

# Analysis of Acoustic Feedback Cancellation Systems based on Direct Closed-Loop Identification

Bruno C. Bispo, César F. Yamamura, Wellington M. S. Nogueira, Edson A. R. Theodoro and Pedro M. Rodrigues

**Abstract**—This work presents, using the least squares estimation theory, a theoretical and experimental analysis on the performance of the standard adaptive filtering algorithms when applied to acoustic feedback cancellation. Expressions for the bias and covariance matrix of the acoustic feedback path estimate provided by these algorithms are derived as a function of the signals statistics as well as derivatives of the cost function. It is demonstrated that, in general, the estimate is biased and presents a large covariance because the closed-loop nature of the system makes the cross-correlation between the loudspeaker and system input signals non-zero. Simulations are carried out to exemplify the results using speech signals, a long acoustic feedback path and the recursive least squares algorithm. The results illustrate that these algorithms converge very slowly to a solution that is not the true acoustic feedback path. The relationship between the performance of the adaptive filtering algorithms and the aforementioned cross-correlation is proven by varying the signal-to-noise ratio and the delay introduced by the forward path.

**Index Terms**—sound reinforcement system, hearing aid, acoustic feedback cancellation, adaptive filtering, least squares.

## I. INTRODUCTION

A sound reinforcement (SR) system essentially comprises microphones, amplifiers and loudspeakers. Its fundamental purpose is to pick up, amplify and play back one or more desired sounds signals in the same acoustic environment. In many situations, mixing consoles and signal processors are employed to combine and enhance the sound signals.

A SR system may be very complex, including hundreds of microphones and multiple loudspeaker arrays. But nearly every system covered in the literature is single-channel, i.e., utilizes one microphone and one loudspeaker [1]. For this reason, only single-channel SR systems will be addressed in this work.

The acoustic coupling between loudspeaker and microphone unavoidably cause the loudspeaker signal to be fed back into the microphone. Thus, a closed signal loop is generated and leads to the so-called acoustic feedback problem [1]. The acoustic coupling and the signal processing circuit are referred to as acoustic feedback path and forward path, respectively.

The acoustic feedback affects the performance of a SR system in two ways. First and foremost, the closed-loop system can exhibit instability, which may lead to oscillations that are

acoustically perceived as howling. This phenomenon is also known as Larsen effect [1]. As a consequence, the achievable amplification is limited [1], [2]. Second, the sound quality is deteriorated by excessive reverberation or even ringing.

Most people have witnessed the howling effect in public address (PA) systems widely used in presentations, lectures, shows, events in general [1], [2], [3]. But it also occurs in hearing aids (HA) [4], [5], [6], [7]. The howling effect is one of the most frequent complaints from users and the reason why many of them give up HA [8]. Industry experts estimate that up to 15% of the HA are returned to the factory within 90 days after manufacture because of feedback problems [9].

In order to control the howling effect and increase the achievable amplification, several methods have been developed over the past decades [1]. The acoustic feedback cancellation (AFC) methods aim at estimating the feedback signal and subtracting it from the microphone signal [1]. The predicted feedback signal is obtained by filtering the loudspeaker signal with a model of the acoustic feedback path. This model is calculated using an adaptive filter that is designed to estimate and track the feedback path by means of some iterative algorithm.

The algorithms of the adaptive filtering theory, which approximate the Wiener solution through the gradient or least squares, identify the feedback path by minimizing in some deterministic sense the error signal [10], [11]. In the target application, the error signal should be defined as the difference between the true and predicted feedback signals. But the microphone signal is actually used as the feedback path output. Besides the feedback signal, the microphone signal also comprises the desired sound signal and a possible ambient noise, which act jointly as interference to the adaptive algorithms.

This approach corresponds to direct closed-loop identification because the data used for identification is collected in closed loop but the identification is performed using an open-loop model [12]. In other words, the feedback path is identified using only measurements of its input and output, and no assumptions are made on how they are generated [13].

The cancellation architecture is similar to the acoustic echo cancellation (AEC) commonly used in teleconference systems. Although it is effective when applied to AEC, this scheme is inefficient when applied to AFC. The adaptive filtering algorithms premise that the signals acting as input and interference to the adaptive filter are uncorrelated [10], [11]. In AEC, these signals are independent. In AFC, on the other hand, these signals are strongly correlated, mainly when the desired sound signal has a high degree of spectral coloration. Therefore, the standard adaptive filtering algorithms perform poorly when applied to AFC [1], [4], [14].

B. C. Bispo is with the Department of Electrical and Electronic Engineering, Federal University of Santa Catarina (UFSC), Florianópolis, Brazil, e-mail: bruno.bispo@ufsc.br.

C. F. Yamamura, W. M. S. Nogueira, E. A. R. Theodoro are with the Department of Electrical Engineering, Federal University of Technology – Paraná (UTFPR), Cornélio Procopio, Brazil, e-mail: edsontheodoro@utfpr.edu.br.

P. M. Rodrigues is with the Universidade Católica Portuguesa, CBQF – Centro de Biotecnologia e Química Fina – Laboratório Associado, Escola Superior de Biotecnologia, Porto, Portugal, e-mail: prodrigues@porto.ucp.pt.  
Digital Object Identifier: 10.14209/jcis.2020.24

Most solutions to overcome this problem introduce a decorrelation procedure in the AFC approach with the aim of reducing the cross-correlation between the loudspeaker and interference signals [1], [14]. Some methods aim to change the loudspeaker signal with the disadvantage of a possible distortion. In this context, addition of different shaped noises to the loudspeaker signal, insertion of delays, nonlinear processing, as half-wave rectification, and time-varying processing, as frequency shifting, phase and delay modulation, into the forward path have been proposed [1], [4], [15], [16].

Other methods aim to create modified versions of the loudspeaker or microphone signals that are used only in updating the adaptive filter. Delaying the loudspeaker signal was first proposed [4]. Subsequent methods are based on the prediction error framework. They assume that the desired sound signal is the output of a filter, the source model, whose input is a zero-mean white noise [12], [17]. The loudspeaker and microphone signals are then prefiltered with the inverse source model in order to be whiten. In [5], [18], the source model is fixed. In [6], an adaptive filter estimates the source model in hearing aids. Short and long-time prediction filters are combined to estimate the source model in [19].

Another possible solution is to not utilize the standard gradient or least-squares based adaptive filtering algorithms to update the adaptive filter. Following this approach, the works in [2], [20] have demonstrated that the cepstra of the microphone and error signals can be defined as a function of the impulse responses of the feedback path, forward path and adaptive filter. And methods that estimate the feedback path impulse response from the signal cepstra to update the adaptive filter in a recursive fashion have been proposed [2], [20].

The poor performance of the standard adaptive filtering algorithms when applied to AFC has already been theoretically studied in some previous works. In [4], [21], [22], it was analyzed in a statistical framework using the Wiener theory. In [1], [14], it was analyzed in a deterministic framework using the least squares (LS) theory. And, in [5], [23], it was analyzed considering the specific prediction error framework. However, none of these works presented experiments using real speech signals and a long acoustic feedback path of a PA system to exemplify the conclusions drawn from the theoretical analysis.

This work has two goals: first, using the LS estimation theory, to present a theoretical analysis on the poor performance of the standard adaptive filtering algorithms when applied to AFC, compiling in detail the results available in the literature; second, to exemplify the conclusions drawn from the theoretical analysis using speech signals, a long acoustic feedback path of a PA system and the recursive LS (RLS) algorithm.

This paper is organized as follows: Section II presents the modelling of both acoustic feedback problem and acoustic feedback cancellation; Section III presents the theoretical analysis and compiles in detail the results available in the literature; Section IV describes the configuration of the simulated experiments; in Section V, the experimental results are presented and discussed based on the statistical properties of the desired sound signal, ambient noise and closed-loop system impulse response; finally, Section VI concludes the paper, emphasizing its main contributions.

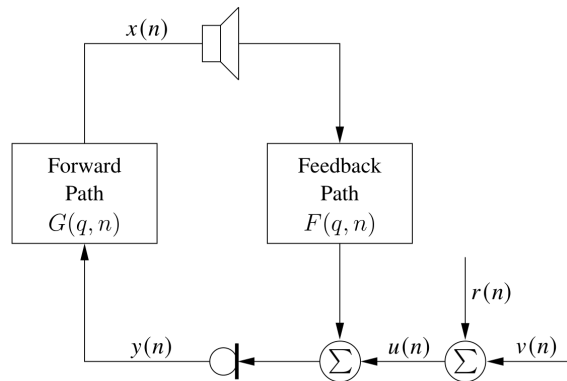


Fig. 1: Acoustic feedback in SR systems.

## II. ACOUSTIC FEEDBACK MODELLING AND CANCELLATION SYSTEMS

The modelling of the acoustic feedback is shown in Figure 1. The source signal  $v(n)$  is the desired sound. The system input signal  $u(n)$  is  $v(n)$  added to the ambient noise  $r(n)$ , i.e.,  $u(n) = v(n) + r(n)$ . For convenience,  $u(n)$  includes the characteristics of the microphone and A/D converter.

The acoustic feedback path models the acoustic coupling between loudspeaker and microphone. For simplicity, it also includes the characteristics of the D/A converter, loudspeaker, microphone and A/D converter. The feedback path may present non-linearities, for example due to loudspeaker saturation, but it is generally assumed to be linear. Hence, the feedback path is represented by the  $N_F$ -order time-varying transfer function

$$F(q, n) = f_0(n) + f_1(n)q^{-1} + \dots + f_{N_F}(n)q^{-N_F}, \quad (1)$$

where  $q$  denotes the discrete-time shift operator [24], i.e.,  $q^{-1}x(n) = x(n-1)$ , or by the impulse response

$$\mathbf{f}(n) = [f_0(n) \ f_1(n) \ \dots \ f_{N_F}(n)]^T. \quad (2)$$

The forward path models the amplification system and any other signal processing device inserted in that part of the signal loop. The forward path may present non-linearities, for example because of frequency compression, but it is usually assumed to be linear. Hence, the forward path is represented by the  $N_G$ -order time-varying transfer function

$$G(q, n) = g_0(n) + g_1(n)q^{-1} + \dots + g_{N_G}(n)q^{-N_G} \quad (3)$$

or by the impulse response

$$\mathbf{g}(n) = [g_0(n) \ g_1(n) \ \dots \ g_{N_G}(n)]^T. \quad (4)$$

The loudspeaker signal  $x(n)$  and the system input signal  $u(n)$  are related by the time-varying closed-loop transfer function of the SR system as follows

$$x(n) = \frac{G(q, n)}{1 - G(q, n)F(q, n)}u(n) = A(q, n)u(n), \quad (5)$$

where the filtering of  $x(n)$  with  $F(q, n)$  is given by

$$F(q, n)x(n) = \mathbf{f}(n) * x(n) = \sum_{k=0}^{N_F} f_k(n)x(n-k) \quad (6)$$

and the symbol  $*$  denotes the convolution operation.

From the control systems theory, the Nyquist's stability criterion states that the closed-loop system is unstable if there is at least one frequency  $\omega$  for which [1]

$$\begin{cases} |G(e^{j\omega}, n)F(e^{j\omega}, n)| \geq 1 \\ \angle G(e^{j\omega}, n)F(e^{j\omega}, n) = 2k\pi, k \in \mathbb{Z}, \end{cases} \quad (7)$$

where  $G(e^{j\omega}, n)$  and  $F(e^{j\omega}, n)$  are the short-term frequency responses of the forward and feedback paths, respectively, and  $\omega \in [0, 2\pi]$  is the normalized angular frequency.

Therefore, if one frequency component is not attenuated and its phase is shifted by an integer multiple of  $2\pi$  by the time-varying open-loop transfer function of the SR system,  $G(q, n)F(q, n)$ , then this frequency component will never disappear from the system even if there is no more input signal  $u(n)$ . If it is amplified, its magnitude will increase after each loop in the system, leading to a howling at that frequency.

With the aim of quantifying the achievable amplification in a SR system, it is customary to define a broadband gain of the forward path as [1], [2]

$$K(n) = \frac{1}{2\pi} \int_0^{2\pi} |G(e^{j\omega}, n)| d\omega \quad (8)$$

and to extract it from  $G(q, n)$  as follows

$$G(q, n) = K(n)J(q, n). \quad (9)$$

Assuming that  $J(q, n)$  is known and  $K(n)$  can be varied, the maximum stable gain (MSG) of the SR system is defined as [1]

$$\begin{aligned} \text{MSG}(n)(\text{dB}) &= 20 \log_{10} K(n) \\ \text{such that } \max_{\omega \in P(n)} |G(e^{j\omega}, n)F(e^{j\omega}, n)| &= 1, \end{aligned} \quad (10)$$

resulting in [1], [2]

$$\begin{aligned} \text{MSG}(n)(\text{dB}) &= \\ &- 20 \log_{10} \left[ \max_{\omega \in P(n)} |J(e^{j\omega}, n)F(e^{j\omega}, n)| \right], \end{aligned} \quad (11)$$

where  $P(n)$  denotes the set of frequencies at which the phase condition in (7) is satisfied, which are also called the critical frequencies of the SR system.

The cancellation architecture of the AFC approach is depicted in Figure 2 [1], [2]. The adaptive filter, which identifies and tracks the acoustic feedback path  $F(q, n)$ , is represented by the  $N_H$ -order time-varying transfer function

$$H(q, n) = h_0(n) + h_1(n)q^{-1} + \dots + h_{N_H}(n)q^{-N_H} \quad (12)$$

or by the impulse response

$$\mathbf{h}(n) = [h_0(n) \ h_1(n) \ \dots \ h_{N_H}(n)]^T. \quad (13)$$

Then, the feedback signal  $\mathbf{f}(n)*x(n)$  is estimated as  $\mathbf{h}(n)*x(n)$  and subtracted from the microphone signal  $y(n)$ , generating the feedback-compensated or error signal

$$\begin{aligned} e(n) &= y(n) - \mathbf{h}(n) * x(n) \\ &= u(n) + [\mathbf{f}(n) - \mathbf{h}(n)] * x(n), \end{aligned} \quad (14)$$

which is effectively the signal fed to the forward path  $G(q, n)$ .

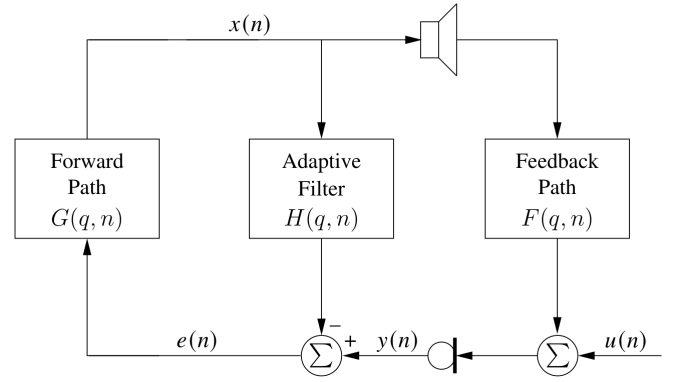


Fig. 2: Structure for acoustic feedback cancellation.

The signals  $x(n)$  and  $u(n)$  are now related by the closed-loop transfer function of the SR system with a AFC method, hereinafter called AFC system, as follows

$$x(n) = \frac{G(q, n)}{1 - G(q, n) [F(q, n) - H(q, n)]} u(n). \quad (15)$$

According to the Nyquist's stability criterion, the closed-loop AFC system is unstable if there is at least one frequency  $\omega$  for which

$$\begin{cases} |G(e^{j\omega}, n) [F(e^{j\omega}, n) - H(e^{j\omega}, n)]| \geq 1 \\ \angle G(e^{j\omega}, n) [F(e^{j\omega}, n) - H(e^{j\omega}, n)] = 2k\pi, k \in \mathbb{Z}, \end{cases} \quad (16)$$

where  $H(e^{j\omega}, n)$  is the short-term frequency response of the adaptive filter.

The MSG of the AFC system is defined as [1], [2]

$$\begin{aligned} \text{MSG}(n)(\text{dB}) &= -20 \log_{10} \\ &\left[ \max_{\omega \in P_H(n)} |J(e^{j\omega}, n) [F(e^{j\omega}, n) - H(e^{j\omega}, n)]| \right] \end{aligned} \quad (17)$$

and the increase in the amount of achievable amplification, the MSG increase, is given by [1], [2]

$$\begin{aligned} \Delta \text{MSG}(n)(\text{dB}) &= -20 \log_{10} \\ &\left[ \frac{\max_{\omega \in P_H(n)} |J(e^{j\omega}, n) [F(e^{j\omega}, n) - H(e^{j\omega}, n)]|}{\max_{\omega \in P(n)} |J(e^{j\omega}, n)F(e^{j\omega}, n)|} \right], \end{aligned} \quad (18)$$

where  $P_H(n)$  denotes the set of frequencies at which the phase condition in (16) is met, which are the critical frequencies of the AFC system.

It is concluded from (18) that the achievable  $\Delta \text{MSG}(n)$  increases as the match between the frequency response magnitudes of the adaptive filter and feedback path at the critical frequencies of the AFC system gets better. If  $H(e^{j\omega}, n) = F(e^{j\omega}, n)$ ,  $\forall \omega \in P_H(n)$ , the MSG of the AFC system can be indefinitely high. But some reverberation may still exist in  $e(n)$  due to the frequency components that were not perfectly matched. If  $H(q, n) = F(q, n)$ , it follows from (14) and (15) that the acoustic feedback will be totally cancelled and the system will no longer have a closed signal loop, respectively.

### III. LEAST SQUARES ANALYSIS OF AFC SYSTEMS

This section presents, using the LS estimation theory, a theoretical analysis on the poor performance of standard adaptive filtering algorithms when applied to AFC and compiles in detail the results already presented in the literature. Thereunto, let a data record  $\{x(k), y(k)\}_{k=1}^n$  of the loudspeaker and microphone signals be available as well as the initial conditions  $\{x(k)\}_{k=1-N_H}^0$  of the loudspeaker signal. And let the feedback path be time-invariant ( $\mathbf{f}(n) = \mathbf{f}$ ) so that no data windowing is employed and the order of the feedback path estimator  $\mathbf{h}(n)$  be equal to that of  $\mathbf{f}$  ( $N_H = N_F$ ).

In the LS approach, an estimate of  $\mathbf{f}$  at the sample index  $n$ ,  $\mathbf{h}(n)$ , is obtained by minimizing the cost function or error criterion<sup>1</sup> defined as [25]

$$\begin{aligned} J_n(\mathbf{h}) &= \frac{1}{2} \sum_{k=1}^n e^2(k) \\ &= \frac{1}{2} \sum_{k=1}^n [y(k) - \mathbf{x}^T(k)\mathbf{h}(n)]^2 \\ &= \frac{1}{2} [\mathbf{y} - \mathbf{X}\mathbf{h}(n)]^T [\mathbf{y} - \mathbf{X}\mathbf{h}(n)], \end{aligned} \quad (19)$$

where the data matrices and vectors are defined as

$$\mathbf{y} = [y(n) \ y(n-1) \ \dots \ y(1)]^T \quad (20)$$

$$\mathbf{X} = [\mathbf{x}(n) \ \mathbf{x}(n-1) \ \dots \ \mathbf{x}(1)]^T \quad (21)$$

$$\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-N_H)]^T. \quad (22)$$

The LS cost function, defined in (19), can be written as

$$J_n(\mathbf{h}) = \frac{1}{2} [\mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\mathbf{h}(n) + \mathbf{h}^T(n) \mathbf{X}^T \mathbf{X}\mathbf{h}(n)] \quad (23)$$

and thus its gradient is given by

$$J'_n(\mathbf{h}) = -\mathbf{X}^T \mathbf{y} + \mathbf{X}^T \mathbf{X}\mathbf{h}(n). \quad (24)$$

The LS estimator of the feedback path is then obtained by making the gradient equal to zero, resulting in [1]

$$\begin{aligned} \mathbf{h}(n) &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{u} + \mathbf{X}\mathbf{f}) \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} + \mathbf{f}, \end{aligned} \quad (25)$$

where

$$\mathbf{u} = [u(n) \ u(n-1) \ \dots \ u(1)]^T. \quad (26)$$

An estimator may be characterized by its bias and variance [1], [25]. In the sequel, the bias and variance of the LS estimator defined in (25) is obtained and discussed separately.

#### A. Bias of the LS Estimator

The bias of an estimator is the difference between its expected value and the parameter true value [25]. Hence, at index  $n$ , the bias of the feedback path estimator is defined as [1]

$$\text{bias}\{\mathbf{h}(n)\} = E\{\mathbf{h}(n)\} - \mathbf{f}, \quad (27)$$

<sup>1</sup>The error criterion in the LS approach is commonly defined as  $J_n(\mathbf{h}) = \sum_{k=1}^n e^2(k)$  [25]. In this work, a constant  $1/2$  is included only for convenience.

where  $E\{\cdot\}$  denotes the statistical expectation operator. The estimator is desired to be unbiased, i.e., to have zero bias [25].

Replacing (25) in (27), the bias of the LS estimator of the feedback path is given by [1], [14]

$$\begin{aligned} \text{bias}\{\mathbf{h}(n)\} &= E\left\{(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} + \mathbf{f}\right\} - \mathbf{f} \\ &= E\left\{(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}\right\}. \end{aligned} \quad (28)$$

In order to conclude on the bias, it is necessary to realize that (28) can be written as

$$\text{bias}\{\mathbf{h}(n)\} = E\left\{\bar{\mathbf{R}}_x^{-1} \bar{\mathbf{p}}_{xu}\right\}, \quad (29)$$

where

$$\bar{\mathbf{R}}_x = \frac{1}{n} \mathbf{X}^T \mathbf{X} \quad (30)$$

is the  $(N_H + 1) \times (N_H + 1)$  time-average autocorrelation matrix of  $x(n)$  [11] and

$$\bar{\mathbf{p}}_{xu} = \frac{1}{n} \mathbf{X}^T \mathbf{u} \quad (31)$$

is the  $(N_H + 1) \times (N_H + 1)$  time-average cross-correlation vector between  $x(n)$  and  $u(n)$  [11].

If  $x(n)$  and  $u(n)$  are at least jointly wide-sense stationary and ergodic processes then, for large  $n$ , the time averages  $\bar{\mathbf{R}}_x$  and  $\bar{\mathbf{p}}_{xu}$  are consistent estimates of the respective statistical averages [26], i.e.,

$$\lim_{n \rightarrow \infty} \bar{\mathbf{R}}_x = \mathbf{R}_x \quad (32)$$

and

$$\lim_{n \rightarrow \infty} \bar{\mathbf{p}}_{xu} = \mathbf{p}_{xu}, \quad (33)$$

where  $\mathbf{R}_x$  is the  $(N_H + 1) \times (N_H + 1)$  autocorrelation matrix of  $x(n)$  and  $\mathbf{p}_{xu}$  is the  $(N_H + 1) \times 1$  cross-correlation vector between  $x(n)$  and  $u(n)$  [11].

In this case, by replacing (32) and (33) in (29), the bias of the feedback path estimator is defined as [21], [22], [4]

$$\text{bias}\{\mathbf{h}(n)\} = \mathbf{R}_x^{-1} \mathbf{p}_{xu}. \quad (34)$$

Since autocorrelation matrices are positive definite for practical signals [11],  $\mathbf{R}_x^{-1}$  is positive definite and thus the only possibility for the bias defined in (34) to be null is that  $\mathbf{p}_{xu} = \mathbf{0}$ , i.e., the loudspeaker signal  $x(n)$  and the system input signal  $u(n)$  must be orthogonal. If at least one of them has zero mean, it is equivalent to say that they must be uncorrelated. However, as indicated in (5),  $x(n)$  and  $u(n)$  are related by the closed-loop transfer function  $A(q, n)$  of the SR system, which generally introduces linear dependence. As will be discussed later, a counterexample to this correlation introduced by the system closed loop is when  $u(n)$  is white Gaussian noise and  $A(q, n)$  has at least 1-sample delay.

Disregarding the assumptions of stationarity and ergodicity of  $x(n)$  and  $u(n)$ , similar conclusion can be drawn directly from (29). The bias can be generally understood to be nonzero because the closed-loop nature of the system tends to cause  $\bar{\mathbf{p}}_{xu} \neq \mathbf{0}$ . Therefore, in general,  $\text{bias}\{\mathbf{h}(n)\} \neq \mathbf{0}$ . And it actually depends on the autocorrelation matrix of the loudspeaker signal  $x(n)$  and the cross-correlation vector between  $x(n)$  and the system input signal  $u(n)$ .

The existence of bias means that, on average, the feedback path estimator  $\mathbf{h}(n)$  will not converge to  $\mathbf{f}$ , the true value of the feedback path, no matter how many process realizations are performed and their time duration [25]. The resulting effect in AFC is twofold: first, the adaptive filter estimates and cancels only part of the feedback signal  $\mathbf{f} * x(n)$ ; second, it also estimates and cancels part of the system input signal  $u(n)$ . As a consequence, the feedback-compensated signal  $e(n)$  is a distorted estimate of the system input signal  $u(n)$  [1], [2].

In addition to the definitions in (28), (29) and (34), a different expression for the bias can be obtained by realizing that

$$\bar{\mathbf{R}}_x = \frac{1}{n} J_n''(\cdot) \quad (35)$$

and

$$\bar{\mathbf{p}}_{xu} = -\frac{1}{n} J_n'(\mathbf{f}), \quad (36)$$

where  $J_n''(\cdot)$  is the second derivative of the error criterion as a function of  $\mathbf{h}(n)$ . Replacing (35) and (36) in (29), the bias can be defined as a function of the error criterion as follows

$$\text{bias}\{\mathbf{h}(n)\} = -E\left\{[J_n''(\cdot)]^{-1} J_n'(\mathbf{f})\right\}. \quad (37)$$

### B. Variance of the LS Estimator

The variance of the LS estimator can be obtained by considering its covariance matrix, also called coefficient-error-vector covariance matrix, which is defined as [11], [10]

$$\text{cov}\{\mathbf{h}(n)\} = E\{[\mathbf{h}(n) - \mathbf{f}][\mathbf{h}(n) - \mathbf{f}]^T\}. \quad (38)$$

Replacing (25) in (38), the covariance matrix is given by

$$\text{cov}\{\mathbf{h}(n)\} = E\left\{\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{u} \left[\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{u}\right]^T\right\}, \quad (39)$$

which using (30) and (31) becomes

$$\text{cov}\{\mathbf{h}(n)\} = E\left\{\bar{\mathbf{R}}_x^{-1} \bar{\mathbf{p}}_{xu} \left[\bar{\mathbf{R}}_x^{-1} \bar{\mathbf{p}}_{xu}\right]^T\right\}. \quad (40)$$

Assuming that  $x(n)$  and  $u(n)$  are at least jointly wide-sense stationary and ergodic processes and thus replacing (32) and (33) in (40), the covariance matrix is defined as

$$\begin{aligned} \text{cov}\{\mathbf{h}(n)\} &= \mathbf{R}_x^{-1} \mathbf{p}_{xu} \left[\mathbf{R}_x^{-1} \mathbf{p}_{xu}\right]^T \\ &= \text{bias}\{\mathbf{h}(n)\} [\text{bias}\{\mathbf{h}(n)\}]^T. \end{aligned} \quad (41)$$

The only possibility for the covariance matrix, defined in (41), to be null is that  $\text{bias}\{\mathbf{h}(n)\} = \mathbf{0}$ , i.e., the estimator  $\mathbf{h}(n)$  must be unbiased. However, as discussed in the previous subsection, this does not normally occur because the closed-loop transfer function  $A(q, n)$  of the SR system makes  $\mathbf{p}_{xu} \neq \mathbf{0}$ , which results in a biased LS estimator of the feedback path.

Disregarding the assumptions of stationarity and ergodicity of  $x(n)$  and  $u(n)$ , similar conclusion can be drawn directly from (40). The covariance matrix can be generally understood to be nonzero because the closed-loop nature of the system tends to cause  $\bar{\mathbf{p}}_{xu} \neq \mathbf{0}$ . Therefore, in general,  $\text{cov}\{\mathbf{h}(n)\} \neq \mathbf{0}$ . And it actually depends on the autocorrelation matrix of the loudspeaker signal  $x(n)$  and the cross-correlation vector between  $x(n)$  and the system input signal  $u(n)$ .

The resulting effect of the covariance matrix in AFC can be understood by realizing that (39) can be written as

$$\text{cov}\{\mathbf{h}(n)\} = E\left\{\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \hat{\mathbf{R}}_u \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right\} \quad (42)$$

or [1], [13]

$$\text{cov}\{\mathbf{h}(n)\} = \left[E\{\mathbf{X}^T \hat{\mathbf{R}}_u^{-1} \mathbf{X}\}\right]^{-1}, \quad (43)$$

where

$$\mathbf{R}_u = E\{\mathbf{u}\mathbf{u}^T\} \quad (44)$$

is the  $n \times n$  autocorrelation matrix of  $u(n)$  [11], [10] and

$$\hat{\mathbf{R}}_u = \mathbf{u}\mathbf{u}^T \quad (45)$$

is its instantaneous estimate [10].

The interpretation of (42) and (43) may be related to the double-talk problem in AEC [1]. In AEC, the signals  $x(n)$  and  $u(n)$ , which are respectively called far-end speaker signal and near-end speaker signal, are independent. Then, when  $x(n)$  is active and  $u(n)$  is not, the covariance matrix of the echo path LS estimator is relatively small because  $\hat{\mathbf{R}}_u \approx \mathbf{R}_u \approx \mathbf{0}$  and thus the adaptive filter works properly. But when both signals are active, situation commonly called double-talk, the covariance matrix can become large because  $\hat{\mathbf{R}}_u \neq \mathbf{0}$  and  $\mathbf{R}_u \neq \mathbf{0}$ , which consequently lead to a decrease in the convergence speed, or even divergence, of the adaptive filter [1]. This problem becomes more severe when  $u(n)$  has a high degree of spectral coloration, as occurs when  $v(n)$  is speech, because  $\mathbf{R}_u$  presents a denser structure in this case [1].

In AFC, on the other hand, the loudspeaker signal  $x(n)$  and the system input signal  $u(n)$  are not independent because they are related by the system closed-loop as indicated in (5). Therefore, except for possible short time intervals dependent on  $F(q, n)$  and  $G(q, n)$ , the system operates in a continuous double-talk situation. And this is further worsened by the aforementioned correlation between  $u(n)$  and  $x(n)$  [1]. The resulting effect in AFC is that the adaptive filter  $H(q, n)$  presents a slow convergence speed throughout its operation [1].

In addition to the previous definitions, the covariance matrix can be also defined as a function of the error criterion by replacing (35) and (36) in (40), resulting in

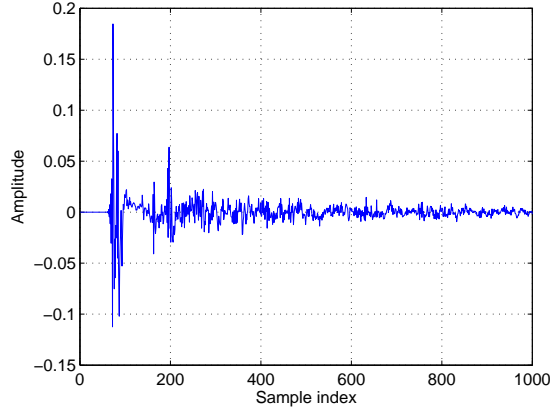
$$\text{cov}\{\mathbf{h}(n)\} = E\left\{[J_n''(\cdot)]^{-1} J_n'(\mathbf{f}) [J_n'(\mathbf{f})]^T [J_n''(\cdot)]^{-1}\right\}. \quad (46)$$

## IV. SIMULATIONS CONFIGURATION

This section describes the configuration of the two experiments carried out in a simulated environment to corroborate and exemplify the conclusions on the bias and covariance matrix of the LS estimator. In the first, the bias is estimated over time. In the second, the slow convergence speed of the standard adaptive filtering algorithms when applied to AFC is exemplified using the RLS algorithm. To this end, the following configuration is used.

### A. Simulated Environment

1) *Feedback Path*: The impulse response  $\mathbf{f}$  of the feedback path is a measured room impulse response (RIR) available in [27]. The RIR was downsampled to the sampling rate  $f_s = 16$  kHz and then truncated so that  $N_F = 1000$ . The impulse response of the acoustic feedback path is shown in Figure 3.


 Fig. 3: Impulse response  $\mathbf{f}$  of the feedback path.

2) *Forward Path*: As in [1], [2], the forward path is a time-invariant filter defined as a delay and a gain, i.e.,

$$G(q) = g_{N_G} q^{-N_G}, \quad (47)$$

which leads to  $K(n) = g_{N_G}$  and  $J(q, n) = q^{-N_G}$ . The delay is inherent to any digital signal processing included in that part of the signal loop and plays a key role in the analysis. Because of that,  $N_G = \{1, 100, 400, 800\}$ . And  $g_{N_G} = 0.7$  so that the initial gain margin, the difference between the MSG of the SR system and the actual broadband gain, is 3 dB as in [1], [2].

3) *Closed-loop System*: According to (5), the closed-loop transfer function of the SR system is defined as

$$A(q, n) = \frac{G(q, n)}{1 - G(q, n)F(q, n)}. \quad (48)$$

If  $|G(e^{j\omega}, n)F(e^{j\omega}, n)| < 1$ , a sufficient condition to ensure the closed-loop system stability, then  $A(q, n)$  can be written as

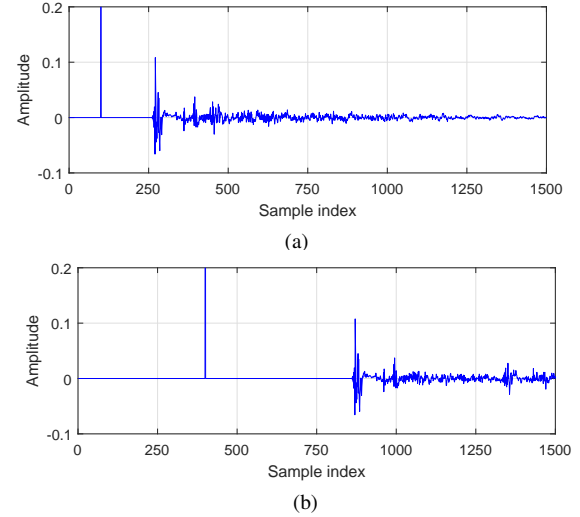
$$A(q, n) = G(q, n) \sum_{k=0}^{\infty} [G(q, n)F(q, n)]^k. \quad (49)$$

It can be concluded that the system closed-loop transfer function is the forward path transfer function multiplied by a power sum of the system open-loop transfer function. Each term on (49) can be interpreted as the transfer function from  $u(n)$  to  $x(n)$  after  $k$  loops through the SR system.

By replacing (47) in (49), the closed-loop transfer function of the simulated SR system is given by

$$A(q) = \sum_{k=0}^{\infty} g_{N_G}^{k+1} q^{-N_G(k+1)} F^k(q). \quad (50)$$

In the time domain, except for a constant, the first two terms ( $k = 0, 1$ ) on (50) correspond to an impulse at the sample index  $N_G$  (the forward path impulse response) and the feedback path impulse response delayed by  $2N_G$  samples, respectively. The other terms ( $k > 1$ ) correspond to  $k$  successive convolutions of the feedback path impulse response delayed by  $(k+1)N_G$  samples. Thus, as  $N_G$  increases, this linear delay as a function of  $N_G$  causes an increasing dilatation of the system closed-loop impulse response  $\mathbf{a}$  along the sample axis. This effect is illustrated in Figure 4 for  $N_G = \{100, 400\}$ , where it should be noted that  $a_k = 0$  for  $k < N_G$ .


 Fig. 4: Impulse response  $\mathbf{a}$  of the closed-loop system for: (a)  $N_G = 100$ ; (b)  $N_G = 400$ .

## B. Evaluation Metrics

1) *Bias energy*: In the first experiment, the bias of the LS estimator is measured through its energy, which is defined as

$$E(n) = \|\text{bias}\{\mathbf{h}(n)\}\|^2 \quad (51)$$

where  $\|\cdot\|$  denotes the Euclidean or Frobenius norm.

2) *Misalignment*: In the second experiment, the performance of the adaptive filter is evaluated through the normalized misalignment (MIS), which is defined as [2]

$$\text{MIS}(n) = \frac{\|\mathbf{f}(n) - \mathbf{h}(n)\|}{\|\mathbf{f}(n)\|} = \frac{\|F(e^{j\omega}, n) - H(e^{j\omega}, n)\|}{\|F(e^{j\omega}, n)\|}. \quad (52)$$

3) *Maximum stable gain*: In the second experiment, the performance of the adaptive filter is also evaluated through the achievable  $\Delta\text{MSG}(n)$  defined in (18). It is noteworthy that the metrics  $\text{MIS}(n)$  and  $\Delta\text{MSG}(n)$  are related because both depend on  $|F(e^{j\omega}, n) - H(e^{j\omega}, n)|$ . With respect to this factor, the difference is that  $\text{MIS}(n)$  takes all frequencies into account while  $\Delta\text{MSG}(n)$  considers only one of the critical frequencies.

## C. Speech Signals

The source signals  $v(n)$  are created from signals of a speech database. Each signal contains one short sentence recorded in a 4 s time slot and at a 48 kHz sampling rate, but downsampled to 16 kHz. The active power level of each signal is normalized to  $-26$  dBov according to the ITU-T Rec. P.56 algorithm. All sentences are spoken by native speakers of the following nationalities and genders: 4 Americans (2 males and 2 females), 2 British (1 male and 1 female), 2 French (1 male and 1 female) and 2 Germans (1 male and 1 female).

Since long-term signals are required, several signals from the same speaker are concatenated and the silence parts are removed through a voice activity detector (VAD), thereby resulting in 10 speech signals (1 signal per speaker).

## V. SIMULATIONS RESULTS

This section is devoted to reporting the results of the two experiments performed. The simulated environment, evaluation metrics and signals described in Section IV are employed.

### A. Experiment 1

In this experiment, the bias  $\{\mathbf{h}(n)\}$  is computed over time using (28) for several signal-to-noise ratios (SNR). The ambient noise  $r(n)$  is a white Gaussian noise with zero mean and variance  $\sigma^2$  such that  $\text{SNR} = \{\infty, 30, 20, 10, -\infty\}$  dB, where  $\text{SNR} = -\infty$  indicates that  $v(n) = 0$  and  $\sigma = 1$ .

The bias energy  $E(n)$  is measured according to (51). In order to simplify the analysis and discussion of the results, it is assumed that  $x(n)$  and  $u(n)$  are stationary and ergodic processes as well as  $n$  is large enough so that  $\bar{\mathbf{R}}_x \approx \mathbf{R}_x$  and  $\bar{\mathbf{p}}_{xu} \approx \mathbf{p}_{xu}$ , thereby making (34) valid. In this case,  $E(n)$  is bounded from above as follows [28]

$$E(n) \leq \|\mathbf{R}_x^{-1}\|^2 \|\mathbf{p}_{xu}\|^2. \quad (53)$$

As demonstrated in Appendices A and B,  $\mathbf{R}_x$  and  $\mathbf{p}_{xu}$  can be written as

$$\mathbf{R}_x = \sum_{k=-\infty}^{\infty} \tilde{r}_a(k) \mathbf{R}_v(k) + \sigma^2 \tilde{\mathbf{R}}_a \quad (54)$$

and

$$\mathbf{p}_{xu} = \sum_{k=0}^{\infty} a_k \mathbf{p}_v(k) + \sigma^2 [a_0 \ 0 \ \dots \ 0]^T, \quad (55)$$

where  $\mathbf{p}_v(k)$ ,  $\tilde{r}_a(k)$ ,  $\mathbf{R}_v(k)$  and  $\tilde{\mathbf{R}}_a$  are defined in (73), (79), (85) and (86), respectively. As  $a_k = 0, k < N_G$ , (55) becomes

$$\mathbf{p}_{xu} = \sum_{k=N_G}^{\infty} a_k \mathbf{p}_v(k) \quad (56)$$

and therefore  $\mathbf{p}_{xu}$  is actually independent of  $r(n)$ .

The results for  $\text{SNR} \rightarrow \infty$  and  $N_G = \{1, 100, 400, 800\}$  are shown in Figure 5. It is observed that  $E(n)$  decreases exponentially over time, becoming nearly constant at the end of the simulation. This occurs because the time averages  $\bar{\mathbf{R}}_x$  and  $\bar{\mathbf{p}}_{xu}$  are not sensitive to any changes in the signal statistics for large  $n$  [10]. But  $E(n) \neq 0$  even after 10 s, which illustrates the bias of the acoustic feedback path estimate provided by the standard adaptive filtering algorithms.

It is also noticed that  $E(n)$  decreases as  $N_G$  increases. This can be explained by noting that  $\|\mathbf{p}_{xu}\|^2$  is bounded from above as follows

$$\|\mathbf{p}_{xu}\|^2 \leq \sum_{k=N_G}^{\infty} |a_k|^2 \|\mathbf{p}_v(k)\|^2. \quad (57)$$

As  $N_G$  increases, fewer correlation vectors are included in right-hand side of (57), decreasing the upper bound of  $\|\mathbf{p}_{xu}\|^2$ . And, since the autocorrelation function of speech signals usually decays with increasing lag, the higher magnitude vectors  $\mathbf{p}_v(k)$  are removed from this computation and consequently the upper bound of  $\|\mathbf{p}_{xu}\|^2$  tends to significantly decrease. On the other hand, as  $N_G$  increases,  $\hat{r}_a(k)$  presents an increasing dilatation along the lags  $k$  caused by the similar dilatation of  $\mathbf{a}$  along the sample axis, as discussed in Section IV-A3.

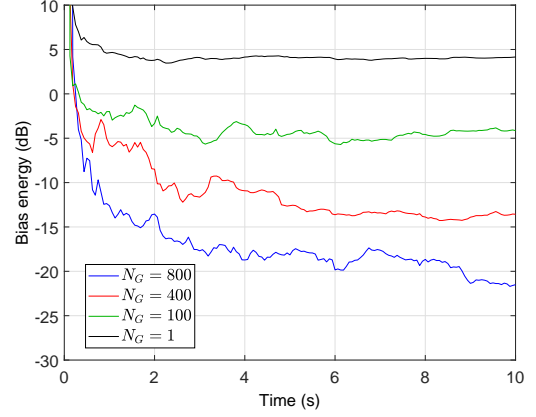


Fig. 5: Bias energy for speech signals,  $\text{SNR} = \infty$  and several values of  $N_G$ .

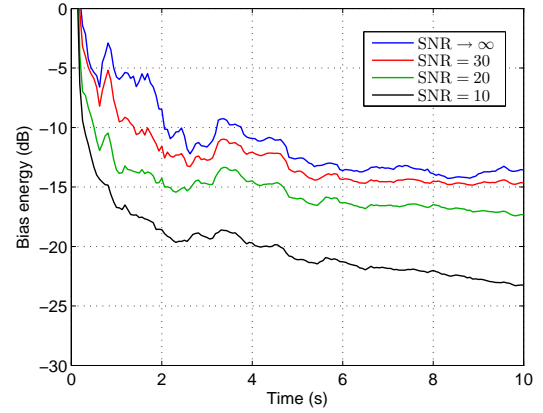


Fig. 6: Bias energy for speech signals,  $N_G = 400$  and several values of SNR.

This effect in  $\mathbf{R}_x$ , defined in (54), is not significant because the aforementioned decreasing behavior of the speech signal autocorrelation function makes  $\mathbf{R}_v(k) \approx \mathbf{0}_{(N_H+1) \times (N_H+1)}$  for high values of  $|k|$ . Therefore, as  $N_G$  increases,  $\|\mathbf{p}_{xu}\|^2$  tends to decrease while  $\|\mathbf{R}_x^{-1}\|^2$  is not significantly affected, leading to a decrease in the upper bound of  $E(n)$ . In practise, although not mathematically guaranteed, the bias energy ultimately decreases as indicated by the obtained results.

In addition, the results for  $N_G = 400$  and  $\text{SNR} = \{\infty, 30, 20, 10\}$  dB are shown in Figure 6. It is observed that  $E(n)$  decreases as SNR decreases. This is justified by observing that  $\mathbf{p}_{xu}$ , defined in (56), is not affected by  $r(n)$  while  $\mathbf{R}_x$ , defined in (54), is affected by its power ( $\sigma$ ). From (54), it is noted that  $\|\mathbf{R}_x\|^2$  is bounded from above as follows

$$\|\mathbf{R}_x\|^2 \leq \sum_{k=-\infty}^{\infty} |\tilde{r}_a(k)|^2 \|\mathbf{R}_v(k)\|^2 + \sigma^2 \|\tilde{\mathbf{R}}_a\|^2 \quad (58)$$

while  $\|\mathbf{R}_x^{-1}\|^2$  has the following lower bound [28]

$$\|\mathbf{R}_x^{-1}\|^2 \geq \frac{\|I\|^2}{\|\mathbf{R}_x\|^2} = \frac{N_H}{\|\mathbf{R}_x\|^2}. \quad (59)$$

As  $\sigma$  increases (SNR decreases), the upper bound of  $\|\mathbf{R}_x\|^2$  increases and consequently the lower bound of  $\|\mathbf{R}_x^{-1}\|^2$  de-



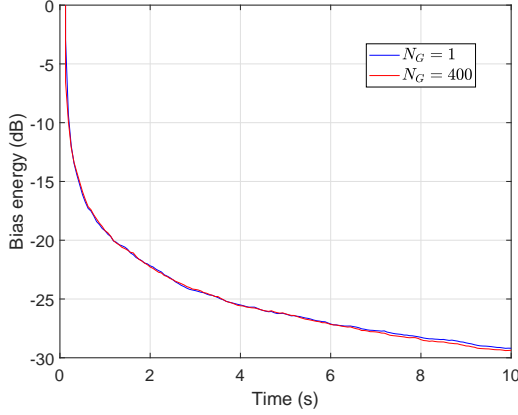


Fig. 7: Bias energy for white Gaussian noise and  $N_G = 1, 400$ .

creases. Therefore, as SNR decreases,  $\|\mathbf{p}_{xu}\|^2$  is not affected while  $\|\mathbf{R}_x^{-1}\|^2$  tends to decrease, leading to a decrease in the upper bound of the bias energy defined in (53). Again in practise, although not mathematically guaranteed, the bias energy ultimately decreases as indicated by the obtained results.

Finally, the results for white Gaussian noise ( $\text{SNR} \rightarrow -\infty$ ) and  $N_G = \{1, 400\}$  are shown in Figure 7. It is observed that the values of  $E(n)$  are practically the same. This occurs because, as can be concluded from (56),  $v(n) = 0$  combined with  $a_0 = 0$  makes  $\mathbf{p}_{xu} = \mathbf{0}$ , i.e., the loudspeaker signal  $x(n)$  and the system input signal  $u(n)$  are uncorrelated. As a consequence, bias  $\{\mathbf{h}(n)\} = \mathbf{0}$ . It is noteworthy that, although it is much smaller than for speech signals, the bias energy is still nonzero after 10s because  $r(n)$  is indeed sequences of pseudorandom values drawn from the standard normal distribution and, consequently,  $\bar{\mathbf{R}}_x$  and  $\bar{\mathbf{p}}_{xu}$  are close but not equal to the statistical  $\mathbf{R}_x$  and  $\mathbf{p}_{xu}$  even for large  $n$ .

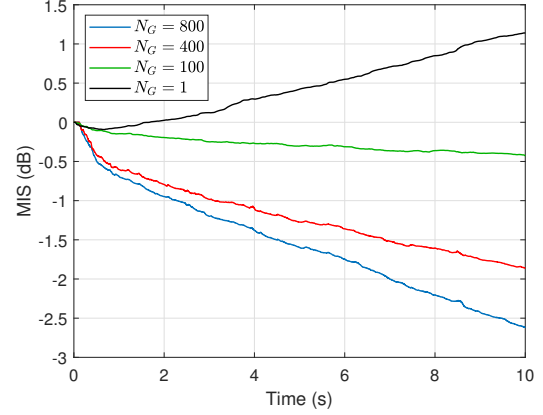
The presented results exemplify and corroborate the theoretical discussion on the existence of bias in the feedback path estimate provided by the standard adaptive filtering algorithms and its direct relationship with the autocorrelation matrix of the loudspeaker signal  $x(n)$  and the cross-correlation between  $x(n)$  and the system input signal  $u(n)$ .

### B. Experiment 2

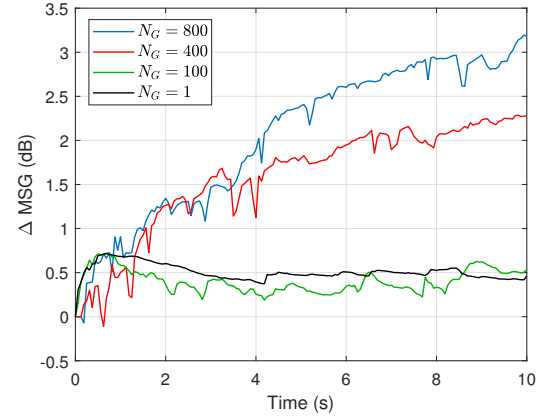
In this experiment, the slow convergence speed of the standard adaptive filtering algorithms when applied to AFC is exemplified using the RLS algorithm,  $N_H = N_F$  and  $\text{SNR} = \{\infty, 30, 20, 10, -\infty\}$ . The RLS algorithm is initialized with  $\mathbf{h}(0) = \mathbf{0}$  and  $\bar{\mathbf{R}}_x^{-1} = \delta \mathbf{I}$ . For each signal, the algorithm is run with  $\delta = 10^{-\alpha}$  and forgetting factor  $\lambda = 1 - 10^{-\beta}$ , where  $\alpha = \{2, 3, 4\}$  and  $\beta = \{0, 2, 3, 4, 5, 6\}$ , and the result that minimize instability time and, secondarily, MIS( $n$ ) average is selected.

As did in Experiment 1, in order to simplify the analysis and discussion of the results, it is assumed that  $x(n)$  and  $u(n)$  are stationary and ergodic processes as well as  $n$  is large enough so that  $\bar{\mathbf{R}}_x \approx \mathbf{R}_x$  and  $\bar{\mathbf{p}}_{xu} \approx \mathbf{p}_{xu}$ , thereby making (37) valid. In this case, the norm of the covariance matrix is given by

$$\begin{aligned} \|\text{cov}\{\mathbf{h}(n)\}\| &= \|\text{bias}\{\mathbf{h}(n)\} [\text{bias}\{\mathbf{h}(n)\}]^T\| \\ &= \|\text{bias}\{\mathbf{h}(n)\}\|^2 = E(n), \end{aligned} \quad (60)$$



(a)



(b)

Fig. 8: Average results of the RLS algorithm for speech signals,  $\text{SNR} \rightarrow \infty$  and several values of  $N_G$ : (a) MIS( $n$ ); (b)  $\Delta \text{MSG}(n)$ .

being equal to the bias energy.

The average results for  $\text{SNR} \rightarrow \infty$  and  $N_G = \{1, 100, 400, 800\}$  are shown in Figure 8. It should be observed from Figure 8a that the adaptive filter has not yet converged even after 10s in all cases. For  $N_G = 1$ , the adaptive filter actually diverged. But, due to the high time required for filter convergence and computational cost of the RLS algorithm, the discussion on the improvement in the convergence speed will be addressed by means of the illustrated transient behavior.

From Figure 8, it is observed that performance improves as  $N_G$  increases. This occurs because, when  $u(n)$  is speech, increasing  $N_G$  decreases the bias energy, as discussed in the previous subsection, and consequently  $\|\text{cov}\{\mathbf{h}(n)\}\|$ , leading to an increase in the convergence speed of the adaptive filter. The lack of improvement from  $N_G = 1$  to  $N_G = 100$  in Figure 8b is due to  $\Delta \text{MSG}(n)$  taking only one frequency into consideration as explained in Section IV-B3.

Moreover, the average results for  $N_G = 400$  and  $\text{SNR} = \{\infty, 30, 20, 10\}$  dB are shown in Figure 9. It is noticed that performance improves as SNR decreases. This occurs because the white Gaussian noise  $r(n)$  decreases the bias energy, as discussed in the previous subsection, and consequently  $\|\text{cov}\{\mathbf{h}(n)\}\|$ , thereby leading to an increase in the convergence speed of the adaptive filter. Note that the results for



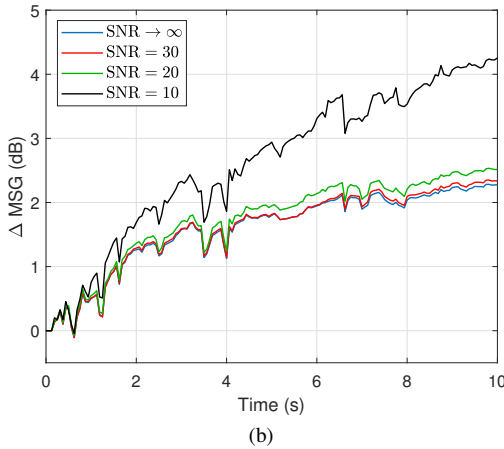
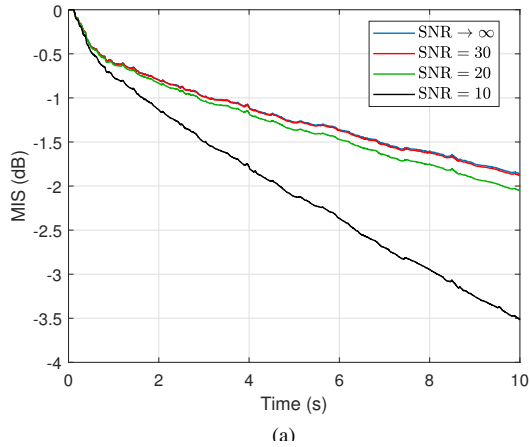


Fig. 9: Average results of the RLS algorithm for speech signals,  $N_G = 400$  and several values of SNR: (a)  $MIS(n)$ ; (b)  $\Delta MSG(n)$ .

SNR = { $\infty$ , 30} dB are very close since the difference between the bias energies is small as indicated in Figure 6.

However, in the cases discussed so far,  $MIS > -3.5$  dB and  $MSG < 4.2$  dB after 10 s. These results exemplify the slow convergence speed of the standard adaptive filtering algorithms when applied to AFC and are explained by the fact that, when  $u(n)$  is speech,  $\mathbf{R}_u$  is such that  $\text{cov}\{\mathbf{h}(n)\}$  is large. These results are even more striking because of the known fast convergence speed of the RLS algorithm when the eigenvalue spread of autocorrelation matrix of the adaptive filter input,  $\mathbf{R}_x$  in this case, is large as occurs for speech signals [10].

Finally, the results for white Gaussian noise (SNR  $\rightarrow -\infty$ ) and  $N_G = \{1, 400\}$  is shown in Figure 10. It is observed that  $MIS(n)$  and  $MSG(n)$  are very similar over time. This occurs because, as discussed in the first experiment,  $a_0 = 0$  is, in theory, sufficient to completely decorrelate  $x(n)$  and  $u(n)$  when the latter is white Gaussian noise, leading to bias  $\{\mathbf{h}(n)\} = \mathbf{0}$  and thus  $\text{cov}\{\mathbf{h}(n)\} = \mathbf{0}$  in both cases. Because of that, the performance of the adaptive filter is much better than for speech signals, achieving  $MIS \approx -12$  dB and  $MSG \approx 12$  dB after 10 s. However, despite the improvement in convergence speed, it is noteworthy the adaptive filter has not yet converged because  $r(n)$  is indeed sequences of pseudorandom values drawn from the standard normal distribution.

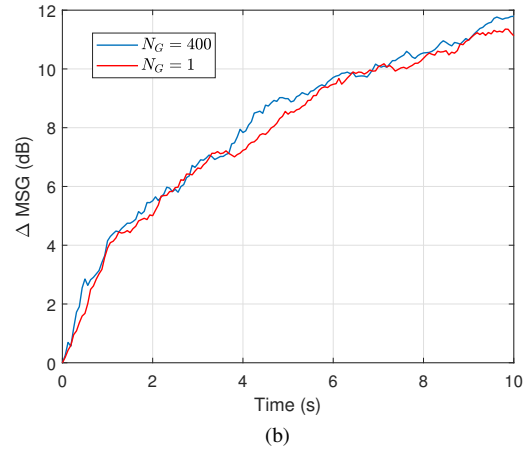
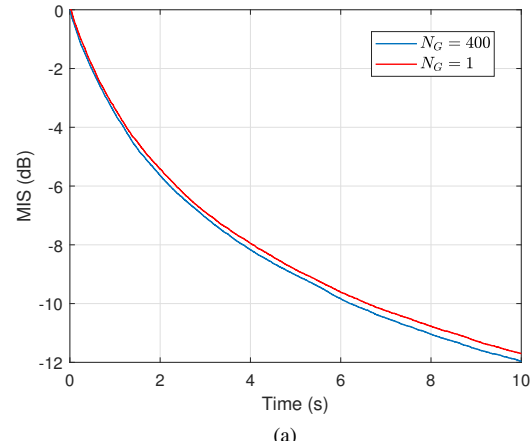


Fig. 10: Average results of the RLS algorithm for white Gaussian noise and  $N_G = \{1, 400\}$ : (a)  $MIS(n)$ ; (b)  $\Delta MSG(n)$ .

This causes bias  $\{\mathbf{h}(n)\} \neq \mathbf{0}$ , as can be seen in Figure 7, and thus  $\text{cov}\{\mathbf{h}(n)\} \neq \mathbf{0}$ , thereby limiting the convergence speed of the adaptive filter.

As a matter of fact, when  $a_0 = 0$  and  $u(n)$  is white Gaussian noise with zero mean and variance  $\sigma^2$ , the signals  $x(n)$  and  $u(n)$  are independent and thus the covariance matrix, defined in (42), can be written as [10]

$$\begin{aligned} \text{cov}\{\mathbf{h}(n)\} &= \mathbb{E}\left\{\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{u} \mathbf{u}^T \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right\} \\ &= \mathbb{E}\left\{\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbb{E}\{\mathbf{u} \mathbf{u}^T\} \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right\} \\ &= \sigma^2 \mathbb{E}\left\{\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right\} \\ &= \sigma^2 \mathbb{E}\left\{\left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right\}, \end{aligned} \quad (61)$$

which, by replacing (30), becomes

$$\text{cov}\{\mathbf{h}(n)\} = \frac{\sigma^2}{n} \mathbb{E}\left\{\overline{\mathbf{R}}_x^{-1}\right\}. \quad (62)$$

The equations (61) and (62) are the Cramer-Rao lower bound (CRLB) of the covariance matrix for any unbiased estimator of  $\mathbf{f}$  [10]. Therefore,  $\mathbf{h}(n)$  is not only the best unbiased linear estimator (BLUE), in the sense that no other unbiased linear

solution generated by any approach has lower variance, but also is the minimum-variance unbiased (MVU) estimator [10].

And considering that  $n$  is large enough so that  $\bar{\mathbf{R}}_x \approx \mathbf{R}_x$ , the covariance matrix defined in (62) can be written as

$$\text{cov}\{\mathbf{h}(n)\} = \frac{\sigma^2}{n} \mathbf{R}_x^{-1}, \quad (63)$$

which, by making use of (54), becomes

$$\text{cov}\{\mathbf{h}(n)\} = \frac{\hat{\mathbf{R}}_a^{-1}}{n}. \quad (64)$$

Therefore, in the problem at hand, the CRLB of the covariance matrix depends solely on the sample autocorrelation matrix of the closed-loop impulse response of the SR system. Moreover, from (62) to (64), it can be concluded that elements of  $\text{cov}\{\mathbf{h}(n)\}$  decrease as time progresses, becoming null as  $n \rightarrow \infty$ . This result is in agreement with the discussion of the results shown in Figure 10 because the assumptions made are the same that led to bias  $\{\mathbf{h}(n)\} = \mathbf{0}$  and thus  $\text{cov}\{\mathbf{h}(n)\} = \mathbf{0}$ .

The presented results corroborate the theoretical discussion on the covariance matrix of the acoustic feedback path estimate provided by the AFC systems based on direct closed-loop identification. In addition, they exemplify the slow convergence speed of the standard adaptive filtering algorithms and its direct relationship with the autocorrelation matrix of the system input signal  $u(n)$ . This fact justifies the development of specific methods to estimate the impulse response  $\mathbf{f}$  of the acoustic feedback path in AFC systems.

## VI. CONCLUSIONS

This work presented, using the least squares estimation theory, a theoretical and experimental analysis on the poor performance of standard adaptive filtering algorithms, which approximate the Wiener solution through the gradient or least squares, when applied to acoustic feedback cancellation with direct closed-loop identification.

Expressions for the bias and covariance matrix of the acoustic feedback path estimate provided by these algorithms were derived as functions of the signals statistics as well as derivatives of the cost function. Stationary and non-stationary environments were considered. It was demonstrated that, in general, the acoustic feedback path estimate is biased and presents a large covariance because the closed-loop nature of the system makes the cross-correlation between the loudspeaker and system input signals non-zero. As a consequence, the adaptive filter converges very slowly to a solution that is not the true acoustic feedback path.

These problems were exemplified using speech signals, a long acoustic feedback path and the RLS algorithm. It was verified that, after 10 s, the bias energy is usually greater than  $-15$  dB while the RLS algorithm achieves a normalized misalignment not less than  $-3.5$  dB and increases the maximum stable gain by no more than just 4.2 dB. The relationship between the performance of the adaptive algorithm and the aforementioned cross-correlation was proven by varying the signal-to-noise ratio and the delay caused by forward path.

## APPENDIX A

### EXPANSION OF THE CROSS-CORRELATION VECTOR $\mathbf{p}_{xu}$

The  $(N_H + 1) \times 1$  cross-correlation vector between  $x(n)$  and  $u(n)$  is defined as [11]

$$\mathbf{p}_{xu} = [r_{xu}(0) \quad r_{xu}(1) \quad \cdots \quad r_{xu}(N_H)]^T, \quad (65)$$

where

$$r_{xu}(i) = \text{E}\{x(n-i)u(n)\} \quad (66)$$

is the cross-correlation function between  $x(n)$  and  $u(n)$  [11].

By replacing (5) in (66), we have

$$\begin{aligned} r_{xu}(i) &= \text{E}\left\{\sum_{k=-\infty}^{\infty} a_k u(n-k-i)u(n)\right\} \\ &= \sum_{k=-\infty}^{\infty} a_k \text{E}\{u(n-k-i)u(n)\}. \end{aligned} \quad (67)$$

As  $u(n) = v(n) + r(n)$ , then (67) becomes

$$\begin{aligned} r_{xu}(i) &= \sum_{k=-\infty}^{\infty} a_k \text{E}\{[v(n-k-i) + r(n-k-i)][v(n) + r(n)]\} \\ &= \sum_{k=-\infty}^{\infty} a_k [\text{E}\{v(n-k-i)v(n)\} + \text{E}\{r(n-k-i)r(n)\} \\ &\quad + \text{E}\{r(n-k-i)v(n)\} + \text{E}\{v(n-k-i)r(n)\}] \\ &= \sum_{k=-\infty}^{\infty} a_k [r_v(k+i) + r_r(k+i) + r_{rv}(k+i) + r_{vr}(k+i)] \end{aligned} \quad (68)$$

where  $r_v(i) = \text{E}\{v(n-i)v(n)\}$  and  $r_r(i) = \text{E}\{r(n-i)r(n)\}$  are the autocorrelation functions of  $v(n)$  and  $r(n)$ , respectively,  $r_{vr}(i) = \text{E}\{v(n-i)r(n)\}$  is the cross-correlation function between  $v(n)$  and  $r(n)$ , and  $r_{rv}(i) = r_{vr}(-i)$  is the cross-correlation function between  $r(n)$  and  $v(n)$ .

Assuming that  $r(n)$  is a white Gaussian noise with zero mean and variance  $\sigma^2$ , then  $r_{rv}(i) = r_{vr}(i) = 0$  and

$$r_r(i) = \begin{cases} \sigma^2, & i = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (69)$$

Hence, in this case, (68) becomes

$$\begin{aligned} r_{xu}(i) &= \sum_{k=-\infty}^{\infty} a_k r_v(k+i) + \sum_{k=-\infty}^{\infty} a_k r_r(k+i) \\ &= \sum_{k=-\infty}^{\infty} a_k r_v(k+i) + \sigma^2 a_{-i}. \end{aligned} \quad (70)$$

Assuming that the closed-loop system is causal, then  $a_k = 0$ ,  $k < 0$ , and (70) becomes

$$r_{xu}(i) = \begin{cases} \sum_{k=0}^{\infty} a_k r_v(k+i) + \sigma^2 a_{-i}, & i \leq 0, \\ \sum_{k=0}^{\infty} a_k r_v(k+i), & i > 0. \end{cases} \quad (71)$$

Therefore, by replacing (71) in (65), the cross-correlation vector between the loudspeaker signal  $x(n)$  and the system input signal  $u(n)$  can be expanded as

$$\mathbf{p}_{xu} = \sum_{k=0}^{\infty} a_k \mathbf{p}_v(k) + \sigma^2 [a_0 \quad 0 \quad \cdots \quad 0]^T, \quad (72)$$

where

$$\mathbf{p}_v(k) = [r_v(k) \quad r_v(k+1) \quad \cdots \quad r_v(k+N_H)]^T. \quad (73)$$

APPENDIX B  
EXPANSION OF THE AUTOCORRELATION MATRIX  $\mathbf{R}_x$

The  $(N_H + 1) \times (N_H + 1)$  autocorrelation matrix of  $x(n)$  is defined as [11]

$$\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(N_H) \\ r_x(1) & r_x(0) & \cdots & r_x(N_H - 1) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(N_H) & r_x(N_H - 1) & \cdots & r_x(0) \end{bmatrix}, \quad (74)$$

where

$$r_x(i) = \mathbb{E} \{x(n)x(n-i)\} \quad (75)$$

is the autocorrelation function of  $x(n)$  [11].

By replacing (5) in (75), we have

$$r_x(i) = \sum_{l=-\infty}^{\infty} a_l \sum_{k=-\infty}^{\infty} a_k \mathbb{E} \{u(n)u(n-i)\}. \quad (76)$$

As  $u(n) = v(n) + r(n)$ , then (76) becomes

$$r_x(i) = \sum_{l=-\infty}^{\infty} a_l \sum_{k=-\infty}^{\infty} a_k [r_v(i+k-l) + r_r(i+k-l) + r_{vr}(i+k-l) + r_{rv}(i+k-l)]. \quad (77)$$

Assuming that  $r(n)$  is a white Gaussian noise with zero mean and variance  $\sigma^2$ , then  $r_{vr}(i) = r_{rv}(i) = 0$  and  $r_v(i)$  is defined as (69). Hence, in this case, (77) becomes

$$\begin{aligned} r_x(i) &= \sum_{l=-\infty}^{\infty} a_l \sum_{k=-\infty}^{\infty} a_k r_v(i+k-l) \\ &+ \sum_{l=-\infty}^{\infty} a_l \sum_{k=-\infty}^{\infty} a_k r_r(i+k-l) \\ &= \sum_{l=-\infty}^{\infty} a_l \sum_{k=-\infty}^{\infty} a_k r_v(i+k-l) + \sigma^2 \sum_{l=-\infty}^{\infty} a_l a_{l-i} \\ &= \sum_{l=-\infty}^{\infty} a_l \sum_{k=-\infty}^{\infty} a_k r_v(i+k-l) + \sigma^2 \tilde{r}_a(i) \end{aligned} \quad (78)$$

where

$$\tilde{r}_a(i) = \sum_{l=-\infty}^{\infty} a_l a_{l-i} \quad (79)$$

is the sample autocorrelation function of  $\mathbf{a}$  [11].

Assuming that the closed-loop system is causal, then  $a_k = 0$ ,  $k < 0$ , and (78) becomes

$$r_x(i) = \sum_{l=0}^{\infty} a_l \sum_{k=0}^{\infty} a_k r_v(i+k-l) + \sigma^2 \tilde{r}_a(i). \quad (80)$$

Expanding the summation and regrouping its terms, (80) may be written as

$$\begin{aligned} r_x(i) &= a_0 [a_0 r_v(i) + a_1 r_v(i+1) + a_2 r_v(i+2) + \dots] \\ &+ a_1 [a_0 r_v(i-1) + a_1 r_v(i) + a_2 r_v(i+1) + \dots] \\ &+ a_2 [a_0 r_v(i-2) + a_1 r_v(i-1) + a_2 r_v(i) + \dots] \\ &+ \dots + \\ &+ a_N [a_0 r_v(i-N) + a_1 r_v(i-N+1) + a_2 r_v(i-N+2) + \dots] \\ &+ \dots + \sigma^2 \tilde{r}_a(i) \\ &= r_v(i) [a_0 a_0 + a_1 a_1 + a_2 a_2 + \dots] \\ &+ r_v(i+1) [a_0 a_1 + a_1 a_2 + a_2 a_3 + \dots] \\ &+ r_v(i-1) [a_1 a_0 + a_2 a_1 + a_3 a_2 + \dots] \\ &+ \dots + \\ &+ r_v(i+N) [a_0 a_N + a_1 a_{N+1} + a_2 a_{N+2} + \dots] \\ &+ r_v(i-N) [a_N a_0 + a_{N+1} a_1 + a_{N+2} a_2 + \dots] \\ &+ \dots + \sigma^2 \tilde{r}_a(i). \end{aligned} \quad (81)$$

Using (79), (81) can be written as

$$\begin{aligned} r_x(i) &= r_v(i) \tilde{r}_a(0) + r_v(i+1) \tilde{r}_a(1) + r_v(i-1) \tilde{r}_a(-1) \\ &+ \dots + r_v(i+N) \tilde{r}_a(N) + r_v(i-N) \tilde{r}_a(-N) \\ &+ \dots + \sigma^2 \tilde{r}_a(i) \\ &= r_v(i) \tilde{r}_a(0) + \sum_{k=1}^{\infty} [r_v(i+k) + r_v(i-k)] \tilde{r}_a(k) + \sigma^2 \tilde{r}_a(i) \\ &= \sum_{k=-\infty}^{\infty} r_v(i+k) \tilde{r}_a(|k|) + \sigma^2 \tilde{r}_a(i). \end{aligned} \quad (82)$$

Finally, as  $\tilde{r}_a(k) = \tilde{r}_a(-k)$ , the autocorrelation function of the loudspeaker signal  $x(n)$  is given by

$$r_x(i) = \sum_{k=-\infty}^{\infty} r_v(i+k) \tilde{r}_a(k) + \sigma^2 \tilde{r}_a(i). \quad (83)$$

Therefore, by replacing (83) in (74), the autocorrelation matrix of the loudspeaker signal  $x(n)$  can be expanded as

$$\mathbf{R}_x = \sum_{k=-\infty}^{\infty} \tilde{r}_a(k) \mathbf{R}_v(k) + \sigma^2 \tilde{\mathbf{R}}_a, \quad (84)$$

where

$$\mathbf{R}_v(k) = \begin{bmatrix} r_v(k) & r_v(k+1) & \cdots & r_v(k+N_H) \\ r_v(k+1) & r_v(k) & \cdots & r_v(k+N_H-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_v(k+N_H) & r_v(k+N_H-1) & \cdots & r_v(k) \end{bmatrix} \quad (85)$$

is an  $(N_H + 1) \times 1$  autocorrelation matrix of the source signal  $v(n)$  so that  $\mathbf{R}_v(0) = \mathbf{R}_v$  and

$$\tilde{\mathbf{R}}_a = \begin{bmatrix} \tilde{r}_a(0) & \tilde{r}_a(1) & \cdots & \tilde{r}_a(N_H) \\ \tilde{r}_a(1) & \tilde{r}_a(0) & \cdots & \tilde{r}_a(N_H-1) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{r}_a(N_H) & \tilde{r}_a(N_H-1) & \cdots & \tilde{r}_a(0) \end{bmatrix} \quad (86)$$

is the  $(N_H + 1) \times (N_H + 1)$  sample autocorrelation matrix of  $\mathbf{a}$ .

## ACKNOWLEDGMENT

The authors would like to acknowledge the FAPPR (Fundação Araucária), SETI-PR (Secretaria de Estado da Ciência, Tecnologia e Ensino Superior) and the Government of the State of Paraná for the financial support received. This work was also supported by National Funds from FCT - Fundação para a Ciência e a Tecnologia through project UIDB/50016/2020.

## REFERENCES

- [1] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: state of the art and future challenges," *Proceedings of the IEEE*, vol. 99, no. 2, pp. 288–327, February 2011, doi: 10.1109/JPROC.2010.2090998.
- [2] B. C. Bispo and D. R. S. Freitas, "On the use of cepstral analysis in acoustic feedback cancellation," *Digital Signal Processing*, vol. 44, pp. 88–101, September 2015, doi: 10.1016/j.dsp.2015.03.003.
- [3] —, "A review of the recent cepstrum-based methods for acoustic feedback cancellation," in *Advances in Signal Processing: Reviews*, S. Y. Yurish, Ed. International Frequency Sensor Association Publishing, 2018, vol. 1, pp. 121–154.
- [4] M. G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 4, pp. 443–453, July 2000, doi: 10.1109/89.848225.
- [5] J. Hellgren and U. Forsell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 7, pp. 906–913, November 2001, doi: 10.1109/89.966094.
- [6] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3749–3763, October 2005, doi: 10.1109/TSP.2005.855108.
- [7] M. H. Costa, "Theoretical transient analysis of a hearing aid feedback canceller with a saturation type nonlinearity in the direct path," *Computers in Biology and Medicine*, vol. 91, no. 1, pp. 243–254, December 2017, doi: 10.1016/j.cmpbiomed.2017.10.018.
- [8] S. Kochkin, "Marketrak vii: Customer satisfaction with hearing instruments in the digital age," *The Hearing Journal*, vol. 58, no. 9, pp. 30–32, 2005, doi: 10.1097/01.HJ.0000286545.33961.e7.
- [9] J. Agnew, "Acoustic feedback and other audible artifacts in hearing aids," *Trends in Amplification*, vol. 1, no. 2, pp. 45–82, June 1996, doi: 10.1177/108471389600100202.
- [10] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 4th ed. New York, New York: Springer Science+Business Media, 2013.
- [11] S. Haykin, *Adaptive Filter Theory*, 5th ed. Harlow, Essex: Pearson, 2014.
- [12] U. Forssell, "Closed-loop identification: methods, theory, and applications," Ph.D. dissertation, Linköpings Universitet, 1999.
- [13] U. Forssell and L. Ljung, "Closed-loop identification revisited," *Automatica*, vol. 35, no. 7, pp. 1215–1241, July 1999, doi: 10.1016/S0005-1098(99)00022-9.
- [14] T. van Waterschoot, G. Rombouts, and M. Moonen, "On the performance of decorrelation by prefiltering for adaptive feedback cancellation in public address system," in *Processing of 4th IEEE Benelux Signal Processing Symposium*, Hilvarenbeek, The Netherlands, April 2004, pp. 167–170.
- [15] M. Guo, S. H. Jensen, and J. Jensen, "Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancement," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 9, pp. 2549–2563, November 2012, doi: 10.1109/TASL.2012.2206025.
- [16] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "On the use of a phase modulation method for decorrelation in acoustic feedback cancellation," in *Proceedings of 20th European Signal Processing Conference*, Bucharest, Romania, August 2012, pp. 2000–2004.
- [17] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, New Jersey: Prentice Hall, 1987.
- [18] J. Hellgren, "Analysis of feedback cancellation in hearing aids with Filtered-X LMS and the direct method of closed loop identification," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 2, pp. 119–131, February 2002, doi: 10.1109/89.985549.
- [19] G. Rombouts, T. van Waterschoot, K. Struyve, and M. Moonen, "Acoustic feedback cancellation for long acoustic paths using a nonstationary source model," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3426–3434, September 2006, doi: 10.1109/TSP.2006.879251.
- [20] B. C. Bispo, P. M. L. Rodrigues, and D. R. S. Freitas, "Acoustic feedback cancellation based on cepstral analysis," in *Proceedings of 17th IEEE Conference on Signal Processing Algorithms, Architectures, Arrangements and Applications*, Poznan, Poland, September 2013, pp. 205–209.
- [21] M. G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation systems for hearing aids with a delayed cancellation path," in *Conference Record of 32nd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, USA, November 1998, pp. 518–522.
- [22] —, "Bias analysis in continuous adaptation systems for hearing aids," in *Proceedings of IEEE Conference on Acoustics, Speech, and Signal Processing*, Phoenix, USA, March 1999, pp. 925–928.
- [23] J. Hellgren and U. Forsell, "Bias of feedback cancellation algorithms based on direct closed loop identification," in *Proceedings of IEEE Conference on Acoustics, Speech, and Signal Processing*, Istanbul, Turkey, June 2000, pp. 869–872.
- [24] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, New Jersey: Prentice Hall, 1999.
- [25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, New Jersey: Prentice Hall, 1993.
- [26] P. Z. Peebles Jr., *Probability, Random Variables, and Random Signal Principles*, 2nd ed. New York: McGraw-Hill, 1987.
- [27] M. Jeub, M. Schäfer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *Proceedings of International Conference on Digital Signal Processing*, Santorini, Greece, July 2009.
- [28] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2013.



**Bruno Catarino Bispo** received the B.Sc. degree in electronics and computer engineering from the Federal University of Rio de Janeiro (UFRJ), Brazil, in 2006, the M.Sc. degree in electrical engineering from UFRJ in 2008, and the Ph.D. degree in electrical and computer engineering from the University of Porto, Portugal, in 2015. From 2009 to 2012, he was a Lecturer at the Faculty of Engineering of the University of Porto. From 2015 to 2019, he was an Assistant Professor at the Federal University of Technology – Paraná (UTFPR), Brazil. Since 2019,

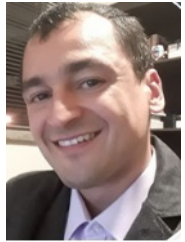
he has been an Assistant Professor in the Department of Electrical and Electronic Engineering at the Federal University of Santa Catarina (UFSC), Florianópolis, Brazil. His research interests are in digital signal processing, particularly speech and biomedical signal processing and adaptive systems.



**César Fumio Yamamura** received the B.Sc. degree in electronics engineering from the Federal University of Technology – Paraná (UTFPR), Cornélio Procopio, Brazil, in 2019. He is currently pursuing a M.Sc. degree in electrical engineering at UTFPR. His research interests are in digital signal processing and machine learning, particularly applied to acoustic feedback cancellation and speech processing.



**Wellington Murilo da Silva Nogueira** received the B.Sc. degree in electronics engineering from the Federal University of Technology – Paraná (UTFPR), Cornélio Procópio, Brazil, in 2018. His research interests are in digital signal processing, specially applied to acoustic feedback cancellation and speech processing.



**Edson Aparecido Rozas Theodoro** received the B.Eng., M.Sc. and D.Sc. degrees in electrical engineering from University of São Paulo (USP), São Carlos, Brazil, in 2007, 2010 and 2013, respectively. He is currently an Assistant Professor (since 2015) with the Electrical Engineering Department at Federal University of Technology – Paraná (UTFPR), Cornélio Procópio, Brazil. His research interests include intelligent systems applications, dynamical systems, and power systems stability.



**Pedro Miguel Rodrigues** received the B.Sc. and M.Sc. degrees in biomedical engineering from the Polytechnic Institute of Bragança, Portugal, in 2009 and 2011, respectively, and the Ph.D. degree in biomedical engineering from the University of Porto, Portugal, in 2017. Since 2018, he is an Assistant Professor at the Faculty of Biotechnology – Universidade Católica Portuguesa (ESB-UCP) and member of the research Centre for Biotechnology and Fine Chemistry (CBQF-UCP), Porto, Portugal. His research areas focus on digital signal processing and artificial intelligence applied to the diagnosis of neurodegenerative, cardiac and speech diseases, and acoustic feedback cancellation.