S}_r\left[\omega, t-1, \hat{\mathbf{f}}(t-2)\right]$$
$$+ (1-\delta)\lambda^2 \hat{S}_d\left[\omega, t-1, \hat{\mathbf{f}}(t-1)\right]$$
$$+ (1-\delta)(1-\lambda)^2 \frac{\hat{S}_r^2\left[\omega, t-1, \hat{\mathbf{f}}(t-2)\right]}{\hat{S}_d\left[\omega, t-1, \hat{\mathbf{f}}(t-1)\right]} \tag{103}$$

where the parameters λ and δ are chosen to be around 0.3 and 0.8, respectively [94]. Consequently, the first term on the right-hand side of (103) dominates the other terms, and hence it can be understood that the initialization of the

$$\left|\tilde{H}(\omega_k, t)\right| = \max\left\{\frac{|Y(\omega_k, t)| - \gamma\left(\left|\hat{Y}\left[\omega_k, t, \hat{\mathbf{f}}(t)\right]\right| + \left|\hat{R}\left[\omega_k, t, \hat{\mathbf{f}}(t)\right]\right|\right)}{\left|D\left[\omega_k, t, \hat{\mathbf{f}}(t)\right]\right|}, 0\right\} \tag{99}$$

$$\left|H(\omega_k, t)\right| = \lambda|H(\omega_k, t-1)| + (1-\lambda)\left|\tilde{H}(\omega_k, t)\right| \tag{100}$$

residual feedback signal PSD estimate at $t = 0$ has a crucial effect on the quality of the estimate in (103).

It should be noted that a postfilter may also be used in the AFC scheme with the aim of preventing closed-loop system instability rather than suppressing the residual feedback signal. In this case, the postfilter should behave as a bank of notch filters, operating at the critical frequencies of the closed-loop system. Schmidt *et al.* [74], [106] propose an ANF postfilter that does not directly use any information from the AFC adaptive filter, and hence does not behave differently from the ANF that operates without an AFC (see Section III-B). Rombouts *et al.* [52], [103] propose a postfilter based on a two-stage NHS method, in which the NHS howling detection is replaced by a proactive detection of critical frequencies by inspecting the estimated loop gain $|G(\omega,t)\hat{F}(\omega,t)|$ using the most recent AFC acoustic feedback path estimate $\hat{F}(q,t)$.

C. Initialization

Similarly to the NHS method, an initialization procedure that is performed during the startup of the sound reinforcement system is useful to improve the performance of the AFC method. The room acoustics information that is gathered during the initialization can be elegantly incorporated in the AFC adaptive filtering algorithm using a technique known as regularization [105], [160]. The most straightforward approach to regularization consists in calculating an offline estimate of the acoustic feedback path impulse response, and subsequently using this estimate as the initial parameter vector $\hat{\mathbf{f}}(0)$ in any of the adaptive algorithms discussed in Section VI-B1. While this approach may lead to a considerable improvement of the adaptive filter's convergence speed, it is nonrobust to changes in the acoustic feedback path impulse response. More particularly, the impulse response may be considerably different during initialization and during operation of the system, e.g., due to the presence of an audience on the room acoustics.

A more advanced approach to regularization consists in identifying the acoustic feedback path model in a Bayesian minimum mean square error (MMSE) framework instead of in an LS framework [105]. The acoustic feedback path impulse response $\mathbf{f}(t)$ is then viewed as a stochastic quantity on which some prior knowledge may be available, e.g., the mean $E\{\mathbf{f}(t)\} = \mathbf{f}_0$ and covariance matrix $\text{cov}\{\mathbf{f}(t)\} = \mathbf{R_f}$. In the Bayesian MMSE framework, the optimal impulse response estimate is then given by [105]

$$\hat{\mathbf{f}}(t) = \mathbf{f}_0 + \left(\mathbf{U}^T\mathbf{R_v}^{-1}\mathbf{U} + \mathbf{R_f}^{-1}\right)^{-1}\mathbf{U}^T\mathbf{R_v}^{-1}(\mathbf{y} - \mathbf{U}\mathbf{f}_0) \quad (104)$$

which, in contrast to the LS estimate in (61), depends both on the acoustic feedback path statistics through $\mathbf{f}_0$ and $\mathbf{R_f}$, and on the source signal statistics through $\mathbf{R_v}$ defined in (70). In the context of adaptive filtering, the mean of the acoustic feedback path impulse response is usually chosen either as

$\mathbf{f}_0 = \mathbf{0}$ or as $\mathbf{f}_0 = \hat{\mathbf{f}}(t-1)$, which results in two well-known types of regularization, more specifically, Tikhonov regularization (TR) and Levenberg–Marquardt regularization (LMR), respectively [105]. On the other hand, the covariance matrix $\mathbf{R_f}$ is constructed using an initial impulse response measurement or using the available room acoustic parameters such as the reverberation time and the loudspeaker–microphone distance [105]. The resulting adaptive filtering algorithms, known as TR-RLS, LMR-RLS, LMR-APA, and LMR-NLMS, do not require significantly more computations as compared to the original RLS, APA, and NLMS algorithms, if the covariance matrix $\mathbf{R_f}$ is constructed to be a diagonal matrix [105], [160].

D. Discussion

The AFC approach is widely considered to be the most promising solution to the acoustic feedback problem. Its most attractive property lies in the fact that the effect of acoustic feedback can be completely canceled, provided that the AFC algorithm converges to the desired solution, and hence the MSG can be increased considerably. Experiments have shown that MSG increases of 15–20 dB are practically achievable [89], [103], which is two to three times more than the MSG increases obtained with the PFC and NHS approaches (see Sections IV-C and V-D). As a consequence, a sound reinforcement system equipped with an AFC method can generally operate at a reasonably large gain margin and hence howling, ringing, and reverberation artifacts can be avoided, resulting in a high sound quality. We should note, however, that in terms of sound quality, the choice of the decorrelation method is of crucial importance. In particular, when applying decorrelation in the closed signal loop, signal distortion appears to be unavoidable, either because the decorrelation itself is perceptible, or because the source signal is partially canceled when the decorrelation is insufficient [81]. From this point of view, it is highly desirable to perform the decorrelation in the adaptive filtering circuit instead of in the closed signal loop. In terms of robustness, the AFC approach has benefited much from recent improvements such as postfiltering [66], [93], [94], notch filtering [74], [103], [106]–[108], adaptation control [103], and regularization [103], [105], [160].

The main disadvantage of the AFC approach is its computational complexity, which is typically much higher than the PFC and NHS complexity. Even when the cheapest adaptive filtering algorithm is applied, i.e., the NLMS algorithm which requires $O(n_{\hat{f}})$ multiplications per time update, the AFC complexity may still exclude a real-time implementation. The reason for this is twofold. First, since the acoustic feedback path is modeled by its impulse response, a very high adaptive filter order is typically required. Second, since a sufficiently high sampling rate should be used to obtain a good sound quality (especially for audio applications), the impulse response is densely sampled hence requiring many coefficients, and moreover, a large number

of adaptive filter iterations has to be performed per second. Nevertheless, several real-time AFC implementations for single-channel systems have been reported. Goertz has tested a real-time AFC setup with a 2646-tap adaptive filter (i.e., modeling the first 60 ms of the acoustic feedback path impulse response at $f_s = 44.1$ kHz) in a room with $T_{60} = 1.2$ s, thereby achieving a 5-dB MSG increase [84]. Rombouts *et al.* have reported MSG increases up to 14 dB in a real-time AFC experiment with a frequency domain adaptive filter of order 2048, operating at a sampling frequency of 16 kHz in a room with $T_{60} = 120$ ms [103].

The high complexity also puts a limit on the generalization of the AFC approach to multichannel systems. Since no results are available on how to exploit the fact that the different acoustic feedback path impulse responses of a multichannel system share some underlying room acoustic properties, the state of the art in multichannel AFC consists in applying $S \cdot L$ single-channel AFC algorithms in a system having S microphones and L loudspeakers, hence the resulting complexity also increases with a factor $S \cdot L$.

VII. EVALUATION

A. Evaluated Algorithms

From the above exposition, it is clear that a multitude of acoustic feedback control methods has been proposed. An experimental evaluation of all the existing methods and realizations is beyond the scope of this paper. We will however provide an evaluation of a selection of methods and realizations that we consider representative for the state of the art. The evaluation is based on computer simulations rather than real-time experiments, to make sure the simulation scenario is exactly reproducible for the different algorithms. From each of the three presented categories of feedback control methods (i.e., PFC, NHS, and AFC methods), we will select three different state-of-the-art algorithms.

As for PFC, we evaluate three of the PM techniques described in Section IV-A: sinusoidal PM, FS, and sinusoidal DM. The corresponding PFC algorithms are denoted as PFC-PM, PFC-FS, and PFC-DM, respectively. The PFC-PM and PFC-FS algorithms are realized as shown in (33) and (35), respectively, where the discrete-time Hilbert transform $\hat{y}(t)$ is estimated using the method proposed in [115]. For the PFC-DM algorithm, we use a Hamming-windowed and truncated linear interpolation filter as given in (38), with an interpolation ratio $D = 8$ and a filter length of $2I = 32$ taps. The PFC parameters are tuned to provide a firm tradeoff between the resulting MSG and signal distortion, and also taking into account the parameter values suggested in the PFC literature. In the PFC-FS algorithm, following [2], the modulation frequency is set to $f_m = 5$ Hz. In the PFC-PM and PFC-DM algorithms, however, a lower value should be used to avoid excessive signal distortion, hence for these two approaches we set $f_m = 1$ Hz. In the PFC-PM algorithm, a modulation

index $\beta = 3.8$ was found to produce better results than $\beta = 2.4$, while the PFC-DM algorithm is implemented with a modulation depth of $\Delta_\tau = 32$ samples and a delay offset of $\tau_0 = \Delta_\tau + 2I = 64$ samples.

In the NHS approach, many different howling detection criteria can be designed by combining the spectral and temporal microphone signal features defined in Section V-B1. An elaborate evaluation of each of these features, both in terms of howling detection accuracy and NHS feedback control performance, can be found in [62] and [63]. Here, we will only consider the following three approaches. In the first algorithm (denoted as NHS-1), the howling detection is performed as suggested in [32] and [34], using a combination of the PHPR and IPMP features defined in (45) and (50), respectively. Howling is then detected if for a certain frequency, both (46) is fulfilled (with $T_{PHPR} = 30$ dB) and IPMP $\geq 3/5$ with $Q_M = 5$. The second algorithm (denoted as NHS-2) uses the PAPR feature (42) for howling detection, following, e.g., [50] and [51]. It was found that the PAPR threshold should preferably have a different value in speech and audio applications [62], [63], e.g., $T_{PAPR}^{(speech)} = 33$ dB and $T_{PAPR}^{(audio)} = 55$ dB. Choosing $T_{PAPR}^{(audio)} > T_{PAPR}^{(speech)}$ is recommended since the tonal components in an audio signal are much more easily misclassified as howling components. Finally, in the third algorithm (denoted as NHS-3), we apply the howling detection criterion proposed in [28]–[30], which combines the PNPR and IMSD features, defined in (47) and (51), respectively. According to [28]–[30], the PNPR and IMSD features are used to calculate two secondary features, namely the "peakness" and the "slopeness," which are subsequently combined into a so-called feedback existence probability (FEP) function as follows:

$$\text{FEP}(\breve{\omega}_i, t) = 0.7 \cdot \text{slopeness}(\breve{\omega}_i, t) + 0.3 \cdot \text{peakness}(\breve{\omega}_i, t). \tag{105}$$

The relation between the PNPR and peakness features is given in (49), while the calculation of the slopeness from the IMSD is performed using a nonlinear mapping (which is not explicitly given in [28]–[30]) that is chosen to be

$$\text{slopeness}(\breve{\omega}_i, t) = e^{-|\text{IMSD}(\breve{\omega}_i, t)|}. \tag{106}$$

Again, we found that a different value of the FEP threshold should be used in speech and audio applications, e.g., $T_{FEP}^{(speech)} = 0.7$ (as suggested in [28]–[30]) and $T_{FEP}^{(audio)} = 0.95$ [with howling being detected if $\text{FEP}(\breve{\omega}_i, t) \geq T_{FEP}$]. Since the howling detection in the NHS-1 and NHS-3 algorithms is more advanced as compared to the NHS-2 algorithm, we can expect a larger false alarm probability when using the latter algorithm [50], [51], [62], [63]. To

compensate for this effect, the notch filters applied in the NHS-2 algorithm are given a very small bandwidth, i.e., 1/60 octave, as suggested in [50] and [51]. The NHS-1 and NHS-3 algorithms work with 1/10 octave notch filters, following [32], [34], [52], and [103]. Also, the maximum number of cascaded notch filters as defined in (53) is set to $n_H/2 = 12$ in the NHS-1 and NHS-3 algorithms, and to $n_H/2 = 48$ in the NHS-2 algorithm. Finally, we should mention that all three NHS algorithms under consideration apply a DFT-based frequency analysis as in (40), with $M = 2048$ at $f_s = 16$ kHz, or $M = 4096$ at $f_s = 44.1$ kHz, and with $P = M/2$, from which $N = 3$ candidate howling components are identified by peak picking.

The AFC approach will be evaluated using three different decorrelation methods; see Section VI-B2. We refer to [82] for an evaluation of the decorrelation methods that are not covered here. The first AFC algorithm (denoted as AFC-NI) includes a decorrelation by noise injection, in which a white noise signal is added to the feedback-compensated signal before amplification, as suggested in [64], [74], and [83]–[85]. The loudspeaker signal is chosen as the input signal to the adaptive filter [i.e., the switch in Fig. 12(a) is set to its vertical position]. The injected noise power is adjusted to be 10 dB below the long-term feedback-compensated signal power, which results in an adaptive filter convergence speed that is comparable to the other AFC algorithms under consideration. The second algorithm (denoted as AFC-FS) features a decorrelation by a time-varying processing, more specifically by FS, following [74] and [87]–[89]. The FS operation is realized as in the PFC-FS algorithm, and the same modulation frequency $f_m = 5$ Hz will be used. The third algorithm (denoted as AFC-PF) is based on decorrelating prefilters, as proposed in [81], [91], and

[97]. We will use a cascade source signal model that consists of a pitch prediction model and an all-pole model, as defined in (93)–(95), which are estimated concurrently with the acoustic feedback path model using the PEM-AFROW algorithm [100], [101], [103]. For this algorithm to be applicable for both speech and audio source signals, the search range for the pitch lag K should be chosen large enough, e.g., $K \in \{\lceil f_s/1000 \rceil, \ldots, \lceil f_s/100 \rceil\}$ [111]. The fractional delay in the pitch prediction model (94) is approximated by a linear interpolation filter similar to the interpolation filter in the PFC-DM algorithm. The all-pole model order is set to $n_C = 20$, and both source signal models are estimated using 50% overlapping data windows of length $M = 320$ at $f_s = 16$ kHz in case of speech source signals [100], [101], and of length $M = 2048$ at $f_s = 44.1$ kHz for audio source signals [111]. Moreover, a processing delay of half the data window length is inserted in the electroacoustic forward path, as suggested in [100], [101], and [111]. In all three AFC algorithms, the acoustic feedback path model order is equal to the length of the feedback path impulse response, i.e., $n_{\hat{F}} = n_F$, and the NLMS algorithm (81)–(83) is used to update the adaptive filter coefficients. The NLMS step size is chosen to be $\mu = 0.02$ for speech source signals and $\mu = 0.005$ for audio source signals, while the regularization parameter is set to $\alpha = 10^{-6}$.

B. Evaluation Procedure

We will evaluate the performance of each of the nine algorithms described above in two simulation scenarios: a 30-s simulation at $f_s = 16$ kHz with a speech source signal, and a 60-s simulation at $f_s = 44.1$ kHz with an audio source signal. The speech signal is plotted in Fig. 14(a) and is taken from an interview with two male Dutch-speaking subjects that was digitally broadcast by the Flemish Radio
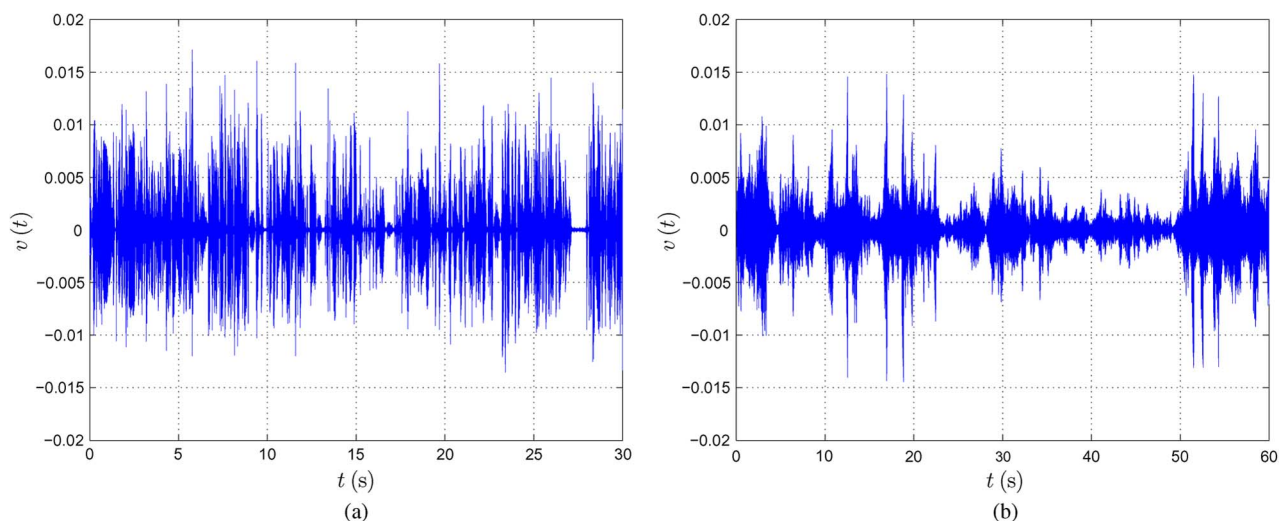

Fig. 14. *Source signals used in the evaluation of acoustic feedback control methods: (a) speech source signal ($f_s = 16$ kHz), (b) music source signal ($f_s = 44.1$ kHz).*
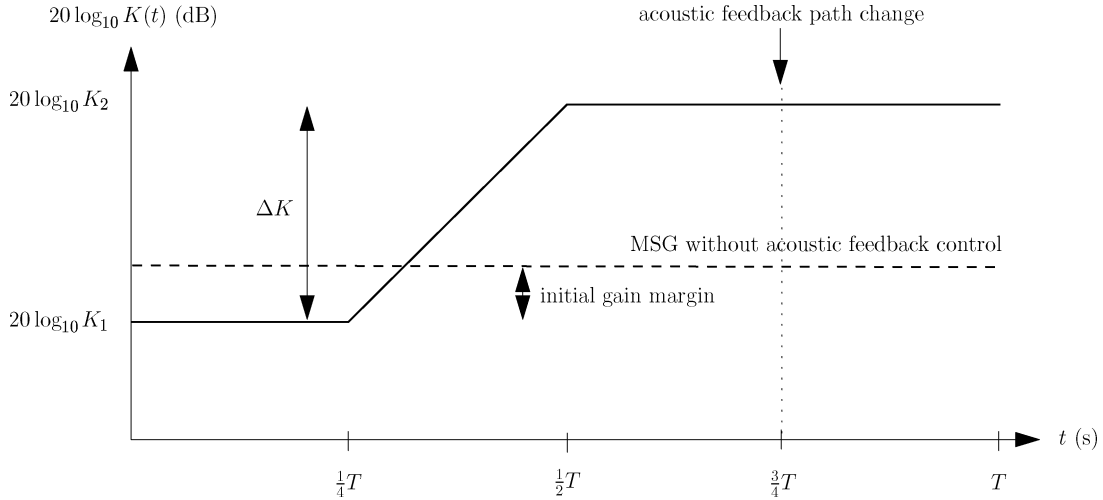
Fig. 15. *Electroacoustic forward path gain 20* $\log_{10} K(t)$ *versus time for the acoustic feedback control simulations.*

and Television Network (VRT), resampled to $f_s = 16$ kHz. The audio signal is an excerpt from a CD recording of the Partita No. 2 in D minor (Allemande) for solo violin by J. S. Bach, and is shown in Fig. 14(b). These signals were scaled to have a root mean square (RMS) value of -55 dBV, which corresponds to the output voltage of a typical microphone used in sound reinforcement applications.

Each simulation consists of four equally long phases, as shown in Fig. 15. In the first phase, the electroacoustic forward path broadband gain factor $K(t)$, defined in (14), is set to a value K_1 that would result in a 3-dB gain margin if no acoustic feedback control were performed. In particular, this first phase should allow the AFC algorithms to partially converge before the gain is increased beyond the point of instability. In the second phase, the gain $20 \log_{10} K(t)$ is then linearly increased up to a value $20 \log_{10} K_2 = 20 \log_{10} K_1 + \Delta K$ beyond the point of instability (where ΔK is defined on a decibel-scale for ease of notation). Since the different acoustic feedback control methods stabilize the closed-loop system to a different degree, the maximum gain increase ΔK that can be allowed while maintaining a stable operation [which should not be confused with the MSG defined in (16)] differs depending on which method is being used. More specifically, we have found that the maximum gain increase is around $\Delta K = 3$ dB for the PFC algorithms, $\Delta K = 5$ dB for the NHS algorithms, and $\Delta K = 10$ dB for the AFC algorithms. In the third and fourth phases of the simulation, the gain factor is fixed to K_2, and at the end of the third phase, an acoustic feedback path change is simulated. The acoustic feedback path used in the first three simulation phases corresponds to the RIR shown in Fig. 3, while the feedback path in the fourth phase is equal to the RIR measured in the same room as the first RIR, after a 1-m displacement of the microphone.

Our goal is to evaluate the acoustic feedback control methods based on three general objectives: the achievable amplification, the sound quality, and the reliability. These objectives can be quantified by a number of performance measures, which are calculated during the third and fourth simulation phases, since these phases correspond to the preferential mode of operation for the sound reinforcement system. The *achievable amplification* is measured by the MSG and the MSG increase, which by using (16) are defined as

$$\text{MSG}(t)[\text{dB}]$$
$$= -20 \log_{10}\left[\max_{\omega \in \mathcal{P}_H}|H(\omega,t)J(\omega,t)F(\omega,t)|\right] \tag{107}$$

$$\Delta\text{MSG}(t)[\text{dB}]$$
$$= -20 \log_{10}\left[\frac{\max_{\omega \in \mathcal{P}_H}|H(\omega,t)J(\omega,t)F(\omega,t)|}{\max_{\omega \in \mathcal{P}}|J(\omega,t)F(\omega,t)|}\right] \tag{108}$$

for the PFC and NHS methods, where $H(\omega,t)$ represents the frequency response of the PM filter or the bank of adjustable notch filters, respectively, and

$$\mathcal{P}_H = \{\omega \,|\, \angle\, H(\omega,t)G(\omega,t)F(\omega,t) = n2\pi\}.$$

In case of the AFC method, these measures are defined using (16) and (60) as follows:

$$\text{MSG}(t)[\text{dB}]$$
$$= -20 \log_{10}\left[\max_{\omega \in \mathcal{P}_{\hat{F}}}\left|J(\omega,t)\left[F(\omega,t) - \hat{F}(\omega,t)\right]\right|\right] \tag{109}$$

$$\Delta\text{MSG}(t)[\text{dB}]$$
$$= -20 \log_{10}\left[\frac{\max_{\omega \in \mathcal{P}_{\hat{F}}}\left|J(\omega,t)\left[F(\omega,t) - \hat{F}(\omega,t)\right]\right|}{\max_{\omega \in \mathcal{P}}|J(\omega,t)F(\omega,t)|}\right]. \tag{110}$$

We will use the instantaneous value of the MSG(t), as well as the mean and maximum value of the ΔMSG(t), as a performance measure in the evaluation.

An objective measure for quantifying the *sound quality* resulting from acoustic feedback control was proposed in the context of HA AFC in [161]. This measure, known as the frequency-weighted log-spectral signal distortion (SD), is defined as[9]

$$\text{SD}(t) = \sqrt{\int_0^{f_s/2} w_{\text{ERB}}(f)\left(10\log_{10}\frac{S_d(f,t)}{S_v(f,t)}\right)^2 df} \quad (111)$$

where $S_d(f,t)$ and $S_v(f,t)$ denote the short-term PSD of the feedback-compensated signal and source signal, respectively, and $w_{\text{ERB}}(f)$ is a weighting function that gives equal weight to each auditory critical band in the Nyquist interval, following Table 2 of the ANSI S3.5-1997 standard [162]. The short-term PSD is estimated as the squared magnitude of the short-term DFT, which is calculated using 50% overlapping data windows of length $M = 2048$ at $f_s = 16$ kHz, or $M = 4096$ at $f_s = 44.1$ kHz. The integration in (111) is then approximated by a summation over the DFT frequency bins. Both the mean and maximum value of the SD measure will be used in the evaluation.

Finally, the *reliability* is quantified using two performance measures: the howling occurence probability (HOP) and the time to recover from instability (TRI). These measures rely on an estimate of the time intervals during which howling occurs in the simulation. Howling occurrences are manually identified using the following procedure:

1) a rough estimate of the howling time intervals is obtained by listening to the feedback-compensated signal;
2) a spectrogram of the feedback-compensated signal is plotted for each of the time intervals identified in the first step, and the frequency bin(s) in which howling occurs are visually identified from the spectrogram;
3) a time-varying PAPR feature is calculated for each of the time intervals identified in the first step, where the peak PSD is estimated by averaging the power in the howling frequency bins identified in the second step;
4) the time interval during which howling occurs is then defined by the time points on either side of the PAPR maximum value, at which the PAPR has decreased to a value that is 3 dB below the maximum value.

[9]Note that in a real-time experiment, the source signal is not available, hence its PSD $S_v(\omega, t)$ cannot be calculated. The SD measure can then be calculated by comparing the loudspeaker signal PSD with the PSD of a reference signal that is obtained in a secondary experiment, in the absence of acoustic feedback [161].

From the time points identified in the last step of the above procedure, we can estimate the time duration Δt_i (s) of each howling occurrence, $i = 1, \ldots, N_{\text{HO}}$, with N_{HO} the number of howling occurrences estimated in the first step of the above procedure. The HOP and TRI measures are then defined as follows:

$$\text{HOP(\%)} = \frac{\sum_{i=1}^{N_{\text{HO}}} \Delta t_i}{T} \quad (112)$$

$$\text{TRI(s)} = \frac{\sum_{i=1}^{N_{\text{HO}}} \Delta t_i}{N_{\text{HO}}} \quad (113)$$

where T (s) denotes the length of the simulation.

C. Simulation Results

The instantaneous value of the MSG(t) measure versus time is displayed in Fig. 16 (where the left column contains the results obtained with the speech source signal, and the right column gives the results for the audio source signal). These MSG(t) curves have been smoothed with a one-pole low-pass filter to improve the clarity of the figures. The instantaneous value of the electroacoustic forward path gain $20\log_{10} K(t)$ and the MSG values obtained without acoustic feedback control are also shown (where "MSG $F_1(q)$" and "MSG $F_2(q)$" denote the MSG before and after the acoustic feedback path change, respectively). In the PFC simulation results shown in Fig. 16(a) and (b), the periodic behavior of the PM filters is clearly visible from the MSG curves. It can also be observed that these algorithms behave in a deterministic way, in the sense that their performance is independent of the instantaneous source signal and electro-acoustic forward path gain values. The PFC-DM algorithm generally performs somewhat worse compared to the other two PFC algorithms, while the PFC-PM algorithm performance can be seen to slightly improve at a higher sampling frequency. From the NHS simulation results shown in Fig. 16(c) and (d), the howling detection performance of the different NHS algorithms can also be judged. An instantaneous increase in the MSG curves indeed corresponds to the activation of a new notch filter (or the adjustment of an existing notch filter), while an MSG decrease occurs at the acoustic feedback path change. Ideally, no notch filters should be activated before the gain value $20\log_{10} K(t)$ exceeds the instantaneous MSG curves. However, this ideal behavior is exhibited only by the NHS-1 algorithm in the speech simulation. In all other cases, some notch filters are activated earlier, which indicates that some tonal source signal components are wrongly identified as howling components. The behavior of the different NHS algorithms in terms of the MSG(t) measure is comparable for speech source signals, while the NHS-1 and NHS-2 algorithms behave quite differently from the NHS-3 algorithm in the audio simulation. We should stress that the high MSG values obtained with the NHS-1 and NHS-2 algorithms in the audio
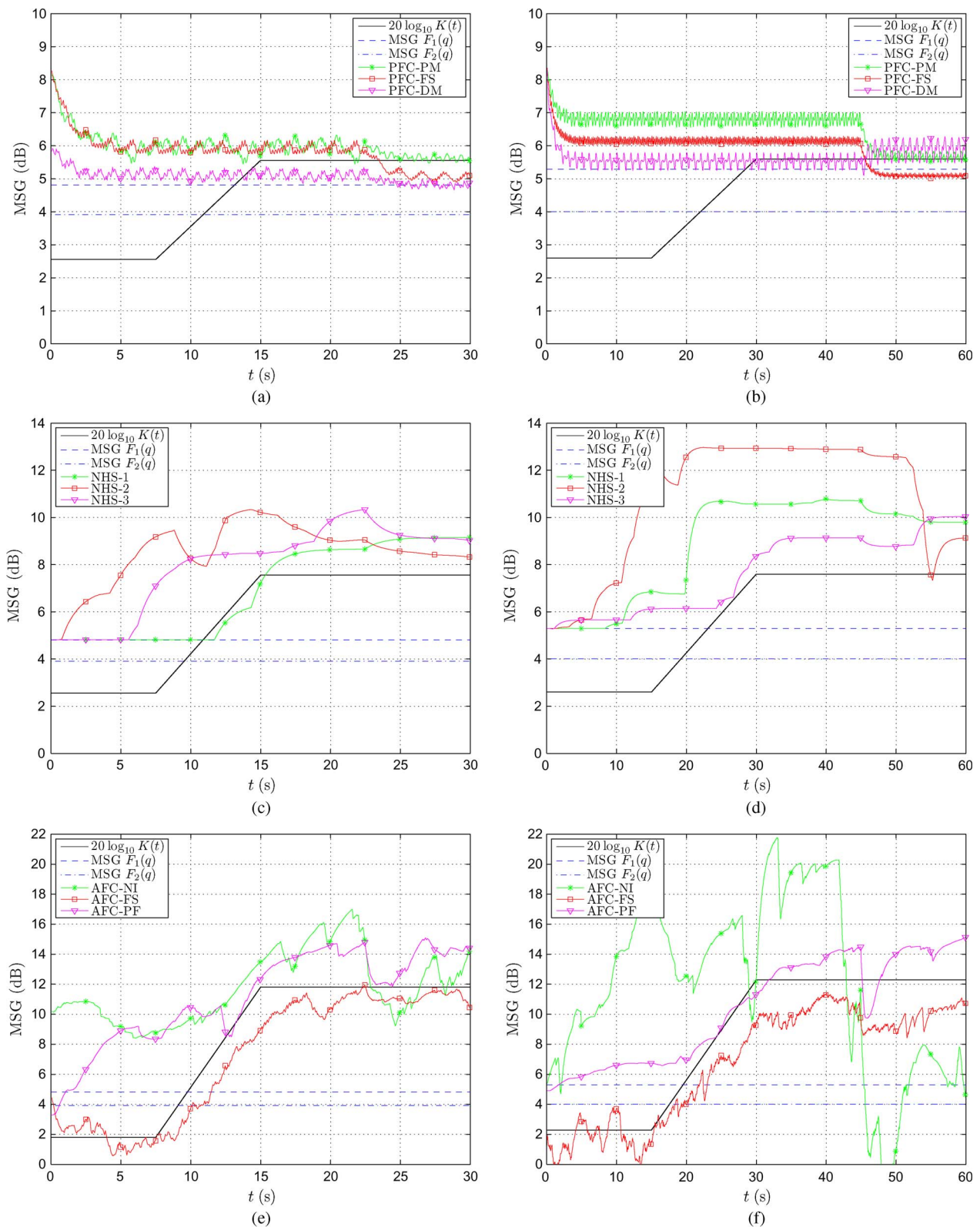
Fig. 16. *Instantaneous MSG versus time for simulations with speech (left column) and audio (right column) source signals: (a) and (b) PFC methods (ΔK = 3 dB), (c), (d) NHS methods (ΔK = 5 dB), (e), (f) AFC methods (ΔK = 10 dB). Note the scale difference on the vertical axis between (a)–(b), (c)–(d), and (e)–(f).*

simulation are in fact caused by an excessive amount of notch filtering that is due to the poor howling detection performance and leads to a broadband attenuation of the microphone signal. Hence the resulting sound quality obtained with these methods is extremely poor for audio applications (see the discussion on the results in Table 2). Finally, the simulation results obtained with the AFC algorithms are shown in Fig. 16(e) and (f). In the speech simulation, the MSG performance of the AFC-NI and AFC-PF algorithms appears to be better compared to the AFC-FS algorithm. In the audio simulation, the AFC-NI algorithm initially outperforms the other algorithms, however, the AFC-PF algorithm eventually provides the highest MSG value. All three AFC algorithms appear to react in a relatively robust way to the acoustic feedback path change in the fourth phase of the simulation, except for the AFC-NI algorithm in the audio simulation.

The performance measures calculated during the third and fourth simulation phases are shown in Tables 1 and 2 for the speech and audio simulations, respectively. Some general observations can be made concerning the performance of the different acoustic feedback control methods. The achievable amplification in terms of the MSG increase is relatively low for the PFC algorithms, and highest for the AFC algorithms, which is consistent with the MSG increase values reported in the literature. It can also be observed that, for the NHS and AFC algorithms, the MSG increase is larger when the electroacoustic forward path gain is raised

to a higher value. This effect can be explained by noting that more notch filters are activated as the gain is increased, while the AFC convergence is known to benefit from a gain increase since the power ratio of the feedback signal and source signal then also increases [103], [111]. In terms of sound quality, the SD performance measure reveals that the perceptual signal distortion is worse for the PFC algorithms and for the AFC-NI algorithm. The other AFC algorithms provide a much higher sound quality, and generally perform somewhat better than the NHS algorithms. As mentioned earlier, the NHS-1 and NHS-2 algorithms result in an extremely poor sound quality in audio applications, which is due to the poor howling detection performance. The reliability of the evaluated algorithms is seen to be slightly worse in the audio simulation as compared to the speech simulation, especially for the PFC algorithms.

Within each acoustic feedback control method, the relative performance of the different algorithms can be compared using the measures in Tables 1 and 2. Among the PFC algorithms, the PFC-PM algorithm should generally be preferred since it performs best in terms of nearly all performance measures. Among the NHS algorithms, the NHS-3 algorithm is the only algorithm that is suited for audio applications, and moreover, in terms of achievable amplification and sound quality, this algorithm outperforms the NHS-1 and NHS-2 algorithms for speech applications also. We should note, however, that the NHS-3 howling detection method is computationally more

Table 1 Performance Measures for Comparative PFC, NHS, and AFC Simulations: Speech Source Signal

			PFC			NHS			AFC		
			PFC-PM	PFC-FS	PFC-DM	NHS-1	NHS-2	NHS-3	AFC-NI	AFC-FS	AFC-PF
achievable amplification	mean(ΔMSG)	$\Delta K = 3$	1.4	1.1	0.6	2.2	4.4	4.2	6.8	1.3	4.5
		$\Delta K = 5$				4.5	4.5	5.0	7.8	3.1	6.9
		$\Delta K = 10$							9.8	6.6	9.6
	max(ΔMSG)	$\Delta K = 3$	4.1	4.1	4.7	3.0	5.2	5.3	9.1	8.6	8.1
		$\Delta K = 5$				5.2	5.2	5.6	10.5	9.5	9.3
		$\Delta K = 10$							13.7	11.1	12.8
sound quality	mean(SD)	$\Delta K = 3$	6.2	7.1	7.9	3.5	3.8	3.1	13.8	5.6	2.4
		$\Delta K = 5$				3.8	4.6	3.7	14.0	5.6	2.6
		$\Delta K = 10$							15.1	6.0	3.9
	max(SD)	$\Delta K = 3$	10.7	11.6	16.2	7.5	6.5	5.6	30.4	8.3	6.6
		$\Delta K = 5$				6.6	7.8	6.9	30.1	8.4	6.5
		$\Delta K = 10$							31.7	10.6	16.2
reliability	HOP (%)	$\Delta K = 3$	0	0	0	3.6	0	0	0	0	0
		$\Delta K = 5$				8.5	0	1.5	0	0	0
		$\Delta K = 10$							1.3	0	2.6
	TRI (s)	$\Delta K = 3$	N/A	N/A	N/A	1.09	N/A	N/A	N/A	N/A	N/A
		$\Delta K = 5$				0.85	N/A	0.22	N/A	N/A	N/A
		$\Delta K = 10$							0.19	N/A	0.77

Table 2 Performance Measures for Comparative PFC, NHS, and AFC Simulations: Audio Source Signal

			PFC			NHS			AFC		
			PFC-PM	PFC-FS	PFC-DM	NHS-1	NHS-2	NHS-3	AFC-NI	AFC-FS	AFC-PF
achievable amplification	mean(ΔMSG)	$\Delta K = 3$	1.6	1.0	1.1	5.7	6.7	3.5	-3.2	0.1	3.0
		$\Delta K = 5$				5.7	7.1	4.8	-2.7	1.8	4.6
		$\Delta K = 10$							6.3	5.4	9.0
	max(ΔMSG)	$\Delta K = 3$	3.9	3.9	4.6	6.1	8.6	3.7	15.0	6.1	4.5
		$\Delta K = 5$				6.1	8.6	6.0	16.0	6.8	6.5
		$\Delta K = 10$							17.2	8.6	11.3
sound quality	mean(SD)	$\Delta K = 3$	8.9	52.1	9.2	6.7	39.1	3.3	19.0	6.4	3.7
		$\Delta K = 5$				6.8	38.3	4.1	19.2	6.5	4.0
		$\Delta K = 10$							19.9	7.1	5.3
	max(SD)	$\Delta K = 3$	23.9	72.7	25.4	17.3	63.7	5.8	27.3	10.8	6.1
		$\Delta K = 5$				16.7	67.4	6.3	27.7	11.1	7.1
		$\Delta K = 10$							28.2	14.7	19.6
reliability	HOP (%)	$\Delta K = 3$	11.1	52.0	19.3	0	0	0	0	2.2	0.5
		$\Delta K = 5$				0	0	2.2	0	2.2	0.5
		$\Delta K = 10$							0.4	4.7	7.2
	TRI (s)	$\Delta K = 3$	0.23	∞	0.27	N/A	N/A	N/A	N/A	0.67	0.33
		$\Delta K = 5$				N/A	N/A	0.67	N/A	0.67	0.33
		$\Delta K = 10$							0.23	0.46	1.44

demanding compared to the other NHS howling detection methods. Among the AFC algorithms, the AFC-NI algorithm yields the highest MSG increase in the speech simulation, which however comes at the cost of a poor sound quality. The AFC-PF algorithm provides the best sound quality and still allows for a relatively high MSG increase. In the audio simulation, the performance of the AFC-NI and AFC-FS algorithms is highly fluctuating, which can be observed from the discrepancy between the mean and maximum MSG values. The AFC-PF algorithm, on the other hand, produces a more steady MSG behavior in the audio simulation. The superior sound quality of the AFC-PF algorithm compared to all other evaluated algorithms results from the fact that the decorrelation is applied in the adaptive filtering circuit instead of in the closed signal loop. Note that the reliability of the AFC algorithms can be further improved by including additional features such as adaptation control, foreground/ background adaptive filtering, regularization, and post-filtering; see [103] for an overview.

The feedback-compensated signals obtained in the different simulations are all available for download,[10] such that the sound quality can be assessed subjectively by the reader. Also, the source signals and acoustic feedback path impulse responses used in the simulations can be downloaded for benchmarking purposes.

[10]ftp://ftp.esat.kuleuven.be/pub/sista/vanwaterschoot/abstracts/08-13.html

VIII. CONCLUSION AND FUTURE CHALLENGES

In this paper, we have attempted to provide a comprehensive overview of five decades of research in acoustic feedback control. The available literature has been reviewed following a classification of the state-of-the-art solutions into four categories: PM methods, gain reduction methods, spatial filtering methods, and room modeling methods. We have also provided an in-depth treatment of three widely used acoustic feedback control methods, namely PFC, NHS, and AFC, thereby discussing conceptual as well as realization issues. Finally, several different realizations of these three methods have been evaluated and compared, in terms of their achievable amplification, sound quality, and reliability.

From the simulation results presented in this paper, we can conclude that the AFC method is superior to the PFC and NHS methods in terms of achievable amplification and sound quality, while its reliability is comparable to the reliability of the PFC and NHS methods. The AFC method should preferably be combined with a decorrelation approach that operates in the adaptive filtering circuit, e.g., using decorrelating prefilters (AFC-PF), since this approach appears to be beneficial w.r.t. the achievable amplification and sound quality. We have found the AFC-PF approach to be capable of providing an average MSG increase of approximately 9 dB, and a maximum MSG increase around 12 dB.

Looking into future research challenges in acoustic feedback control, it appears that there is little room for improvement in the PFC and NHS methods. Since these methods aim at smoothing the loop gain, a theoretical upper bound for the achievable MSG increase is given by the ratio of the peak and average magnitude response of the acoustic feedback path, which was found to be typically around 10 dB [2]. In practice, however, this upper bound is generally not achieved since the allowable values of the PFC modulation frequency and modulation depth are bounded by constraints on the signal distortion, while the number of active notch filters in the NHS method should be limited to avoid a broadband attenuation that ultimately affects sound quality. From our comparative simulation results, we may conclude that the best PFC solution consists in the use of a sinusoidal PM at low modulation frequency, while the preferable NHS solution is based on combining the howling detection method proposed by Osmanovic et al. [28]–[30] with a state-of-the-art biquadratic notch filter design method, e.g., the pole-zero placement technique recently proposed in [125].

On the other hand, we believe that since the AFC method appears to produce promising results, the main challenges for future research in acoustic feedback control lie in further increasing the AFC reliability and reducing its computational complexity. In terms of reliability, recent research has pointed out that so-called hybrid AFC methods, in which AFC is combined with other methods for acoustic feedback control, are far more robust compared to the traditional AFC approach. However, we believe that in the existing hybrid AFC methods, the cooperation between the different methods is still suboptimal. For example, in the combined AFC and postfiltering methods proposed in [66], [93], and [94], the postfilter design is solely based on the feedback-compensated signal spectrum, while it is known from AEC that the joint design of a cancellation filter and a postfilter generally results in a better performance [158], [159]. A related issue is the combination of AFC with a gain reduction method: in [74] and [106]–[108], the AFC and ANF filters are adapted independently, while in the combined AFC and AEQ approach proposed in [66] and in the combined AFC and NHS approach proposed in [103], the AEQ/NHS design is based on the most recent AFC estimate. Similarly to the joint AFC and postfilter design, it can be expected that a joint estimation of the AFC and gain reduction filter coefficients is to be preferred over a decoupled estimation. Finally, a similar remark can be made on the joint design of an AFC and a spatial filtering method, which would probably outperform the state-of-the-art approach of AFC combined with a fixed beamformer [74] or an adaptive beamformer steered by the feedback-compensated signal [66].

The greatest challenge in AFC, however, consists in reducing the computational complexity. Since typically an already cheap NLMS-type algorithm is used, a significant complexity reduction in the AFC adaptive filtering algorithm cannot be expected. The fundamental problem lies in the fact that in AFC, the acoustic feedback path is traditionally modeled using its impulse response, which typically has a large number of coefficients. This is especially so when a high sampling frequency is applied (e.g., in audio applications). The impulse response is then more densely sampled and in addition more adaptive filter updates have to be performed per second. However, from a stability point of view, it may suffice to only model the peaks in the acoustic feedback path magnitude response instead of the complete impulse response. This may be achieved with frequency domain adaptive filtering (FDAF). However, since the frequency domain models currently used in FDAF have a fixed and uniform frequency resolution, the required FDAF filter order should still be high to guarantee that the magnitude peaks are modeled with sufficient accuracy; see, e.g., the FDAF experimental results in [103]. Another possibility for reducing the acoustic feedback path model complexity consists in using a time domain model different from the FIR model. Since the peaks in the acoustic feedback path magnitude response can be modeled as narrowband resonances, an IIR (or pole-zero) model seems to be an appropriate choice. The use of such models in room acoustics has both been recommended [134], [163], [164] and discouraged [165], [166], however, no results on the use of IIR models in AFC are available. The appeal of using such models in room acoustic applications is related to the conjecture that the IIR model denominator coefficients can in fact be assumed time invariant in a certain acoustic environment, regardless of the loudspeaker and microphone positions [134]. A related model, which also exploits the assumption of time-invariant room acoustic resonance frequencies, is based on the use of orthogonal basis functions such as the discrete-time Laguerre or Kautz functions, which have been evaluated in an AEC context in [167] and [168].

Another great challenge in acoustic feedback control, and in AFC in particular, is to generalize the methods proposed in a single-channel context to multichannel systems. Since the number of acoustic feedback paths in a multichannel system equals the number of loudspeakers times the number of microphones, the AFC computational complexity can be expected to increase very quickly in a multichannel context. Again, the use of IIR models or models based on orthogonal basis functions may bring some relief, since, following the arguments in [134] and [168], these models could then share a common denominator. Another problem arising in multichannel AFC is related to the identifiability of the acoustic feedback path models in case the loudspeaker signals are correlated. A similar problem occurs in multichannel AEC, and has received quite some attention in the literature; see, e.g., [153] and [169]. ∎

REFERENCES

[1] D. A. Bohn, "Pro audio reference," ©Rane Corporation, Mukilteo, WA, Aug. 2010. [Online]. Available: http://www.rane.com/digi-dic.html

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

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

[4] J. C. Willems, *The Analysis of Feedback Systems*. Cambridge, MA: MIT Press, 1971.

[5] R. W. Guelke and A. D. Broadhurst, "Reverberation time control by direct feedback," *Acustica*, vol. 24, pp. 33–41, 1971.

[6] P. Mapp and C. Ellis, "Improvements in acoustic feedback margin in sound reinforcement systems," in *Preprints AES 105th Convention*, San Francisco, CA, Sep. 1998, AES Preprint 4850.

[7] M. R. Schroeder, "Improvement of acoustic feedback stability in public address systems," in *Proc. 3rd Int. Congr. Acoust.*, Stuttgart, Germany, 1959, pp. 771–775.

[8] M. R. Schroeder, "Stop feedback in public address systems," *Radio Electron.*, vol. 31, pp. 40–42, Feb. 1960.

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

[10] A. J. Prestigiacomo and D. J. MacLean, "A frequency shifter for improving acoustic feedback stability," *J. Audio Eng. Soc.*, vol. 10, no. 2, pp. 110–113, Apr. 1962.

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

[12] C. Vila Deutschbein, "Digital frequency shifting for electroacoustic feedback suppression," in *Preprints AES 118th Conv.*, Barcelona, Spain, May 2005, AES Preprint 6505.

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

[14] L. N. Mishin, "A method for increasing the stability of sound amplification systems," *Akust. Z.*, vol. 4, pp. 64–72, Jan./Mar. 1958.

[15] G. Nishinomiya, "Improvement of acoustic feedback stability of public address system by warbling," in *Proc. 6th Int. Congr. Acoust.*, Tokyo, Japan, 1968, vol. E, pp. 93–96.

[16] P. U. Svensson, "Computer simulations of periodically time-varying filters for acoustic feedback control," *J. Audio Eng. Soc.*, vol. 43, no. 9, pp. 667–677, Sep. 1995.

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

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

[19] P. Svensson, "On reverberation enhancement in auditoria," Ph.D. dissertation, Dept. Appl. Acoust., Chalmers Univ. Technol., Gothenburg, Sweden, 1994.

[20] E. T. Patronis, Jr., "Acoustic feedback detector and automatic gain control," U.S. Patent 4 079 199, Mar. 1978.

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

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

[23] Y. Nagata, S. Suzuki, M. Yamada, M. Yoshida, M. Kitano, K. Kuroiwa, and S. Kimura, "Howling remover having cascade connected equalizers suppressing multiple noise peaks," U.S. Patent 5 710 823, Jan. 1998.

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

[25] M. Hanajima, M. Yoneda, and T. Okuma, "Howling eliminator," WIPO Patent Appl. WO/1999/021 396, Apr. 1999.

[26] M. Hanajima, M. Yoneda, and T. Okuma, "Howling eliminating apparatus," U.S. Patent 6 125 187, Sep. 2000.

[27] Y. Terada and A. Murase, "Howling control device and howling control method," U.S. Patent 7 190 800, Mar. 2007.

[28] N. Osmanovic, V. E. Clarke, and E. Velandia, "An in-flight low latency acoustic feedback cancellation algorithm," in *Preprints AES 123rd Conv.*, New York, Oct. 2007, AES Preprint 7266.

[29] N. Osmanovic and V. Clarke, "Acoustic feedback cancellation system," WIPO Patent Appl. WO/2007/013 981, Feb. 2007.

[30] N. Osmanovic and V. Clarke, "Acoustic feedback cancellation system," U.S. Patent 7 664 275, Feb. 2010.

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

[32] D. M. Oster, M. P. Lewis, and T. J. Tucker, "Method and apparatus for adaptive audio resonant frequency filtering," WIPO Patent Appl. WO/1991/020 134, Dec. 1991.

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

[34] M. P. Lewis, T. J. Tucker, and D. M. Oster, "Method and apparatus for adaptive audio resonant frequency filtering," U.S. Patent 5 245 665, Sep. 1993.

[35] M. H. Er, T. H. Ooi, L. S. Li, and C. J. Liew, "A DSP-based acoustic feedback canceller for public address systems," in *Proc. Int. Conf. Signal Process.*, Beijing, China, Oct. 1993, pp. 1251–1254.

[36] M. H. Er, T. H. Ooi, L. S. Li, and C. J. Liew, "A DSP-based acoustic feedback canceller for public address systems," *Microprocessors Microsyst.*, vol. 18, no. 1, pp. 39–47, Jan./Feb. 1994.

[37] A. Kawamura, M. Matsumoto, M. Serikawa, and H. Numazu, "Sound amplifying apparatus with automatic Howl-suppressing function," Eur. Patent Appl. EP0 599 450 A2, Jun. 1994.

[38] A. Kawamura, M. Matsumoto, M. Serikawa, and H. Numazu, "Sound amplifying apparatus with automatic Howl-suppressing function," U.S. Patent 5 442 712, Aug. 1995.

[39] M. Tahernezhadi and L. Liu, "An adaptive notch filter for howling cancellation," *Acoust. Lett.*, vol. 18, no. 8, pp. 142–145, 1995.

[40] W. Staudacher, "Acoustic feedback cancellation for equalized amplifying systems," U.S. Patent 5 533 120, Jul. 1996.

[41] J. Timoney and F. B. Foley, "Robust performance of the adaptive periodic noise canceller in a closed-loop system," in *Proc. 9th Eur. Signal Process. Conf.*, Rhodes, Greece, Sep. 1998, pp. 1177–1180.

[42] J. E. Lane, D. Hoory, and J. Choe, "Method and apparatus for suppressing acoustic feedback in an audio system," U.S. Patent 5 717 772, Feb. 1998.

[43] R. Porayath and D. J. Mapes-Riordan, "Acoustic feedback elimination using adaptive notch filter algorithm," U.S. Patent 5 999 631, Dec. 1999.

[44] P. R. Williams, "Method and system for elimination of acoustic feedback," WIPO Patent Appl. WO/2002/021 817, Mar. 2002.

[45] P. R. Williams, "Method and system for elimination of acoustic feedback," U.S. Patent Appl. 2010/0 046 768 A1, Feb. 2010.

[46] W. Loetwassana, R. Punchalard, and W. Silaphan, "Adaptive howling canceller using adaptive IIR notch filter: Simulation and implementation," in *Proc. IEEE Int. Conf. Neural Netw. Signal Process.*, Nanjing, China, Dec. 2003, pp. 848–851.

[47] J. Timoney, F. B. Foley, and A. T. Schwarzbacher, "An explicit criterion for adaptive periodic noise canceller robustness applied to feedback cancellation," in *Proc. 4th Electron. Circuits Syst. Conf.*, Bratislava, Slovakia, Sep. 2003, pp. 23–26.

[48] J. Wei, L. Du, Z. Chen, and F. Yin, "A new algorithm for howling detection," in *Proc. IEEE Int. Symp. Circuits Syst.*, Bangkok, Thailand, May 2003, vol. 4, pp. 409–411.

[49] A. F. Rocha and A. J. S. Ferreira, "An accurate method of detection and cancellation of multiple acoustic feedbacks," in *Preprints AES 118th Conv.*, Barcelona, Spain, May 2005, AES Preprint 6335.

[50] M. Börsch, "Method for constraining electroacoustic feedback," Eur. Patent Appl. EP1 684 543 A1, Jul. 2006.

[51] M. Börsch, "Method for suppressing electro-acoustic feedback," U.S. Patent Appl. 2006/0 159 282 A1, Jul. 2006.

[52] G. Rombouts, T. van Waterschoot, and M. Moonen, "Proactive notch filtering for acoustic feedback cancellation," in *Proc. 2nd Annu. IEEE Benelux/DSP Valley Signal Process. Symp.*, Antwerp, Belgium, Mar. 2006, pp. 169–172. [Online]. Available: ftp.esat.kuleuven.be/pub/sista/vanwaterschoot/abstracts/06-81.html

[53] R. Abe, "Howling suppression device and howling suppression method," U.S. Patent 7 295 670, Nov. 2007.

[54] W. Loetwassana, R. Punchalard, A. Lorsawatsiri, J. Koseeyaporn, and P. Wardkein, "Adaptive howling suppressor in an audio amplifier system," in *Proc. Asia-Pacific Conf. Commun.*, Bangkok, Thailand, Oct. 2007, pp. 445–448.

[55] D. Somasundaram, "Feedback cancellation in a sound system," Eur. Patent Appl. EP1 903 833 A1, Mar. 2008.

[56] D. Somasundaram, "Feedback cancellation in a sound system," U.S. Patent Appl. 2008/0 085 013 A1, Apr. 2008.

[57] T. Kawamura and T. Kanamori, "Howling detection device and method," U.S. Patent Appl. 2008/0 021 703 A1, Jan. 2008.

[58] P. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Regularized

adaptive notch filters for acoustic howling suppression," in *Proc. 17th Eur. Signal Process. Conf.*, Glasgow, Scotland, U.K., Aug. 2009, pp. 2574–2578.

[59] T. Ito, "Apparatus detecting howling by decay profile of impulse response in sound system," U.S. Patent 6 442 280, Aug. 2002.

[60] S. Ibaraki, H. Furukawa, and H. Naono, "Pre-howling howlback detection method," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Tokyo, Japan, Apr. 1986, pp. 941–944.

[61] Y. Takahashi, M. Tohyama, and Y. Yamasaki, "Cumulative spectral analysis for transient decaying signals in a transmission system including a feedback loop," *J. Audio Eng. Soc.*, vol. 54, no. 7/8, pp. 620–629, Jul./Aug. 2006.

[62] T. van Waterschoot and M. Moonen, "Comparative evaluation of howling detection criteria in notch-filter-based howling suppression," in *Preprints AES 126th Conv.*, Munich, Germany, May 2009 AES Preprint 7752.

[63] T. van Waterschoot and M. Moonen, "Comparative evaluation of howling detection criteria in notch-filter-based howling suppression," *J. Audio Eng. Soc.*, Nov. 2010.

[64] J. H. Stott and N. D. Wells, "Method and apparatus for reduction of unwanted feedback," U.S. Patent 6 269 165, Jul. 2001.

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

[66] C. P. Janse and H. J. W. Belt, "Sound reinforcement system having an echo suppressor and loudspeaker beamformer," WIPO Patent Appl. WO/2003/010 996, Feb. 2003.

[67] K. Kobayashi, K. Furuya, and A. Kataoka, "An adaptive microphone array for howling cancellation," *Acoust. Sci. Technol.*, vol. 24, no. 1, pp. 45–47, Jan. 2003.

[68] K. Kobayashi, K. Furuya, and A. Kataoka, "A microphone array for howling cancellation," *J. Acoust. Soc. Jpn.*, vol. 60, no. 3, pp. 115–125, Mar. 2004, (in Japanese).

[69] G. Rombouts, A. Spriet, and M. Moonen, "Generalized sidelobe canceller based acoustic feedback cancellation," in *Proc. 14th Eur. Signal Process. Conf.*, Firenze, Italy, Sep. 2006.

[70] G. Rombouts, A. Spriet, and M. Moonen, "Generalized sidelobe canceller based combined acoustic feedback- and noise cancellation," *Signal Process.*, vol. 88, no. 3, pp. 571–581, Mar. 2008.

[71] M. Goodwin and G. Elko, "Beam dithering: Acoustic feedback control using a modulated-directivity loudspeaker array," in *Preprints AES 93rd Conv.*, San Francisco, CA, Oct. 1992, AES Preprint 3384.

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

[73] S. Ushiyama, T. Hirai, M. Tohyama, and Y. Shimizu, "Howling suppression by smoothing the open-loop transfer function," IEICE Tech. Rep., vol. 94, no. 20, pp. 23–28, Apr. 1994, (in Japanese).

[74] G. Schmidt and T. Haulick, "Signal processing for in-car communication systems," *Signal Process.*, vol. 86, *Special Issue on Applied Speech and Audio Processing*, no. 6, pp. 1307–1326, Jun. 2006.

[75] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-36, no. 2, pp. 145–152, Feb. 1988.

[76] S. J. Elliot and P. A. Nelson, "Multiple-point equalization in a room using adaptive filters," *J. Audio Eng. Soc.*, vol. 37, no. 11, pp. 899–907, Nov. 1989.

[77] P. A. Nelson, H. Hameda, and S. J. Elliot, "Adaptive inverse filters for stereophonic sound reproduction," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1621–1632, Jul. 1992.

[78] J. C. Sarris, F. Jacobsen, and G. E. Cambourakis, "Sound equalization in a large region of a rectangular enclosure," *J. Acoust. Soc. Amer.*, vol. 116, no. 6, pp. 3271–3274, Dec. 2004.

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

[80] J. Hellgren and U. Forssell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 7, pp. 906–913, Nov. 2001.

[81] T. van Waterschoot, G. Rombouts, and M. Moonen, "On the performance of decorrelation by prefiltering for adaptive feedback cancellation in public address systems," in *Proc. 4th IEEE Benelux Signal Process. Symp.*, Hilvarenbeek, The Netherlands, Apr. 2004, pp. 167–170. [Online]. Available: ftp.esat.kuleuven.be/pub/sista/vanwaterschoot/abstracts/04-24.html

[82] T. van Waterschoot and M. Moonen, "Assessing the acoustic feedback control performance of adaptive feedback cancellation in sound reinforcement systems," in *Proc. 17th Eur. Signal Process. Conf.*, Glasgow, Scotland, U.K., Aug. 2009, pp. 1997–2001.

[83] S. Ibaraki, H. Furukawa, and H. Naono "Howling canceller," U.S. Patent 4 747 132, May 1988.

[84] A. Goertz, "An adaptive subtraction filter for feedback cancellation in public address sound systems," in *Proc. 15th Int. Congr. Acoust.*, Trondheim, Norway, Jun. 1995, pp. 69–72.

[85] T. van Waterschoot, "Akoestische feedbackonderdrukker," M.S. thesis, Dept. Electr. Eng., Katholieke Universiteit Leuven, Leuven, Belgium, Jun. 2001, (in Dutch).

[86] C. P. Janse and C. C. Tchang, "Acoustic feedback suppression," WIPO Patent Appl. WO/2005/079 109, Aug. 2005.

[87] C. P. Janse and P. A. A. Timmermans, "Signal amplifier system with improved echo cancellation," WIPO Patent Appl. WO/1995/28 034, Oct. 1995.

[88] C. P. Janse and P. A. A. Timmermans, "Signal amplifier system with improved echo cancellation," U.S. Patent 5 748 751, May 1998.

[89] S. Kamerling, K. Janse, and F. van der Meulen, "A new way of acoustic feedback suppression," in *Preprints AES 104th Conv.*, Amsterdam, The Netherlands, May 1998, AES Preprint 4735.

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

[91] T. van Waterschoot, K. Eneman, and M. Moonen, "Instrumental variable methods for acoustic feedback cancellation," Katholieke Universiteit Leuven, Leuven, Belgium, Tech. Rep. ESAT-SISTA TR 05-14. [Online]. Available: ftp.esat.kuleuven.be/pub/sista/vanwaterschoot/abstracts/05-14.html

[92] P. Estermann and A. Kaelin, "Feedback cancellation in hearing aids: Results from using frequency-domain adaptive filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, London, U.K., May/Jun. 1994, vol. 2, pp. 257–260.

[93] F. Gallego, E. Lleida, E. Masgrau, and A. Ortega, "Method and system for suppressing echoes and noises in environments under variable acoustic and highly feedback conditions," WIPO Patent Appl. WO/2002/101 728, Dec. 2002.

[94] A. Ortega, E. Lleida, and E. Masgrau, "Speech reinforcement system for car cabin communications," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 5, pp. 917–929, Sep. 2005.

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

[96] A. Spriet, M. Moonen, and I. Proudler, "Feedback cancellation in hearing aids: An unbiased modelling approach," in *Proc. 11th Eur. Signal Process. Conf.*, Toulouse, France, Sep. 2002, pp. 531–534.

[97] A. Ortega, E. Lleida, E. Masgrau, L. Buera, and A. Miguel, "Acoustic feedback cancellation in speech reinforcement systems for vehicles," in *Proc. Interspeech*, Lisbon, Portugal, Sep. 2005, pp. 2061–2064.

[98] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge, MA: MIT Press, 1986.

[99] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[100] G. Rombouts, T. van Waterschoot, K. Struyve, and M. Moonen, "Acoustic feedback suppression for long acoustic paths using a nonstationary source model," in *Proc. 13th Eur. Signal Process. Conf.*, Antalya, Turkey, Sep. 2005.

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

[102] G. Rombouts, P. Verhoeve, K. Struyve, T. van Waterschoot, and M. Moonen, "Circuit and method for estimating a room impulse response," Eur. Patent Appl. EP1 675 374 A1, Jun. 2006.

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

[104] T. van Waterschoot, G. Rombouts, and M. Moonen, "Dually regularized recursive prediction error identification for acoustic feedback and echo cancellation," in *Proc. 15th Eur. Signal Process. Conf.*, Poznań, Poland, Sep. 2007, pp. 1610–1614.

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

[106] T. Haulick, G. U. Schmidt, and H. Lenhardt, "Feedback reduction in communication systems," Eur. Patent EP1 679 874, May 2008.

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

[108] S. Cifani, R. Rotili, E. Principe, S. Squartini, and F. Piazza, "Real-time implementation of robust PEM-AFROW based solutions for acoustic feedback control," in Preprints AES 127th Conv., New York, Oct. 2009, AES Preprint 7899.

[109] H. Okumura and H. Fujita, "Adaptive howling canceller," Eur. Patent Appl. EP1 615 463 A2, Jan. 2006.

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

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

[112] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "On stationary linear time-varying systems," IEEE Trans. Circuits Syst., vol. CAS-29, no. 3, pp. 169–184, Mar. 1982.

[113] A. Papoulis, Signal Analysis. New York: McGraw-Hill, 1977.

[114] B. Boashash and A. P. Reilly, "Algorithms for time-frequency signal analysis," in Methods and Applications of Time-Frequency Signal Analysis, B. Boashash, Ed. Melbourne, Australia: Longman Cheshire, 1992.

[115] S. L. Marple, Jr., "Computing the discrete-time 'analytic' signal via FFT," IEEE Trans. Signal Process., vol. 47, no. 9, pp. 2600–2603, Sep. 1999.

[116] A. Reilly, G. Frazer, and B. Boashash, "Analytic signal generation-tips and traps," IEEE Trans. Signal Process., vol. 42, no. 11, pp. 3241–3245, Nov. 1994.

[117] J. Dattorro, "Effect design—Part 2: Delay-line modulation and chorus," J. Audio Eng. Soc., vol. 45, no. 10, pp. 764–788, Oct. 1997.

[118] S. Disch and U. Zölzer, "Modulation and delay line based digital audio effects," in Proc. 2nd COST G-6 Workshop Digital Audio Effects, Trondheim, Norway, Dec. 1999, pp. 5–8.

[119] P. Dutilleux and U. Zölzer, "Delays," in DAFX: Digital Audio Effects, U. Zölzer, Ed. New York: Wiley, 2002.

[120] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay: Tools for fractional delay filter design," IEEE Signal Process. Mag., vol. 13, no. 1, pp. 30–60, Jan. 1996.

[121] D. Griesinger, "Improving room acoustics through time-variant synthetic reverberation Preprints AES 90th Conv., Paris, France, Feb. 1991, AES Preprint 3014.

[122] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," Proc. IEEE, vol. 66, no. 1, pp. 51–83, Jan. 1978.

[123] J. O. Smith, Mathematics of the Discrete Fourier Transform (DFT), accessed Nov. 2008. [Online]. Available: http://ccrma.stanford.edu/~jos/mdft/

[124] A. V. Oppenheim, D. H. Johnson, and K. Steiglitz, "Computation of spectra with unequal resolution using the fast Fourier transform," Proc. IEEE, vol. 59, no. 2, pp. 299–301, Feb. 1971.

[125] T. van Waterschoot and M. Moonen, "A pole-zero placement technique for designing second-order IIR parametric equalizer filters," IEEE Trans. Audio Speech Lang. Process., vol. 15, no. 8, pp. 2561–2565, Nov. 2007.

[126] J. A. Moorer, "The manifold joys of conformal mapping: Applications of digital filtering in the studio," J. Audio Eng. Soc., vol. 31, no. 11, pp. 826–841, Nov. 1983.

[127] K. Hirano, S. Nishimura, and S. K. Mitra, "Design of digital notch filters," IEEE Trans. Commun., vol. COM-22, no. 7, pp. 964–970, Jul. 1974.

[128] S. A. White, "Design of a digital biquadratic peaking or notch filter for digital audio equalization," J. Audio Eng. Soc., vol. 34, no. 6, pp. 479–483, Jun. 1986.

[129] P. A. Regalia and S. K. Mitra, "Tunable digital frequency response equalization filters," IEEE Trans. Acoust. Speech Signal Process., vol. ASSP-35, no. 1, pp. 118–120, Jan. 1987.

[130] D. J. Shpak, "Analytical design of biquadratic filter sections for parametric filters," J. Audio Eng. Soc., vol. 40, no. 11, pp. 876–885, Nov. 1992.

[131] D. C. Massie, "An engineering study of the four-multiply normalized ladder filter," J. Audio Eng. Soc., vol. 41, no. 7/8, pp. 564–582, Jul./Aug. 1993.

[132] R. Bristow-Johnson, "The equivalence of various methods of computing biquad coefficients for audio parametric equalizers," in Preprints AES 97th Conv., San Francisco, CA, Nov. 1994, AES Preprint 3906.

[133] S. J. Orfanidis, "Digital parametric equalizer design with prescribed Nyquist-frequency gain," J. Audio Eng. Soc., vol. 45, no. 6, pp. 444–455, Jun. 1997.

[134] Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," IEEE Trans. Speech Audio Process., vol. 2, no. 2, pp. 320–328, Apr. 1994.

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

[136] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Upper Saddle River, NJ: Prentice-Hall, 1993.

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

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

[139] S. Haykin, Adaptive Filter Theory. Englewood Cliffs, NJ: Prentice-Hall, 1996.

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

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

[142] S. H. Jensen, "Acoustic echo canceller for hands-free mobile radiotelephony," in Proc. 6th Eur. Signal Process. Conf., Brussels, Belgium, Aug. 1992, pp. 1629–1632.

[143] S. L. Gay, "Dynamically regularized fast RLS with application to echo cancellation," in Proc. IEEE Int. Conf. Acoust. Speech Signal Process., Atlanta, GA, May 1996, vol. 2, pp. 957–960.

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

[145] B. Baykal and A. G. Constantinides, "Underdetermined-order recursive least-squares adaptive filtering: The concept and algorithms," IEEE Trans. Signal Process., vol. 45, no. 2, pp. 346–362, Feb. 1997.

[146] M. Rupp, "A family of adaptive filter algorithms with decorrelating properties," IEEE Trans. Signal Process., vol. 46, no. 3, pp. 771–775, Mar. 1998.

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

[148] M. Tanaka, S. Makino, and J. Kojima, "A block exact fast affine projection algorithm," IEEE Trans. Speech Audio Process., vol. 7, no. 1, pp. 79–86, Jan. 1999.

[149] G. Rombouts and M. Moonen, "A sparse block exact affine projection algorithm," IEEE Trans. Speech Audio Process., vol. 10, no. 2, pp. 100–108, Feb. 2002.

[150] G. Rombouts, T. van Waterschoot, K. Struyve, P. Verhoeve, and M. Moonen, "Identification of undermodelled room impulse responses," in Proc. Int. Workshop Acoust. Echo Noise Control, Eindhoven, The Netherlands, Sep. 2005, pp. 153–156.

[151] R. D. Poltmann, "Stochastic gradient algorithm for system identification using adaptive FIR-filters with too low number of coefficients," IEEE Trans. Circuits Syst., vol. CAS-35, no. 2, pp. 247–250, Feb. 1988.

[152] C. Paleologu, S. Ciochina, and J. Benesty, "Variable step-size NLMS algorithm for under-modeling acoustic echo cancellation," IEEE Signal Process. Lett., vol. 15, pp. 5–8, 2008.

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

[154] Y. Qian, G. Chahine, and P. Kabal, "Pseudo-multi-tap pitch filters in a low bit-rate CELP speech coder," Speech Commun., vol. 14, no. 4, pp. 339–358, Sep. 1994.

[155] J. Makhoul, "Linear prediction: A tutorial review," Proc. IEEE, vol. 63, no. 4, pp. 561–578, Apr. 1975.

[156] T. van Waterschoot and M. Moonen, "Comparison of linear prediction models for audio signals," EURASIP J. Audio Speech Music Process., vol. 2008, 2008, Article ID 706935.

[157] S. Gustafsson, R. Martin, and P. Vary, "Combined acoustic echo control and noise reduction for hands-free telephony," Signal Process., vol. 64, Special Issue on Acoustic Echo and Noise Control, no. 1, pp. 21–32, Jan. 1998.

[158] E. Haensler and G. U. Schmidt, "Hands-free telephones-joint control of echo cancellation and postfiltering," Signal Process., vol. 80, no. 11, pp. 2295–2305, Nov. 2000.

[159] G. Enzner and P. Vary, "Frequency-domain adaptive Kalman filter for acoustic echo control in hands-free telephones," *Signal Process.*, vol. 86, *Special Issue on Applied Speech and Audio Processing*, no. 6, pp. 1140–1156, Jun. 2006.

[160] T. van Waterschoot, G. Rombouts, and M. Moonen, "MSE optimal regularization of APA and NLMS algorithms in room acoustic applications," in *Proc. Int. Workshop Acoust. Echo Noise Control*, Paris, France, Sep. 2006.

[161] A. Spriet, K. Eneman, M. Moonen, and J. Wouters, "Objective measures for real-time evaluation of adaptive feedback cancellation algorithms in hearing aids," in *Proc. 16th Eur. Signal Process. Conf.*, Lausanne, Switzerland, Aug. 2008.

[162] *American National Standard Methods for Calculation of the Speech Intelligibility Index*, ANSI Std. S3.5-1997, 1997.

[163] J. Pongsiri, P. Amin, and C. Thompson, "Modeling the acoustic transfer function of a room," in *Proc. 12th Int. Conf. Math. Comput. Model. Sci. Comput.*, Chicago, IL, Aug. 1999, pp. 44–51.

[164] T. Gustafsson, J. Vance, H. R. Pota, B. D. Rao, and M. M. Trivedi, "Estimation of acoustical room transfer functions," in *Proc. 39th IEEE Conf. Decision Control*, Sydney, Australia, Dec. 2000, pp. 5184–5189.

[165] S. Gudvangen and S. J. Flockton, "Comparison of pole-zero and all-zero modelling of acoustic transfer functions," *IEE Electron. Lett.*, vol. 28, no. 21, pp. 1976–1978, Oct. 1992.

[166] A. P. Liavas and P. A. Regalia, "Acoustic echo cancellation: Do IIR models offer better modeling capabilities than their FIR counterparts?" *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2499–2504, Sep. 1998.

[167] G. W. Davidson and D. D. Falconer, "Reduced complexity echo cancellation using orthonormal functions," *IEEE Trans. Circuits Syst.*, vol. 38, no. 1, pp. 20–28, Jan. 1991.

[168] L. S. H. Ngia, "Recursive identification of acoustic echo systems using orthonormal basis functions," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 3, pp. 278–293, May 2003.

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

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

ABOUT THE AUTHORS

Toon van Waterschoot was born in Lier, Belgium, on June 11, 1979. He received the M.S. and Ph.D. degrees in electrical engineering from Katholieke Universiteit Leuven (K.U.Leuven), Leuven, Belgium, in 2001 and 2009, respectively.

Since 2010, he has been a Postdoctoral Research Fellow with Delft University of Technology (TU Delft), Delft, The Netherlands. In 2002, he spent a year as a Teaching Assistant with the Antwerp Maritime Academy (Hogere Zeevaartschool Antwerpen), Belgium. From 2002 to 2003, and from 2008 to 2009, he was a Research Assistant with K.U.Leuven, Belgium, while from 2004 to 2007, he was a Research Assistant with the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT), Belgium. After his Ph.D. graduation, he was a Postdoctoral Research Fellow with K.U.Leuven, Belgium, until 2010. Since 2005, he has been a Visiting Teaching Assistant at the Advanced Learning and Research Institute, University of Lugano (Università della Svizzera Italiana), Switzerland, where he is teaching Digital Signal Processing. His research interests are in adaptive signal processing and parameter estimation, with application to acoustic signal enhancement, speech and audio processing, and wireless communications.

Dr. van Waterschoot served as a Technical Program Committee (TPC) Track Chair for Speech Processing at the 18th European Signal Processing Conference (EUSIPCO-2010), and has been a technical reviewer and TPC member for numerous journals and conferences.

Marc Moonen received the electrical engineering degree and the Ph.D. degree in applied sciences from Katholieke Universiteit Leuven (K.U.Leuven), Leuven, Belgium, in 1986 and 1990, respectively.

Since 2004, he has been a Full Professor at the Electrical Engineering Department, K.U.Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL and audio signal processing.

Dr. Moonen received the 1994 K.U.Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), the 2004 Alcatel Bell (Belgium) Award (with Raphael Cendrillon), and was a 1997 "Laureate of the Belgium Royal Academy of Science." He received a journal best paper award from the IEEE Transactions on Signal Processing (with G. Leus) and from *Elsevier Signal Processing* (with S. Doclo). He was Chairman of the IEEE Benelux Signal Processing Chapter (1998–2002) and is currently Past-President of the European Association for Signal Processing (EURASIP) and a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications. He has served as Editor-in-Chief for the *EURASIP Journal on Applied Signal Processing* (2003–2005), and has been a member of the editorial board of the IEEE Transactions on Circuits and Systems—Part II: Analog and Digital Signal Processing (2002–2003), the IEEE Signal Processing Magazine (2003–2005), and the *Integration, the VLSI Journal*. He is currently a member of the editorial board of the *EURASIP Journal on Applied Signal Processing*, the *EURASIP Journal on Wireless Communications and Networking*, and *Signal Processing*.