Adaptive Feedback Cancellation Using a Partitioned-Block Frequency-Domain Kalman Filter Approach With PEM-Based Signal Prewhitening

Giuliano Bernardi, *Student Member, IEEE*, Toon van Waterschoot, *Member, IEEE*, Jan Wouters, and Marc Moonen, *Fellow, IEEE*

***Abstract*—Adaptive filtering based feedback cancellation is a widespread approach to acoustic feedback control. However, traditional adaptive filtering algorithms have to be modified in order to work satisfactorily in a closed-loop scenario. In particular, the undesired signal correlation between the loudspeaker signal and the source signal in a closed-loop scenario is one of the major problems to address when using adaptive filters for feedback cancellation. Slow convergence speed and limited tracking capabilities are other important limitations to be considered. Additionally, computationally expensive algorithms as well as long delays should be avoided, for instance, in hearing aid applications, because of power constraints, important to extend battery life, and real-time implementations requirements, respectively. We present an algorithm combining good decorrelation properties, by means of the prediction-error method based signal prewhitening, fast convergence, good tracking behavior, and low computational complexity by means of the frequency-domain Kalman filter, and low delay by means of a partitioned-block implementation.**

***Index Terms*—Adaptive feedback cancellation (AFC), acoustic feedback control, frequency-domain adaptive filter (FDAF), Kalman filter, prediction-error method (PEM).**

I. Introduction

ACOUSTIC feedback control is of critical importance in several systems dealing with acoustic signals, such as public address (PA) systems and HAs. A lack of acoustic feedback control can lead to system instabilities causing annoying artifacts and sound degradation. Mainly in the last three decades, several methods have been developed to cope with the problem of acoustic feedback [1]. An important class of such methods is characterized by the use of adaptive filters and, more specifically, by the use of the adaptive filters to model the unknown feedback path. adaptive feedback cancellation (AFC) is the usual name by which these methods are identified. An illustration of a typical acoustic feedback scenario including an AFC approach is shown in Fig. 1; the adaptive filter $\hat{F}(q,t)$ represents the estimated feedback path model which should, ideally, perfectly match the true feedback path $F_{\mathrm{t}}(q,t)$, in order to reduce the feedback artifact. $\hat{F}(q,t)$ and $F_{\mathrm{t}}(q,t)$ are assumed linear and possibly time-varying and will be further defined in Section II. Here, t is the discrete time index and q^{-1} is the delay operator, i.e., $q^{-k}u(t) = u(t-k)$, which allows a compact definition of the different TFs and will be used throughout the paper. The nature of the problem can be seen by noticing that the microphone signal $y(t)$ is not only composed of the source signal $v(t)$, i.e., the desired signal to be amplified and sent to the loudspeaker, but also of the undesired interference $x(t)$, originating from the presence of the acoustic feedback, i.e., $y(t) = x(t) + v(t)$. A similar situation also characterizes the standard acoustic echo scenario; what differentiates the acoustic feedback scenario from the acoustic echo scenario is the presence of the forward path transfer function $G(q,t)$, turning the system into a closed-loop system, and introducing a signal correlation between the loudspeaker signal $u(t)$ and the source

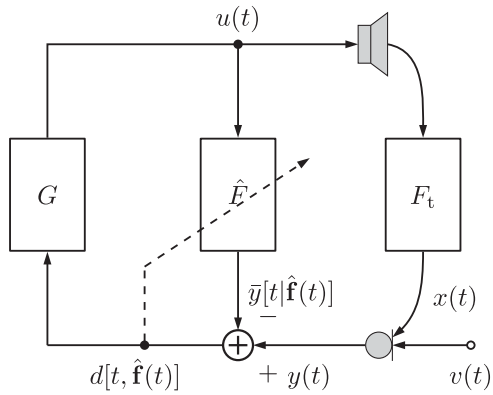

Fig. 1. General AFC scenario.

Manuscript received April 20, 2017; accepted June 13, 2017. Date of publication June 15, 2017; date of current version July 12, 2017. This research work was carried out at the ESAT Laboratory and at the ExpORL Laboratory of KU Leuven, in the frame of the IWT O&O Project 110722 "Signal processing and automatic fitting for next generation cochlear implants," KU Leuven Research Council CoE PFV/10/002 (OPTEC), the Interuniversity Attractive Poles Programme initiated by the Belgian Science Policy Office: IUAP P7/19 "Dynamical systems control and optimization" (DYSCO) 2012–2017. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Richard Christian Hendriks. *(Corresponding author: Giuliano Bernardi.)*

G. Bernardi and M. Moonen are with the Department of Electrical Engineering, ESAT-STADIUS, KU Leuven, B-3001 Leuven, Belgium (e-mail: giuliano.bernardi@esat.kuleuven.be; marc.moonen@esat.kuleuven.be).

T. van Waterschoot is with the Department of Electrical Engineering, ESAT-STADIUS, KU Leuven, B-3001 Leuven, Belgium, and also with the Department of Electrical Engineering, ESAT-ETC, AdvISe Lab, B-2440 Geel, Belgium (e-mail: toon.vanwaterschoot@esat.kuleuven.be).

J. Wouters is with the Department of Neurosciences, Laboratory ExpORL, KU Leuven, B-3000 Leuven, Belgium (e-mail: jan.wouters@med.kuleuven.be).

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

Digital Object Identifier 10.1109/TASLP.2017.2716188

signal $v(t)$. This correlation makes the estimation of the feedback path more problematic than in the acoustic echo scenario and, as a consequence, employing a standard adaptive filtering algorithm, e. g. the normalized least mean squares (NLMS), returns a biased estimate of $F_{\mathrm{t}}(q,t)$ [2], [3], thus limiting the cancellation properties of $\hat{F}(q,t)$. Additionally, system instabilities can be induced by the closed-loop, leading to a series of acoustic artifacts such as howling. In order to reduce these problems and obtain a reliable estimate, a procedure for decorrelating $v(t)$ and $u(t)$ should be included.

Different approaches have been proposed in the literature to reduce the signal correlation in the acoustic feedback scenario, and thus produce better estimates of the feedback path transfer function (TF), such as the introduction of an external probe noise [4], [5], modifications of the forward path TF by means of nonlinear processing [6], [7], time-varying processing [6], [8] and added delays [9], two-microphones strategies [10], and, more recently, the use of a prewhitening filter used for decorrelation [11]–[15]. The latter approach relies on the use of an appropriate model for the disturbance of the identification procedure which, in the AFC context, is represented by the source signal $v(t)$.

The prewhitening filter-based AFC has been shown to be advantageous since it provides limited perceptual distortions, unlike the other aforementioned approaches [1], [16]. However, the need for a source signal model introduces a new challenge from the identification point of view, since the unknown source signal $v(t)$ is usually a nonstationary speech or audio signal. Nonstationarity implies that the source signal model for $v(t)$ must be concurrently estimated alongside the estimation of the feedback path model. Therefore, the identifiability conditionss (ICs) of the system, which now counts two models to be identified, are inevitably changed [3], [17].

The application of the prediction-error method (PEM) to prewhitening filter-based AFC has been widely studied [3], [17]–[21], resulting in several different algorithms, e. g. the PEM-based adaptive filtering with row operations (PEM-AFROW), as well as interesting results regarding model identifiability. In the time-invariant case, with a true source signal generation system $H_{\mathrm{t}}(q)$ defined by an autoregressive (AR) process with a white noise excitation signal $e(t)$, see Fig. 2, Spriet *et al.* [3] have proved that identifiability can be achieved if sufficient delay is included in the forward path or in the feedback cancellation path, as well as if a time-varying or nonlinear processing forward path TF is considered. This identifiability analysis has subsequently been extended to a wider range of source signal models [17].

The AFC has been also formulated in the frequency domain, i.e., as a frequency-domain adaptive filter (FDAF), and combined with a time-domain prewhitening filter, i.e., the PEM-based frequency-domain adaptive filter (PEM-FDAF) [19], [22]. More recently, a PEM-based prewhitening filter has been used in combination with a frequency-domain Kalman filter (FDKF) applied to a state-space structure, leading to the so-called PEM-based frequency-domain Kalman filter (PEM-FDKF) [18], to achieve better convergence and tracking properties compared to the PEM-FDAF [22], [23]. An advantage of the PEM-FDKF

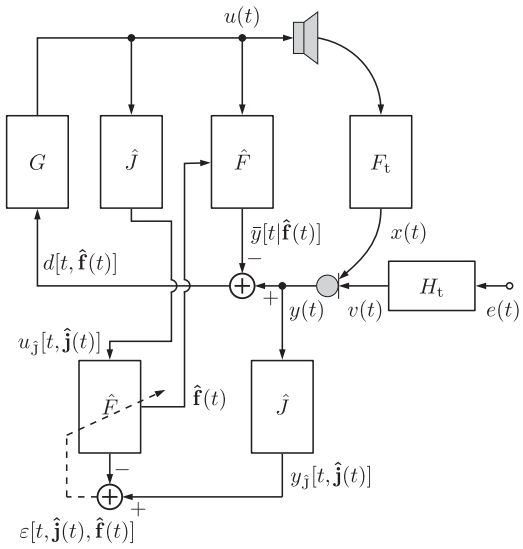

Fig. 2. Complete AFC algorithm with prediction-error method (PEM) stage.

is the inherent optimal choice of the step-size parameter [24] which usually needs to be fixed as design parameter of the PEM-FDAF algorithm or adaptively estimated using variable step-size algorithms [14], [23], [25].

In this paper, we provide the complete derivation of the PEM-FDKF algorithm, which was not included in [18], together with a complexity analysis and a study of the ICs for the closed-loop identification. Additionally, we propose an extension of the PEM-FDKF by means of a partitioned-block (PB) frequency-domain implementation, referred to as the PEM-based partitioned-block frequency-domain Kalman filter (PEM-PBFDKF), allowing to reduce the algorithmic delay, as needed in, e. g., hearing aid (HA) applications. The paper is organized as follows. In Section II, we review the PEM for direct closed-loop system identification. In Section III, we introduce the PEM-FDKF, providing a complete derivation of the algorithm. In Section IV, we study the ICs, allowing to obtain a unique and unbiased model estimate for both the feedback path and the source signal generation system. In Section V, we present the extension of the PEM-FDKF relying on partitioned-block (PB) processing, the PEM-PBFDKF. In Section VI, we provide a computational complexity and memory requirements comparison of the proposed algorithms. In Section VII, we illustrate the performance of the proposed algorithms in terms of convergence speed, added stability and sound quality by means of simulation results. Finally, the conclusions are drawn in Section VIII.

II. Prediction Error Method Identification

The PEM is widely used in direct closed-loop system identification [2], [3]. For the case illustrated in Fig. 2, the PEM can be used to provide a direct closed-loop identification [2], [26] of both the true feedback path $F_{\mathrm{t}}(q,t)$ and the true source signal generation system $H_{\mathrm{t}}(q,t)$. Throughout the paper, we use the following notation system: a symbol with the subscript $_{\mathrm{t}}$ refers to the true system, a regular symbol refers to the model, and a

symbol with a hat refers to the model estimate; e. g., $F_{\mathrm{t}}(q,t)$ is the true feedback path, $F(q,t)$ is the feedback path model, and $\hat{F}(q,t)$ is the feedback path model estimate.

Assuming $F(q,t)$ and $H(q,t)$ to be parametric difference equation models, and defining a new model $J(q,t)$ satisfying the equation $J(q,t)H(q,t)=1$ for later use, we introduce the parameter vectors $\boldsymbol{\theta}(t)$, $\mathbf{f}(t)$, and $\mathbf{j}(t)$:

$$\boldsymbol{\theta}(t) = [\mathbf{f}^{\mathrm{T}}(t) \quad \mathbf{j}^{\mathrm{T}}(t)]^{\mathrm{T}} \tag{1}$$

$$\mathbf{f}(t) = [f_0(t) \quad f_1(t) \quad \dots \quad f_{n_F-1}(t)]^{\mathrm{T}} \tag{2}$$

$$\mathbf{j}(t) = [1 \quad j_1(t) \quad \dots \quad j_{n_J-1}(t)]^{\mathrm{T}}, \tag{3}$$

where $n_\theta = n_F + n_J$. Assuming the true system is contained in the model set [26], the true system can be described using the true values of $\mathbf{f}(t)$, i.e., $\mathbf{f}_{\mathrm{t}}(t)$, as

$$F_{\mathrm{t}}(q,t) = F(q,t)\big|_{\mathbf{f}(t)=\mathbf{f}_{\mathrm{t}}(t)} \tag{4a}$$

$$= f_{\mathrm{t},0}(t) + f_{\mathrm{t},1}(t)q^{-1} + \ldots + f_{\mathrm{t},n_F-1}(t)q^{-n_F+1} \tag{4b}$$

and thus

$$y(t) = F_{\mathrm{t}}(q,t)u(t) + H_{\mathrm{t}}(q,t)e(t). \tag{5}$$

Similarly to (4), the true value of $\mathbf{j}(t)$, i.e., $\mathbf{j}_{\mathrm{t}}(t)$, can be used to write $J_{\mathrm{t}}(q,t) = J(q,t)\big|_{\mathbf{j}(t)=\mathbf{j}_{\mathrm{t}}(t)}$, with $J_{\mathrm{t}}(q,t)H_{\mathrm{t}}(q,t)=1$.

We can now define the prediction error (PE) using the one-step ahead predictor for $y(t)$, $\bar{y}[t|\mathbf{f}(t),\mathbf{j}(t)]$, as

$$\varepsilon[t,\boldsymbol{\theta}(t)] = y(t) - \bar{y}[t|\boldsymbol{\theta}(t)] \tag{6a}$$

$$= J(q,t)\left[y(t) - F(q,t)u(t)\right], \tag{6b}$$

and find the true values of the parameter vectors $\mathbf{f}(t)$ and $\mathbf{j}(t)$, by minimizing the variance of the PE

$$\min_{\boldsymbol{\theta}(t)} \quad \mathbb{E}\{\varepsilon^2[t,\boldsymbol{\theta}(t)]\}, \tag{7}$$

where $\mathbb{E}\{\cdot\}$ denotes statistical expectation and the measured $\varepsilon[t,\boldsymbol{\theta}(t)]$ is considered to be a realization of the PE, deriving from a realization of the white noise excitation $e(t)$, being the only random variable in this scenario.

The ICs, i.e., the conditions that allow to uniquely estimate $F_{\mathrm{t}}(q,t)$ and $J_{\mathrm{t}}(q,t)$, have been derived in literature [3] by converting the nonlinear PEM cost function (7) into a linear cost function, by means of the transformation:

$$A(q,t) = J(q,t) \tag{8}$$

$$B(q,t) = -J(q,t)F(q,t) \tag{9}$$

with $A(q,t)$ and $B(q,t)$ parameterized by

$$\boldsymbol{\xi}(t) = [\mathbf{a}^{\mathrm{T}}(t) \quad \mathbf{b}^{\mathrm{T}}(t)]^{\mathrm{T}} \tag{10}$$

$$\mathbf{a}(t) = [1 \quad a_1(t) \quad \dots \quad a_{n_A}(t)]^{\mathrm{T}} \tag{11}$$

$$\mathbf{b}(t) = [b_0(t) \quad b_1(t) \quad \dots \quad b_{n_B}(t)]^{\mathrm{T}}, \tag{12}$$

allowing to rewrite (6b) as

$$\varepsilon[t,\boldsymbol{\xi}(t)] = A(q,t)y(t) + B(q,t)u(t). \tag{13}$$

and (7) as the constrained optimization problem

$$\min_{\boldsymbol{\xi}(t)} \quad \mathbb{E}\{\varepsilon^2[t,\boldsymbol{\xi}(t)]\} \tag{14a}$$

$$\text{subject to} \quad A(q,t) \text{ is a divisor of } B(q,t), \tag{14b}$$

where the constraint follows from (8) and (9). The first tap parameter of $A(q,t)$ is always set to 1, i.e., $A(q,t) = 1 + q^{-1}\bar{A}(q,t)$, in order to avoid the trivial solution $A(q,t) = B(q,t) = 0$ in (14). The PE definition in (13) is used in the first part of Section IV, where we provide the ICs for the optimization problem leading to the proposed PEM-FDKF. Unlike in the derivation in [3], here $A(q,t)$ and $B(q,t)$ are considered to be time-varying quantities. Later in Section IV, we use again the PE definition in (6b) from the derivation [it will also be clear that the constraint (14b) may be removed] as the solution of (14a) alone satisfies the constraint (14b) and hence the solution of (14a) is also equal to the solution of (7).

In the next section, describing the algorithmic derivation of the PEM-FDKF, we also use (6b) to describe $\varepsilon[t,\boldsymbol{\theta}(t)]$. Following the PEM, the optimization in (7), w.r.t. $\mathbf{f}(t)$ and $\mathbf{j}(t)$, is carried out in an alternating fashion [21], i.e., estimating, at each iteration, in a first step the coefficients of $J(q,t)$ with fixed estimates for $F(q,t)$ and in a second step the coefficients of $F(q,t)$ with fixed estimates for $J(q,t)$.

III. The PEM-Based Frequency Domain Kalman Filter (PEM-FDKF)

The PEM-FDKF algorithm [18] is an extension of the algorithm proposed by Enzner and Vary [24] for acoustic echo cancellation (AEC). It relies on a dynamical model for the feedback path and a model for the recorded microphone signal to define a state-space representation to which the Kalman filter procedure can be applied. The simple frequency-domain dynamical model chosen for the feedback path employs a first-order Markov model as an abstraction of the true feedback path dynamics. Similar models have also been proposed in the time-domain [27]–[29]; however, the use of a block-wise procedure has an impact on the calculation of the model time constant. The main change introduced to the algorithm by Enzner and Vary consists of the decorrelation stage, by means of a PEM-based prewhitening filter already seen in the PEM-FDAF [23], [30], [31]. In this way, the frequency-domain Kalman filter (FDKF) framework can be successfully applied to AFC, leading to the PEM-FDKF.

The two main advantages of a frequency-domain approach are the lower computational complexity and the good decorrelation properties of the discrete Fourier transform (DFT) [19], [32], [33]. In HA applications, where short filters are usually employed, the computational complexity advantage is smaller than in PA applications, but it can be still relevant, as we will show in Section VI, especially if the DFT calculations are shared by other processing stages of the HA. In addition to these two advantages, the formulation of the problem based on a state-space representation allows to optimally estimate the FDAF stepsize as part of the Kalman filter procedure [24], [34], leading to a significantly improved convergence [18].

It should be emphasized that this approach employs the implicit assumption that $F_{\mathrm{t}}(q,t)$ and $H_{\mathrm{t}}(q,t)$ are slowly time-varying [with $F_{\mathrm{t}}(q,t)$ varying more slowly than $H_{\mathrm{t}}(q,t)$], since these are modeled by $F(q,\kappa)$ and $H(q,\kappa)$, where $\kappa \in \mathbb{Z}$ is the time frame index, which hence can only vary at the frame rate [24], [35]. Therefore, we will assume time invariance of $F_{\mathrm{t}}(q,t)$ and $H_{\mathrm{t}}(q,t)$ over each frame, i.e., effectively over each frame shift of R samples, and use the notation $F_{\mathrm{t}}(q,\kappa)$ and $H_{\mathrm{t}}(q,\kappa)$. Similarly, $J_{\mathrm{t}}(q,\kappa)$ is defined such that $J_{\mathrm{t}}(q,\kappa)H_{\mathrm{t}}(q,\kappa)=1$.

The discrete frequency index l will be used, together with the frame index κ, to describe the time-frequency components of the different variables. The first introduced variable is the M-samples loudspeaker signal for frame κ,

$$\mathbf{u}(\kappa) = [u(\kappa R - M + 1) \quad \dots \quad u(\kappa R)]^{\mathrm{T}}, \tag{15}$$

where R denotes the frame shift.

Assuming the true value $J_{\mathrm{t}}(q,\kappa)$ to be available, the prefiltered version of the loudspeaker signal for frame κ is

$$\begin{aligned}\mathbf{u}_{\mathrm{J_t}}(\kappa) &= [J_{\mathrm{t}}(q,\kappa)u(\kappa R - M + 1) \ \dots \ J_{\mathrm{t}}(q,\kappa)u(\kappa R)]^{\mathrm{T}} \\ &= [u_{\mathrm{J_t}}(\kappa R - M + 1) \quad \dots \quad u_{\mathrm{J_t}}(\kappa R)]^{\mathrm{T}}. \end{aligned} \tag{16}$$

The frequency-domain version of the prefiltered loudspeaker signal is then given in matrix form as

$$\mathbf{U}_{\mathrm{J_t}}(\kappa) = \mathrm{diag}\left\{\mathbf{F}_M \mathbf{u}_{\mathrm{J_t}}(\kappa)\right\}, \tag{17}$$

where $\mathbf{F}_M$ is the unitary DFT matrix of size $M \times M$, i.e., $\mathbf{F}_M^{-1} = \mathbf{F}_M^{\mathrm{H}}$, and the $\mathrm{diag}\{\cdot\}$ operator either maps an $M \times 1$ vector to the diagonal of an $M \times M$ diagonal matrix, or maps an $M \times M$ matrix to the $M \times 1$ vector given by its diagonal. The dimension parameters R and M should be chosen properly, taking into consideration the length of the true feedback path or an estimate thereof. A common choice is $R = n_{\hat{F}}$ (assuming $n_{\hat{F}} = n_F$) and $M = 2R$ [23], [32]; if the algorithmic delay, equal to $2R-1$, is not acceptable, a PB solution (also known as multidelay filter) [22], [33], [37] can be chosen, see Section V.

With $R = n_{\hat{F}}$, the frequency-domain version of the true feedback path parameter vector $\mathbf{f}_{\mathrm{t}}(\kappa)$ is

$$\mathbf{F}_{\mathrm{t}}(\kappa) = \mathbf{G}^{10}_{M\times R}\mathbf{F}_R\mathbf{f}_{\mathrm{t}}(\kappa) \tag{18a}$$

$$= \mathbf{F}_M\mathbf{W}^{01}_{M\times R}\mathbf{f}_{\mathrm{t}}(\kappa). \tag{18b}$$

The rectangular matrices $\mathbf{G}^{10}_{M\times R}$ and $\mathbf{W}^{01}_{M\times R}$ are used to obtain the $M \times 1$ frequency-domain version of the $R \times 1$ vector $\mathbf{f}_{\mathrm{t}}(\kappa)$. The smaller of the two matrix dimensions R always indicates the dimension of the identity matrix $\mathbf{I}_{R\times R}$ appearing in the matrix definitions. The simplified definitions of these matrices, considering the standard case $M = 2R$, are shown in Table I along with the other matrices needed to compactly describe the algorithm in the frequency-domain as defined by Benesty *et al.* [36]. Despite the two classes of definitions for frequency-domain quantities, e. g. using (18a) or (18b), we will only use the class of definitions including the linearization matrices $\mathbf{G}^{\cdot\cdot}_{\cdot\times\cdot}$ in the rest of the paper.

Finally, we introduce the R-samples prefiltered microphone signal and the source excitation signal for frame κ, i.e.,

$$\begin{aligned}\mathbf{y}_{\mathrm{J_t}}(\kappa) &= [J_{\mathrm{t}}(q,\kappa)y(\kappa R - R + 1) \ \dots \ J_{\mathrm{t}}(q,\kappa)y(\kappa R)]^{\mathrm{T}} \\ &= [y_{\mathrm{J_t}}(\kappa R - R + 1) \quad \dots \quad y_{\mathrm{J_t}}(\kappa R)]^{\mathrm{T}} \end{aligned} \tag{19}$$

$$\mathbf{e}(\kappa) = [e(\kappa R - R + 1) \quad \dots \quad e(\kappa R)]^{\mathrm{T}}, \tag{20}$$

and their frequency-domain versions:

$$\mathbf{Y}_{\mathrm{J_t}}(\kappa) = (\mathbf{G}^{01}_{R\times M})^{\mathrm{H}}\mathbf{F}_R\mathbf{y}_{\mathrm{J_t}}(\kappa) \tag{21}$$

$$\mathbf{E}(\kappa) = (\mathbf{G}^{01}_{R\times M})^{\mathrm{H}}\mathbf{F}_R\mathbf{e}(\kappa). \tag{22}$$

The quantities introduced so far can be combined into a state-space representation using the frequency-domain Markov model for the feedback path $\mathbf{F}_{\mathrm{t}}(\kappa)$ as a state equation and the linear model for the frequency-domain prefiltered microphone signal $\mathbf{Y}_{\mathrm{J_t}}(\kappa)$ as a measurement equation:

$$\mathbf{F}_{\mathrm{t}}(\kappa+1) = \alpha_{\mathrm{t}}\mathbf{F}_{\mathrm{t}}(\kappa) + \mathbf{N}_{\mathrm{t}}(\kappa) \tag{23a}$$

$$\mathbf{Y}_{\mathrm{J_t}}(\kappa) = \mathbf{C}_{\mathrm{J_t}}(\kappa)\mathbf{F}_{\mathrm{t}}(\kappa) + \mathbf{E}(\kappa) \tag{23b}$$

where $\mathbf{C}_{\mathrm{J_t}}(\kappa) = \mathbf{G}^{01}_{M\times M}\mathbf{U}_{\mathrm{J_t}}(\kappa)$ includes the linear transformation to be applied to $\mathbf{U}_{\mathrm{J_t}}(\kappa)$ to linearize the circular convolution between $\mathbf{u}_{\mathrm{J_t}}(\kappa)$ and $\mathbf{f}_{\mathrm{t}}(\kappa)$, cf. Table I, $\mathbf{N}_{\mathrm{t}}(\kappa)$ is the process noise describing the unpredictability of the feedback path dynamics and α_{t} is the transition factor accounting for the time-variability of the feedback path [24]. The total number of linear convolution samples resulting from the circular convolution, given the chosen signal dimension parameters, is $M - R + 1 = R + 1$ [32]. However, we retain only R samples to match the frame shift and simplify the notation, as commonly done in the literature [24], [32], [36], [38].

Assuming $J_{\mathrm{t}}(q,\kappa)$ is indeed available, the use of prefiltered variables in (23b) guarantees the decorrelation between the prefiltered loudspeaker signal $\mathbf{u}_{\mathrm{J_t}}(\kappa)$ and the source excitation signal $\mathbf{e}(\kappa)$ in the measurement equation, thus achieving the necessary requirements to employ a Kalman filter for the estimation of $\mathbf{F}_{\mathrm{t}}(\kappa)$. The linear minimum mean-square error (MMSE) estimate of the state vector $\mathbf{F}_{\mathrm{t}}(\kappa)$ corresponds to the solution of a Bayesian optimization problem [39, ch. 13], and is given by the well-known set of equations referred to as the (block) Kalman filter:

$$\mathbf{K}(\kappa) = \mathbf{P}(\kappa)\mathbf{C}^{\mathrm{H}}_{\mathrm{J_t}}(\kappa)\left[\mathbf{C}_{\mathrm{J_t}}(\kappa)\mathbf{P}(\kappa)\mathbf{C}^{\mathrm{H}}_{\mathrm{J_t}}(\kappa) + \mathbf{\Psi}_{\mathrm{EE}}(\kappa)\right]^{-1} \tag{24a}$$

$$\hat{\mathbf{F}}^{+}(\kappa) = \hat{\mathbf{F}}(\kappa) + \mathbf{K}(\kappa)[\mathbf{Y}_{\mathrm{J_t}}(\kappa) - \mathbf{C}_{\mathrm{J_t}}(\kappa)\hat{\mathbf{F}}(\kappa)] \tag{24b}$$

$$\mathbf{P}^{+}(\kappa) = [\mathbf{I}_{M\times M} - \mathbf{K}(\kappa)\mathbf{C}_{\mathrm{J_t}}(\kappa)]\,\mathbf{P}(\kappa) \tag{24c}$$

$$\hat{\mathbf{F}}(\kappa+1) = \alpha_{\mathrm{t}} \cdot \hat{\mathbf{F}}^{+}(\kappa) \tag{24d}$$

$$\mathbf{P}(\kappa+1) = \alpha_{\mathrm{t}}^2 \cdot \mathbf{P}^{+}(\kappa) + \mathbf{\Psi}_{\mathrm{N_tN_t}}(\kappa), \tag{24e}$$

where $\mathbf{K}(\kappa)$ is the frequency-domain Kalman gain, the superscript $^{+}$ indicates a posteriori estimates and, finally, $\mathbf{\Psi}_{\mathrm{EE}}(\kappa)$ and $\mathbf{\Psi}_{\mathrm{N_tN_t}}(\kappa)$ are the covariance matrices of $\mathbf{E}(\kappa)$ and $\mathbf{N}_{\mathrm{t}}(\kappa)$ [24], [35], respectively, assumed to be known.

For implementation purposes, we drop some of the assumptions initially made for the state space model, similarly to what

TABLE I
DEFINITIONS OF THE CONSTRAINT AND LINEARIZATION MATRICES USED IN THE PAPER, AS DEFINED IN BENESTY *et al.* [36]

Constraint rectangular matrix	Constraint square matrix	Linearization rectangular matrix	Linearization square matrix
$\mathbf{W}^{01}_{R\times M} = \begin{bmatrix}\mathbf{0}_{R\times R} & \mathbf{I}_{R\times R}\end{bmatrix}$	$\mathbf{W}^{01}_{M\times M} = \begin{bmatrix}\mathbf{0}_{R\times R} & \mathbf{0}_{R\times R}\\ \mathbf{0}_{R\times R} & \mathbf{I}_{R\times R}\end{bmatrix}$	$\mathbf{G}^{01}_{R\times M} = \mathbf{F}_R\,\mathbf{W}^{01}_{R\times M}\,\mathbf{F}^{-1}_M$	$\mathbf{G}^{01}_{M\times M} = (\mathbf{G}^{01}_{R\times M})^{\mathrm{H}}\mathbf{G}^{01}_{R\times M} = \mathbf{F}_M\,\mathbf{W}^{01}_{M\times M}\,\mathbf{F}^{-1}_M$
$\mathbf{W}^{01}_{M\times R} = \begin{bmatrix}\mathbf{I}_{R\times R}\\ \mathbf{0}_{R\times R}\end{bmatrix}$	$\mathbf{W}^{01}_{M\times M} = \begin{bmatrix}\mathbf{I}_{R\times R} & \mathbf{0}_{R\times R}\\ \mathbf{0}_{R\times R} & \mathbf{0}_{R\times R}\end{bmatrix}$	$\mathbf{G}^{10}_{M\times R} = \mathbf{F}_M\,\mathbf{W}^{01}_{M\times R}\,\mathbf{F}^{-1}_R$	$\mathbf{G}^{10}_{M\times M} = \mathbf{G}^{10}_{M\times R}(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}} = \mathbf{F}_M\,\mathbf{W}^{01}_{M\times M}\,\mathbf{F}^{-1}_M$

is done in the PEM-FDAF in order to carry out the optimization in an alternating fashion [22]:

1) The prefiltering operation is performed by means of the estimated $\hat{J}(q,\kappa)$, instead of the true $J_{\mathrm{t}}(q,\kappa)$; therefore, the prefiltered variables will be, from now on, indicated using the subscript $_{\hat{\mathrm{J}}}$, instead of $_{\mathrm{J_t}}$. Following a common assumption found in literature [22], [40], [41], the system $H_{\mathrm{t}}(q,\kappa)$ generating the source signal $v(t)$ is assumed to be time-varying, monic, inversely stable, and AR. The $J_{\mathrm{t}}(q,\kappa)$ is then estimated as the linear prediction filter for the time-domain error signal frame, i.e., $\mathbf{d}[\kappa|\hat{\mathbf{f}}(\kappa)] = \left[d[\kappa R - R + 1|\hat{\mathbf{f}}(\kappa)]\ \ldots\ d[\kappa R|\hat{\mathbf{f}}(\kappa)]\right]^{\mathrm{T}}$ at the current and the previous frame, i.e., $[\mathbf{d}^{\mathrm{T}}[\kappa|\hat{\mathbf{f}}(\kappa)]\ \ \mathbf{d}^{\mathrm{T}}[\kappa - 1|\hat{\mathbf{f}}(\kappa-1)]]^{\mathrm{T}}$, using the Levinson-Durbin algorithm [42, pp. 254-264] and represented by $\hat{\mathbf{j}}(\kappa)$.
2) The use of estimated prewhitening filter parameters motivates the replacement of $[\mathbf{Y}_{\mathrm{J_t}}(\kappa) - \mathbf{C}_{\mathrm{J_t}}(\kappa)\hat{\mathbf{F}}(\kappa)]$ in (24b) by the frequency-domain PE frame $\boldsymbol{\mathcal{E}}[\kappa,\hat{\boldsymbol{\Theta}}(\kappa)]$, related to the time-domain PE frame $\boldsymbol{\varepsilon}[\kappa,\hat{\boldsymbol{\theta}}(\kappa)]$ via

$$\boldsymbol{\varepsilon}[\kappa,\hat{\boldsymbol{\theta}}(\kappa)] = \left[\boldsymbol{\varepsilon}[\kappa R - R + 1,\hat{\boldsymbol{\theta}}(\kappa)]\ \ldots\ \boldsymbol{\varepsilon}[\kappa R,\hat{\boldsymbol{\theta}}(\kappa)]\right]^{\mathrm{T}} \tag{25}$$

$$\boldsymbol{\mathcal{E}}[\kappa,\hat{\boldsymbol{\Theta}}(\kappa)] = (\mathbf{G}^{01}_{R\times M})^{\mathrm{H}}\mathbf{F}_R\,\boldsymbol{\varepsilon}[\kappa,\hat{\boldsymbol{\theta}}(\kappa)] \tag{26}$$

with $\hat{\boldsymbol{\Theta}}(\kappa)$ containing the frequency-domain versions of $\hat{\mathbf{f}}(\kappa)$ and $\hat{\mathbf{j}}(\kappa)$ [cf. (46) to (50)]. Furthermore, we add the linearization constraint $\mathbf{G}^{10}_{M\times M}$ (cf. Table I) in (24b) [cf. (31b)], similarly to what is done in the FDAF, given the improved sound quality provided by a constrained FDAF version [32]. Naturally, such a constraint causes a computational complexity increase.

Additionally, we use the approximations introduced by Enzner and Vary [24], in order to address both the problem of high computational complexity and the possible ill-posedness of the solution, via diagonal operations, as follows:

1) The linearization square matrix $\mathbf{G}^{01}_{M\times M}$ in the definition of $\mathbf{C}_{\hat{\mathrm{J}}}(\kappa)$ is approximated by a diagonal matrix, allowing to write $\mathbf{G}^{01}_{M\times M} \approx (R/M)\mathbf{I}_{M\times M}$ and, less intuitively, $\mathbf{G}^{01}_{M\times M}\boldsymbol{\Delta}(\mathbf{G}^{01}_{M\times M})^{\mathrm{H}} \approx (R/M)\boldsymbol{\Delta}$, if $\boldsymbol{\Delta}$ is a diagonal matrix [36, ch. 8].
2) The covariance matrices $\boldsymbol{\Psi}_{\mathrm{EE}}(\kappa)$ and $\boldsymbol{\Psi}_{\mathrm{N_tN_t}}(\kappa)$ are replaced by the estimates $\boldsymbol{\Psi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)$ and $\boldsymbol{\Psi}_{\hat{\mathrm{N}}\hat{\mathrm{N}}}(\kappa)$, respectively, which are assumed to be diagonal [24]. These are related to the corresponding time-varying power spectral densities $\boldsymbol{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)$ and $\boldsymbol{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}}}(\kappa)$, adaptively estimated using the procedures described in [35], [38], [43], assuming that the feedback path is slowly time varying.
3) Given the assumed low correlation between different frequency components of the estimation error [24], the matrix $\mathbf{P}(\kappa)$ is a nearly diagonal matrix; the diagonalilty is enforced by initalizing $\mathbf{P}(0) \propto \mathbf{I}_{M\times M}$.

The last three approximations are used to write the following simplified expressions:

$$\mathbf{C}_{\hat{\mathrm{j}}}(\kappa) \approx \frac{R}{M}\cdot\mathbf{U}_{\hat{\mathrm{j}}}(\kappa) \tag{27}$$

$$\mathbf{C}_{\hat{\mathrm{j}}}(\kappa)\mathbf{P}(\kappa)\mathbf{C}^{\mathrm{H}}_{\hat{\mathrm{j}}}(\kappa) \approx \frac{R}{M}\cdot\mathbf{U}_{\hat{\mathrm{j}}}(\kappa)\mathbf{P}(\kappa)\mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{j}}}(\kappa) \tag{28}$$

$$\boldsymbol{\Psi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa) \approx R\cdot\mathrm{diag}\{\boldsymbol{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)\} \tag{29}$$

$$\boldsymbol{\Psi}_{\hat{\mathrm{N}}\hat{\mathrm{N}}}(\kappa) \approx M\cdot\mathrm{diag}\{\boldsymbol{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}}}(\kappa)\}. \tag{30}$$

With the approximations discussed so far, the complete set of equations describing the PEM-FDKF update is as follows:

$$\mathbf{K}(\kappa) = \mathbf{P}(\kappa)\mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{j}}}(\kappa)\left[\mathbf{U}_{\hat{\mathrm{j}}}(\kappa)\mathbf{P}(\kappa)\mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{j}}}(\kappa) + M\cdot\mathrm{diag}\{\boldsymbol{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)\}\right]^{-1} \tag{31a}$$

$$\hat{\mathbf{F}}^{+}(\kappa) = \hat{\mathbf{F}}(\kappa) + \mathbf{G}^{10}_{M\times M}\mathbf{K}(\kappa)\boldsymbol{\mathcal{E}}[\kappa,\hat{\boldsymbol{\Theta}}(\kappa)] \tag{31b}$$

$$\mathbf{P}^{+}(\kappa) = \left[\mathbf{I}_{M\times M} - \frac{R}{M}\mathbf{K}(\kappa)\mathbf{U}_{\hat{\mathrm{j}}}(\kappa)\right]\mathbf{P}(\kappa) \tag{31c}$$

$$\hat{\mathbf{F}}(\kappa+1) = \alpha\cdot\hat{\mathbf{F}}^{+}(\kappa) \tag{31d}$$

$$\mathbf{P}(\kappa+1) = \alpha^2\cdot\mathbf{P}^{+}(\kappa) + M\cdot\mathrm{diag}\{\boldsymbol{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}}}(\kappa)\}. \tag{31e}$$

IV. PEM-FDKF IDENTIFIABILITY CONDITIONS

The ICs for the optimization problem solved by the PEM-FDKF will be derived in three steps, as follows.

The first step, providing ICs for (7) and (14), is similar to the derivation in [3], but now introducing time variability in the signal models. The following expressions for $y(t)$ and $u(t)$ hold (cf. Fig. 2):

$$y(t) = F_{\mathrm{t}}(q,t)u(t) + v(t) \tag{32}$$

$$u(t) = G(q)\left[y(t) - \hat{F}(q,t)u(t)\right], \tag{33}$$

where $G(q) = q^{-d_G}\bar{G}(q)$ and $d_G \geq 1$, i.e., the forward path has to have at least a one-sample delay to avoid a delay-less loop.

Using (32) and (33), the PE in (13) can be rewritten as

$$\begin{aligned}\varepsilon[t,\boldsymbol{\xi}(t)] &= A(q,t)v(t) + [A(q,t)F_{\mathrm{t}}(q,t) + B(q,t)] \times u(t) \\ &= A(q,t)v(t) + [A(q,t)F_{\mathrm{t}}(q,t) + B(q,t)] \\ &\quad \times q^{-d_G}\bar{G}(q)\left[y(t) - \hat{F}(q,t)u(t)\right]. \end{aligned} \tag{34}$$

The PE $\varepsilon[t,\boldsymbol{\xi}(t)]$ can be expressed as a function of $v(t)$ only, by using (32) and (33) repeatedly, in (34), i.e.,

$$\begin{aligned}\varepsilon[t,\boldsymbol{\xi}(t)] &= A(q,t)v(t) + Z(q,t)\bar{G}(q)v(t-d_G) \\ &\quad + Z(q,t)\sum_{\ell=2}^{\infty}\bar{G}^{\ell}(q)\left[\prod_{i=1}^{\ell-1}F_{\mathrm{r}}(q,t-id_G)\right] \\ &\quad \times v(t-\ell d_G) \end{aligned} \tag{35}$$

where

$$F_{\mathrm{r}}(q,t) = F_{\mathrm{t}}(q,t) - \hat{F}(q,t) \tag{36}$$

$$Z(q,t) = A(q,t)F_{\mathrm{t}}(q,t) + B(q,t). \tag{37}$$

As is done for the time-invariant case in [3], we consider the sufficient order condition for $A(q,t)$ and $B(q,t)$, i.e., $n_A \geq n_J$ and $n_B \geq n_J + n_F - 1$, and the causality of $\bar{G}(q)$, $F_{\mathrm{t}}(q,t)$, $J_{\mathrm{t}}(q,t)$ and $\hat{F}(q,t)$ and study the conditions under which the minimization in (14a) leads to the unique solution

$$A(q,t) = J_{\mathrm{t}}(q,t) \tag{38}$$

$$B(q,t) = -J_{\mathrm{t}}(q,t)F_{\mathrm{t}}(q,t). \tag{39}$$

This solution satisfies (14b), hence making this constraint indeed redundant.

The unique desired solution can be derived if at least one of the following conditions, similar to the conditions found in the time-invariant case [3], is fulfilled:

C1: The forward path delay d_G satisfies $d_G \geq n_A$;

C2: The cancellation path delay d_F, where $B(q,t) = q^{-d_F}\bar{B}(q,t)$ and hence $F_{\mathrm{t}}(q,t) = q^{-d_F}\bar{F}_{\mathrm{t}}(q,t)$, satisfies $d_G + d_F \geq n_A$;

C3: The TF $\bar{G}(q)$ is nonlinear.

The proofs resemble those in [3], with the difference that the assumed time variability of the signal models does not allow to compactly rewrite (35). Specifically, the first condition, C1, turns the minimization in (14) into a linear prediction of $v(t)$, i.e., $A(q,t) = J_{\mathrm{t}}(q,t)$, given that $a_0(t) = 1$. Additionally, since $n_A \geq n_J$, the equation $Z(q,t) = 0$ must hold, leading to $B(q,t) = -J_{\mathrm{t}}(q,t)F_{\mathrm{t}}(q,t)$. The second condition, C2, can be proved similarly, by including $B(q,t) = q^{-d_F}\bar{B}(q,t)$, and hence $F_{\mathrm{t}}(q,t) = q^{-d_F}\bar{F}_{\mathrm{t}}(q,t)$, in (35), i.e.,

$$\begin{aligned}\varepsilon[t,\boldsymbol{\xi}(t)] &= A(q,t)v(t) + q^{-d_G-d_F}\left[A(q,t)\bar{F}_{\mathrm{t}}(q,t) + \bar{B}(q,t)\right] \\ &\quad \times \bar{G}(q)\left[y(t) - \hat{F}(q,t)u(t)\right]. \end{aligned} \tag{40}$$

If $d_G + d_F \geq n_A$, the unique solution of (40) is given by (38) and (39). The last condition, C3, can be understood considering that a nonlinear $\bar{G}(q)$ introduces additional decorrelation between the first and the other terms in (35), thus decoupling these terms in the minimization of (7). Such decoupled minimization yields, again, the values of $A(q,t)$ and $B(q,t)$ in (38) and (39). Following these results, which hold for any time-varying behavior of $J_{\mathrm{t}}(q,t)$ and $F_{\mathrm{t}}(q,t)$ fulfilling the initial assumptions as well as the ICs, we can go back to the simplified notation of the unconstrained optimization problem in (7), parameterized in $\boldsymbol{\theta}(t)$. Under the same ICs, the minimization in (7) then leads to the solution $J(q,t) = J_{\mathrm{t}}(q,t)$ and $F(q,t) = F_{\mathrm{t}}(q,t)$ [the equivalent of (38) and (39)].

The second step describes the transition from the current optimization problem (7) to a new optimization problem including the specific model (23), for which then the Kalman filter (alternating with the Levinson-Durbin algorithm) is seen to provide a suitable algorithm, and derive the ICs for this new optimization problem. We start by expressing (7) as a length-R frame-based expression, through the frame index κ. Assuming time invariance of $F_{\mathrm{t}}(q,t)$ and $H_{\mathrm{t}}(q,t)$ over each frame, the sample-based estimation of the parameter vector $\boldsymbol{\theta}(t)$ is replaced by the frame-based estimation of the parameter vector $\boldsymbol{\theta}(\kappa)$, found by solving at each κ

$$\min_{\boldsymbol{\theta}(\kappa)} \quad \mathbb{E}\{\|\boldsymbol{\varepsilon}[\kappa,\boldsymbol{\theta}(\kappa)]\|^2_{\boldsymbol{\zeta}^{-1}(\kappa)}\}, \tag{41}$$

where $\|\mathbf{z}\|_{\mathbf{W}} = \sqrt{\mathbf{z}^{\mathrm{H}}\mathbf{W}\mathbf{z}} = \left\|\mathbf{W}^{1/2}\mathbf{z}\right\|_2$ is the weighted norm of $\mathbf{z}$ induced by the positive definite matrix $\mathbf{W}$, and $\boldsymbol{\zeta}^{-1}(\kappa) = \zeta^{-1}(\kappa)\mathbf{I}_{R\times R}$ will be used to compensate for power variations in the excitation signal $\mathbf{e}(\kappa)$. $\boldsymbol{\varepsilon}[\kappa,\boldsymbol{\theta}(\kappa)]$ is the length-R PE frame obtained from (6b) assuming the parameter vector $\boldsymbol{\theta}(\kappa)$ is constant in $t = \kappa R - R + 1,\ldots,\kappa R$. Under the same time-invariance assumption, it has been pointed out in [32] that minimizing the frame-based cost function vs. the sample-based cost function leads to the same mean-square error performance; therefore, the same ICs hold for the (41) and (7).

We now assume a frame-based state-space model for the true system using a simple Markov model to describe the state (i.e., the true feedback path) dynamics [12], and the linear model for the time-domain prefiltered microphone signal frame $\mathbf{y}_{\mathrm{J_t}}(t)$ as the measurement equation:

$$\mathbf{f}_{\mathrm{t}}(\kappa+1) = \alpha_{\mathrm{t}}\mathbf{f}_{\mathrm{t}}(\kappa) + \mathbf{n}_{\mathrm{t}}(\kappa) \tag{42a}$$

$$\mathbf{y}_{\mathrm{J_t}}(\kappa) = F_{\mathrm{t}}(q,\kappa)\mathbf{u}_{\mathrm{J_t}}(\kappa) + \mathbf{e}(\kappa), \tag{42b}$$

cf. (16), (19) and (20), and we describe the random variables of (42) as $\mathbf{n}_{\mathrm{t}}(\kappa) \sim \mathcal{N}(0,\boldsymbol{\Lambda}_{\mathrm{t}}(\kappa))$, $\mathbf{e}(\kappa) \sim \mathcal{N}(0,\boldsymbol{\Sigma}_{\mathrm{t}}(\kappa))$ and $\mathbf{f}_{\mathrm{t}}(0) \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{f}_{\mathrm{t}}(0)},\boldsymbol{\Pi}_{\mathrm{t}}(0))$, where $\mathcal{N}(\cdot,\cdot)$ indicates a normal distribution with specified mean and covariance.

The Kalman filter corresponding to the state-space model (42) effectively solves the optimization problem [44]–[46]

$$\begin{aligned}\min_{\{\mathbf{f}(n),\,\mathbf{n}(n)\}_{n=0}^{\kappa}} \quad & \frac{1}{2}\left\|\mathbf{f}(0) - \boldsymbol{\mu}_{\mathbf{f}_{\mathrm{t}}(0)}\right\|^2_{\boldsymbol{\Pi}_{\mathrm{t}}^{-1}(0)} \\ & + \frac{1}{2}\sum_{n=0}^{\kappa-1}\|\mathbf{n}(n)\|^2_{\boldsymbol{\Lambda}_{\mathrm{t}}^{-1}(n)} \\ & + \frac{1}{2}\sum_{n=0}^{\kappa}\|\boldsymbol{\varepsilon}[n,\boldsymbol{\theta}(n)]\|^2_{\boldsymbol{\Sigma}_{\mathrm{t}}^{-1}(n)} \end{aligned} \tag{43a}$$

$$\text{subject to} \qquad \mathbf{f}(n+1) = \alpha_{\mathrm{t}}\mathbf{f}(n) + \mathbf{n}(n). \tag{43b}$$

This formulation combines the different optimization problems in (41) for successive frames $n = 0, \ldots, \kappa$ into a single optimization framework, with the constraint (43b) defining the time evolution of $\mathbf{f}(n)$, and three terms characterizing the cost function:

1) A regularization term depending on $\mathbf{f}(0)$ and on the matrix $\mathbf{\Pi}_{\mathrm{t}}(0) = \mathbb{E}\{\left[\mathbf{f}(0) - \boldsymbol{\mu}_{\mathbf{f}_{\mathrm{t}}(0)}\right]\left[\mathbf{f}(0) - \boldsymbol{\mu}_{\mathbf{f}_{\mathrm{t}}(0)}\right]^{\mathrm{T}}\}$, with $\boldsymbol{\mu}_{\mathbf{f}_{\mathrm{t}}(0)}$ an initial guess of the initial state $\mathbf{f}(0)$;
2) A term depending on the unknown state noise process $\mathbf{n}(n)$, with $\mathbf{n}(n)$ being a new variable to optimize, and on $\mathbf{\Lambda}_{\mathrm{t}}(n) = \mathbb{E}\{\mathbf{n}(n)\mathbf{n}(n)^{\mathrm{T}}\}$;
3) A term depending on the PE $\boldsymbol{\varepsilon}[n, \boldsymbol{\theta}(n)]$, similar to the term in (41), where the expectation on the single frame is replaced by a summation over the different frames, and $\boldsymbol{\zeta}^{-1}(\kappa)$ in (41) is replaced by the matrix $\mathbf{\Sigma}_{\mathrm{t}}(n)$ [44], [47].

With (43) effectively $\mathbf{f}(n)$ is estimated under the assumption that $\mathbf{j}(n)$ is known [as $\mathbf{j}(n)$ is included in $\boldsymbol{\theta}(n)$ in the third term of (43a)]. If $\mathbf{j}(n) = \mathbf{j}_{\mathrm{t}}(n)$, then the Kalman filter (subject to technical full-rank conditions) provides the unique MMSE estimate of $\mathbf{f}_{\mathrm{t}}(\kappa)$, which itself (subject to the above ICs) is included in the unique minimizer of the third term of (43a) with the expectation reintroduced.

When the Kalman Filter operations [to estimate $\mathbf{f}(n)$] are alternated with the Levinson-Durbin algorithm [to estimate $\mathbf{j}(n)$], as in Section III, the optimization problem that is effectively solved is

$$\min_{\{\boldsymbol{\theta}(n),\, \mathbf{n}(n)\}_{n=0}^{\kappa}} \quad \frac{1}{2}\left\|\mathbf{f}(0) - \boldsymbol{\mu}_{\mathbf{f}_{\mathrm{t}}(0)}\right\|^2_{\mathbf{\Pi}_{\mathrm{t}}^{-1}(0)} + \frac{1}{2}\sum_{n=0}^{\kappa-1}\|\mathbf{n}(n)\|^2_{\mathbf{\Lambda}_{\mathrm{t}}^{-1}(n)} + \frac{1}{2}\sum_{n=0}^{\kappa}\|\boldsymbol{\varepsilon}[n, \boldsymbol{\theta}(n)]\|^2_{\mathbf{\Sigma}_{\mathrm{t}}^{-1}(n)} \tag{44a}$$

$$\text{subject to} \quad \mathbf{f}(n+1) = \alpha_{\mathrm{t}}\mathbf{f}(n) + \mathbf{n}(n). \tag{44b}$$

The alternation minimization then provides a (possibly suboptimal) estimate of $\boldsymbol{\theta}_{\mathrm{t}}(\kappa) = [\mathbf{f}_{\mathrm{t}}^{\mathrm{T}}(\kappa)\, \mathbf{j}_{\mathrm{t}}^{\mathrm{T}}(\kappa)]^{\mathrm{T}}$, which itself (subject to the above ICs) is included in the unique minimizer of the third term of (44a) with the expectation reintroduced. If the Kalman filter is applied with $\mathbf{j}(n) = \mathbf{j}_{\mathrm{t}}(n)$, then (subject to technical full-rank conditions) it provides the unique MMSE estimate of $\mathbf{f}_{\mathrm{t}}(\kappa)$. Similarly, if the Levinson-Durbin algorithm is applied with $\mathbf{f}(n) = \mathbf{f}_{\mathrm{t}}(n)$ it provides the unique MMSE estimate of $\mathbf{j}_{\mathrm{t}}(\kappa)$.

The third step is that of formulating the optimization problem in the frequency domain, i.e.,

$$\min_{\{\mathbf{\Theta}(n),\, \mathbf{N}(n)\}_{n=0}^{\kappa}} \quad \frac{1}{2}\left\|\mathbf{F}(0) - \boldsymbol{\mu}_{\mathbf{F}_{\mathrm{t}}(0)}\right\|^2_{\mathbf{P}_{\mathrm{t}}^{-1}(0)} + \frac{1}{2}\sum_{n=0}^{\kappa-1}\|\mathbf{N}(n)\|^2_{\mathbf{L}_{\mathrm{t}}^{-1}(n)} + \frac{1}{2}\sum_{n=0}^{\kappa}\|\boldsymbol{\mathcal{E}}[n, \mathbf{\Theta}(n)]\|^2_{\mathbf{S}_{\mathrm{t}}^{-1}(n)} \tag{45a}$$

$$\text{subject to} \quad \mathbf{F}(n+1) = \alpha_{\mathrm{t}}\mathbf{F}(n) + \mathbf{N}(n). \tag{45b}$$

We will now show that the ICs of the frequency-domain problem in (45) correspond to those derived for the frame-based time-domain problem in (44). To this end, we introduce some suitable variable transformations in the optimization problem; namely, the different frequency-domain variables are related to their time-domain counterparts using different constraint matrices defined in Table I [36], as follows:

$$\mathbf{f}(n) = \mathbf{F}_R^{-1}(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}}\,\mathbf{F}(n) \tag{46}$$

$$\mathbf{F}(n) = \mathbf{G}^{10}_{M\times R}\mathbf{F}_R\mathbf{f}(n) \tag{47}$$

$$\mathbf{j}(n) = \mathbf{F}_{n_J}^{-1}(\mathbf{G}^{10}_{M\times n_J})^{\mathrm{H}}\mathbf{J}(n) \tag{48}$$

$$\mathbf{J}(n) = \mathbf{G}^{10}_{M\times n_J}\mathbf{F}_{n_J}\mathbf{j}(n) \tag{49}$$

$$\mathbf{\Theta}(n) = [\mathbf{F}^{\mathrm{T}}(n) \quad \mathbf{J}^{\mathrm{T}}(n)]^{\mathrm{T}} \tag{50}$$

$$\mathbf{n}(n) = \mathbf{F}_R^{-1}(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}}\,\mathbf{N}(n), \tag{51}$$

$$\mathbf{N}(n) = \mathbf{G}^{10}_{M\times R}\mathbf{F}_R\mathbf{n}(n). \tag{52}$$

Here we have assumed that $\mathbf{f}(n)$, $\mathbf{j}(n)$ and $\mathbf{n}(n)$ have lengths R, n_J and R, respectively, while all the frequency-domain variables have length M.

In addition to the transformations (47), (49) and (52), the four steps necessary to rewrite (45) as (44) are the following:

1) The frequency-domain and time-domain feedback path model and process noise are related via (47) and (52); substituting (47) and (52) in (45b) and premultiplying with $\mathbf{F}_R^{-1}(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}}$ leads to (44b).
2) The first term of (45a) can be rewritten in terms of $\mathbf{f}(0)$ and $\mathbf{\Pi}_{\mathrm{t}}^{-1}(0)$ using (47) and the following relation:
$$(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}}\,\mathbf{P}_{\mathrm{t}}^{-1}(0) = \mathbf{F}_R\mathbf{\Pi}_{\mathrm{t}}^{-1}(0)\mathbf{F}_R^{-1}(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}}, \tag{53}$$
resulting in the first term of (44a).
3) The second term of (45a) can be rewritten in terms of $\mathbf{n}(n)$ and $\mathbf{\Lambda}_{\mathrm{t}}^{-1}(n)$ using (52) and the following relation:
$$(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}}\,\mathbf{L}_{\mathrm{t}}^{-1}(n) = \mathbf{F}_R\mathbf{\Lambda}_{\mathrm{t}}^{-1}(n)\mathbf{F}_R^{-1}(\mathbf{G}^{10}_{M\times R})^{\mathrm{H}}, \tag{54}$$
resulting in the second term of (44a).
4) The third term of (45a) can be rewritten in terms of $\boldsymbol{\varepsilon}[n, \boldsymbol{\theta}(n)]$ and $\mathbf{\Sigma}_{\mathrm{t}}^{-1}(n)$, using the following relations:
$$\boldsymbol{\mathcal{E}}[n, \mathbf{\Theta}(n)] = (\mathbf{G}^{01}_{R\times M})^{\mathrm{H}}\mathbf{F}_R\boldsymbol{\varepsilon}[n, \boldsymbol{\theta}(n)] \tag{55}$$
$$\mathbf{G}^{01}_{R\times M}\mathbf{S}_{\mathrm{t}}^{-1}(n) = \mathbf{F}_R\mathbf{\Sigma}_{\mathrm{t}}^{-1}(n)\mathbf{F}_R^{-1}\mathbf{G}^{10}RM, \tag{56}$$
resulting in the third term of (44a).

Overall, the transformations (47), (49) and (52) to (56) can be used to link the frequency-domain problem and solutions in (45) to the frame-based time-domain problem and solutions in (44). Thus the ICs C1 to C3 and the considerations drawn for the frame-based time-domain problem (44) hold for the frequency-domain problem (45), too.

V. The PEM-Based Partitioned-Block Frequency Domain Kalman Filter (PEM-PBFDKF)

Even though the use of FDAF algorithms has been shown to be beneficial compared to time-domain algorithms for several

aspects, an important issue that might limit the applicability of FDAF algorithms is the use of excessive filter lengths. High-order filters, motivated by long echo or feedback paths and/or high sampling frequencies, can lead to algorithmic noise [43]. Additionally, high-order filters increase the algorithmic delay, potentially making real-time solutions unfeasible. Even though the feedback path of a HA is usually relatively short, the very low delay requirements generally make FDAF algorithms unsuitable for HA applications.

A way to overcome this problem involves the use of a PB structure, i.e., the so-called partitioned-block frequency-domain adaptive filter (PBFDAF) [33], [37]. The PBFDAF requires the division of the feedback path model into P partitions of length $L \leq R = n_F$, thus allowing to lower the algorithmic delay from $2n_F - 1$ to $2L - 1$ [22]. The PBFDAF has been successfully applied in both AEC [34], [48] and AFC [22], [49], [50]. Specifically, for AFC, the PBFDAF has been combined with a PEM-based prewhitening filter, giving rise to the (PEM-PBFDAF) [22], [23]. More recently, a state-space version of the PBFDAF algorithm for AEC has been proposed [38], [43], which will be referred to as partitioned-block frequency-domain Kalman filter (PBFDKF) in the following.

In this section, we propose a modified version of the PBFDKF including the same decorrelation stage employed for the FDKF presented in Section III, by means of a PEM-based prewhitening, i.e., the PEM-PBFDKF.

We first define the partitioned version of the time-domain M-samples loudspeaker signal and the L-samples true feedback path, at frame κ for block $p = 0, \ldots, P-1$, as follows

$$\mathbf{u}_p(\kappa) = [u(\kappa R - pL - M + 1) \quad \ldots \quad u(\kappa R - pL)]^{\mathrm{T}} \tag{57}$$

$$\mathbf{f}_{\mathrm{t},p}(\kappa) = [f_{\mathrm{t}}(pL, \kappa) \quad \ldots \quad f_{\mathrm{t}}(pL + L - 1, \kappa)]^{\mathrm{T}}. \tag{58}$$

The partitioned signal vectors can be defined in a similar way to the non-partitioned ones, adding the specific block index; in the following definitions, we are always considering the time frame κ for block p. The time- and frequency-domain version of the prefiltered loudspeaker signal can be defined as follows:

$$\mathbf{u}_{\mathrm{J_t},p}(\kappa) = [u_{\mathrm{J_t}}(\kappa R - pL - M + 1) \quad \ldots \quad u_{\mathrm{J_t}}(\kappa R - pL)]^{\mathrm{T}} \tag{59}$$

$$\mathbf{U}_{\mathrm{J_t},p}(\kappa) = \mathrm{diag}\left\{\mathbf{F}_M \mathbf{u}_{\mathrm{J_t},p}(\kappa)\right\}, \tag{60}$$

with the constraint $M \geq R + L - 1$, to ensure proper operations [22]. The partitioned-block frequency-domain (PBFD) representation of the true feedback path can be defined, similarly to (18a), as follows:

$$\mathbf{F}_{\mathrm{t},p}(\kappa) = \mathbf{G}^{10}_{M \times L} \mathbf{F}_L \mathbf{f}_{\mathrm{t},p}(\kappa). \tag{61}$$

Finally, the time- and frequency-domain versions of the prefiltered microphone signal and source signal frame can be defined as follows, with $V = M - L$:

$$\mathbf{y}_{\mathrm{J_t}}(\kappa) = [y_{\mathrm{J_t}}(\kappa R - V + 1) \quad \ldots \quad y_{\mathrm{J_t}}(\kappa R)]^{\mathrm{T}} \tag{62}$$

$$\mathbf{Y}_{\mathrm{J_t}}(\kappa) = (\mathbf{G}^{01}_{V \times M})^{\mathrm{H}} \mathbf{F}_V \mathbf{y}_{\mathrm{J_t}}(\kappa) \tag{63}$$

$$\mathbf{e}(\kappa) = [e(\kappa R - V + 1) \quad \ldots \quad e(\kappa R)]^{\mathrm{T}} \tag{64}$$

$$\mathbf{E}(\kappa) = (\mathbf{G}^{01}_{V \times M})^{\mathrm{H}} \mathbf{F}_V \mathbf{e}(\kappa). \tag{65}$$

As in the non-partitioned case, one of the $M - L + 1 = V + 1$ samples from the fast frequency-domain linear convolution is dropped to simplify the notation, resulting in a V-samples length for both $\mathbf{y}_{\mathrm{J_t}}(\kappa)$ and $\mathbf{e}(\kappa)$ [38]. A common choice [33], [36], [38] for the parameters is $L = V = M/2$, recalling that now $L = n_F/P$. The frame shift R and the signal block length V can be related via $R = V/\gamma$, where γ is the overlapping factor, usually chosen to be an integer [33].

The resulting PB state-space model is the following:

$$\mathbf{F}_{\mathrm{t},p}(\kappa + 1) = \alpha_{\mathrm{t}} \mathbf{F}_{\mathrm{t},p}(\kappa) + \mathbf{N}_{\mathrm{t},p}(\kappa), \quad p = 0, \ldots, P-1 \tag{66a}$$

$$\mathbf{Y}_{\mathrm{J_t}}(\kappa) = \sum_{p=0}^{P-1} \mathbf{C}_{\mathrm{J_t},p}(\kappa) \mathbf{F}_{\mathrm{t},p}(\kappa) + \mathbf{E}(\kappa) \tag{66b}$$

where $\mathbf{C}_{\mathrm{J_t},p}(\kappa) = \mathbf{G}^{01}_{M \times M} \mathbf{U}_{\mathrm{J_t},p}(\kappa)$ is used to linearize the pth circular convolution between the partitions $\mathbf{u}_{\mathrm{J_t},p}(\kappa)$ and $\mathbf{f}_{\mathrm{t},p}(\kappa)$, similarly to the non-partitioned case. The partitioning of the state equation (66a) requires the definition of $\mathbf{N}_{\mathrm{t},p}(\kappa)$, i.e., the process noise for the pth partition; the transition factor α_{t}, however, is still partition invariant [38], [43].

As for the non-partitioned case, we can apply Kalman filter to the model defined in (66) to obtain the linear MMSE estimate of the state $\mathbf{F}_{\mathrm{t},p}(\kappa)$; a set of equations very similar to (24) can be written as follows, for each partition $p = 0, \ldots, P-1$:

$$\mathbf{K}_p(\kappa) = \mathbf{P}_p(\kappa) \mathbf{C}^{\mathrm{H}}_{\mathrm{J_t},p}(\kappa) \left[\mathbf{C}_{\mathrm{J_t},p}(\kappa) \mathbf{P}_p(\kappa) \mathbf{C}^{\mathrm{H}}_{\mathrm{J_t},p}(\kappa) + \mathbf{\Psi}_{\mathrm{EE}}(\kappa)\right]^{-1} \tag{67a}$$

$$\hat{\mathbf{F}}^{+}_p(\kappa) = \hat{\mathbf{F}}_p(\kappa) + \mathbf{K}_p(\kappa) \left[\mathbf{Y}_{\mathrm{J_t}}(\kappa) - \sum_{p=0}^{P-1} \mathbf{C}_{\mathrm{J_t},p}(\kappa) \hat{\mathbf{F}}_p(\kappa)\right] \tag{67b}$$

$$\mathbf{P}^{+}_p(\kappa) = \left[\mathbf{I}_{M \times M} - \mathbf{K}_p(\kappa) \mathbf{C}_{\hat{\mathrm{J}},p}(\kappa)\right] \mathbf{P}_p(\kappa) \tag{67c}$$

$$\hat{\mathbf{F}}_p(\kappa + 1) = \alpha_{\mathrm{t}} \hat{\mathbf{F}}^{+}_p(\kappa) \tag{67d}$$

$$\mathbf{P}_p(\kappa + 1) = \alpha^2_{\mathrm{t}} \mathbf{P}^{+}_p(\kappa) + \mathbf{\Psi}_{\mathrm{N_t N_t},p}(\kappa). \tag{67e}$$

The simplified form of (67) relies on similar approximations as in the non-partitioned case, where R is replaced by V, i.e., $\mathbf{G}^{01}_{M \times M} \approx (V/M)\mathbf{I}_{M \times M}$ and $\mathbf{G}^{01}_{M \times M} \mathbf{\Delta} (\mathbf{G}^{01}_{M \times M})^{\mathrm{H}} \approx (V/M)\mathbf{\Delta}$, if $\mathbf{\Delta}$ is a diagonal matrix [38], [43]. This allows to adapt the approximations listed in Section III accordingly, yielding the diagonalized version of the algorithm, i.e., the PEM-PBFDKF:

$$\mathbf{K}_p(\kappa) = \mathbf{P}_p(\kappa) \mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{J}},p}(\kappa) \left[\mathbf{U}_{\hat{\mathrm{J}},p}(\kappa) \mathbf{P}_p(\kappa) \mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{J}},p}(\kappa) + M\, \mathrm{diag}\{\mathbf{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)\}\right]^{-1} \tag{68a}$$

$$\hat{\mathbf{F}}^{+}_p(\kappa) = \hat{\mathbf{F}}_p(\kappa) + \mathbf{G}^{10}_{M \times M} \mathbf{K}_p(\kappa) \boldsymbol{\mathcal{E}}[\kappa, \hat{\mathbf{\Theta}}(\kappa)] \tag{68b}$$

Algorithm 1: PEM-based partitioned-block frequency-domain Kalman filter (PEM-PBFDKF). The PEM-FDKF can be obtained by setting $P = 1$.

$\mathbf{P}_p(0) = \sigma\mathbf{I}_{M\times M}, \hat{\mathbf{F}}_p(\kappa) = \mathbf{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}},p}(0) = \tilde{\mathbf{\Phi}}_{\hat{\mathrm{F}}\hat{\mathrm{F}},p}(0) = \mathbf{0}_{M\times 1}, \hat{\mathbf{j}}(0) = \begin{bmatrix} 1 & \mathbf{0}_{n_J-1\times 1} \end{bmatrix}^{\mathrm{T}}$
for $\kappa = 0, 1, 2, \ldots$ **do**
  for $p = 0, \ldots, P-1$ **do**
    $\mathbf{U}_p(\kappa) = \mathrm{diag}\{\mathbf{F}_M \mathbf{u}_p(\kappa)\},$
  end for
  $\hat{\mathbf{y}}(\kappa) = \mathbf{W}^{01}_{V\times M}\mathbf{F}^{-1}_M \sum_{p=0}^{P-1} \mathbf{U}_p(\kappa)\hat{\mathbf{F}}_p(\kappa)$
  $\mathbf{d}(\kappa) = \mathbf{y}(\kappa) - \hat{\mathbf{y}}(\kappa)$
  $u_{\hat{\mathrm{j}}}(\kappa R - pL - i) = \hat{J}(q,\kappa)u(\kappa R - pL - i)$
    $i = 0, \ldots, M-1,\ p = 0, \ldots, P-1$
  $y_{\hat{\mathrm{j}}}(\kappa R - i) = \hat{J}(q,\kappa)y(\kappa R - i),$
    $i = 0, \ldots, V-1$
  for $p = 0, \ldots, P-1$ **do**
    $\mathbf{U}_{\hat{\mathrm{j}},p}(\kappa) = \mathrm{diag}\left\{\mathbf{F}_M \mathbf{u}_{\hat{\mathrm{j}},p}(\kappa)\right\},$
  end for
  $\hat{\mathbf{y}}_{\hat{\mathrm{j}}}(\kappa) = \mathbf{W}^{01}_{V\times M}\mathbf{F}^{-1}_M \sum_{p=0}^{P-1} \mathbf{U}_{\hat{\mathrm{j}},p}(\kappa)\hat{\mathbf{F}}_p(\kappa)$
  $\boldsymbol{\varepsilon}[\kappa, \hat{\boldsymbol{\theta}}(\kappa)] = \mathbf{y}_{\hat{\mathrm{j}}}(\kappa) - \hat{\mathbf{y}}_{\hat{\mathrm{j}}}(\kappa)$
  $\boldsymbol{\mathcal{E}}[\kappa, \hat{\mathbf{\Theta}}(\kappa)] = \mathbf{F}_M(\mathbf{W}^{01}_{V\times M})^{\mathrm{H}}\boldsymbol{\varepsilon}[\kappa, \hat{\boldsymbol{\theta}}(\kappa)]$
  $\mathbf{\Phi}_{\hat{\mathcal{E}}\hat{\mathcal{E}}}(\kappa) = \mathrm{diag}\{\boldsymbol{\mathcal{E}}[\kappa, \hat{\mathbf{\Theta}}(\kappa)]\boldsymbol{\mathcal{E}}^{\mathrm{T}}[\kappa, \hat{\mathbf{\Theta}}(\kappa)]\}/V$
  $\mathbf{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa) = \mathbf{\Phi}_{\hat{\mathcal{E}}\hat{\mathcal{E}}}(\kappa) - \frac{\mathrm{diag}\{\sum_{p=0}^{P-1} \mathbf{U}_{\hat{\mathrm{j}},p}(\kappa)\mathbf{P}_p(\kappa)\mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{j}},p}(\kappa)\}}{M}$
  Threshold $\mathbf{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)$ with $\hat{\sigma}^2_{\boldsymbol{\varepsilon}}(\kappa)\mathbf{1}_{M\times 1}$
  for $p = 0, \ldots, P-1$ **do**
    $\mathbf{K}_p(\kappa) = \mathbf{P}_p(\kappa)\mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{j}},p}(\kappa) \left[\mathbf{U}_{\hat{\mathrm{j}},p}(\kappa)\mathbf{P}_p(\kappa)\mathbf{U}^{\mathrm{H}}_{\hat{\mathrm{j}},p}(\kappa) + M\ \mathrm{diag}\{\mathbf{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)\}\right]^{-1}$
    $\hat{\mathbf{F}}^+_p(\kappa) = \hat{\mathbf{F}}_p(\kappa) + \mathbf{G}^{10}_{M\times M}\mathbf{K}_p(\kappa)\boldsymbol{\mathcal{E}}[\kappa, \hat{\mathbf{\Theta}}(\kappa)]$
    $\hat{\mathbf{F}}_p(\kappa+1) = \alpha\hat{\mathbf{F}}^+_p(\kappa)$
    $\mathbf{P}^+_p(\kappa) = [\mathbf{I}_{M\times M} - \frac{V}{M}\mathbf{K}_p(\kappa)\mathbf{U}_{\hat{\mathrm{j}},p}(\kappa)]\mathbf{P}_p(\kappa)$
    Threshold $\mathbf{P}^+_p(\kappa)$ with $\mathbf{0}_{M\times M}$
    $\mathbf{P}_p(\kappa+1) = \alpha^2\mathbf{P}^+_p(\kappa) + M\ \mathrm{diag}\{\mathbf{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}},p}(\kappa)\}$
    $\tilde{\mathbf{\Phi}}_{\hat{\mathrm{F}}\hat{\mathrm{F}},p}(\kappa+1) = \beta\tilde{\mathbf{\Phi}}_{\hat{\mathrm{F}}\hat{\mathrm{F}},p}(\kappa) + \frac{(1-\beta)\mathrm{diag}\{\hat{\mathbf{F}}_p(\kappa)\hat{\mathbf{F}}^{\mathrm{H}}_p(\kappa)\}}{M}$
    $\mathbf{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}},p}(\kappa+1) = (1-\alpha^2)\tilde{\mathbf{\Phi}}_{\hat{\mathrm{F}}\hat{\mathrm{F}},p}(\kappa+1)$
  end for
  $\hat{\mathbf{j}}(\kappa+1) = \texttt{Levinson}-\texttt{Durbin}\left([\mathbf{d}(\kappa); \mathbf{d}(\kappa-1)]\right)$
end for

$$\mathbf{P}^+_p(\kappa) = \left[\mathbf{I}_{M\times M} - \frac{V}{M}\mathbf{K}_p(\kappa)\mathbf{U}_{\hat{\mathrm{j}},p}(\kappa)\right]\mathbf{P}_p(\kappa) \quad (68c)$$

$$\hat{\mathbf{F}}_p(\kappa+1) = \alpha\hat{\mathbf{F}}^+_p(\kappa) \quad (68d)$$

$$\mathbf{P}_p(\kappa+1) = \alpha^2\mathbf{P}^+_p(\kappa) + M\ \mathrm{diag}\{\mathbf{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}},p}(\kappa)\}. \quad (68e)$$

A summary of the PEM-PBFDKF is given in (1), where the explicit calculations of $\mathbf{\Phi}_{\hat{\mathrm{E}}\hat{\mathrm{E}}}(\kappa)$ and $\mathbf{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}}}(\kappa)$ follow the procedures described in [35], [38], [43]; in particular, $\mathbf{\Phi}_{\hat{\mathrm{N}}\hat{\mathrm{N}}}(\kappa)$ is estimated using a first-order recursive filter, with forgetting factor $\beta = 0.91$ [38]. Additionally, $\mathbf{1}_{M\times 1}$ is defined as an $M \times 1$ vector of ones.

Usually, in PBFDAF algorithms increasing the number of partitions reduces the convergence speed [51], since this increase results in smaller partitions, thus lowering the degree of decorrelation introduced when working in the frequency domain. However, this behavior is not always observed in the AFC case due to the closed-loop nature of the system. In a feedback scenario, if the system is effectively subject to or close to instability, the high power of the loudspeaker signal leads to a faster identification of the unknown feedback path and hence increases the convergence speed [12].

Finally, as pointed out by Buchner *et al.* [52], [53], PBFDAF algorithms relying on diagonal approximations require a stronger regularization than non-partitioned algorithms. For this purpose, in our implementation we introduce an additional thresholding operation in the calculation of $\mathbf{P}^+_p(\kappa)$, since the subtraction in (68c) may give rise to negative values in some frequency bins with low signal-to-noise ratio (SNR).

VI. Computational Complexity and Memory Requirements

In this section, we provide a complexity analysis of the proposed PEM-PBFDKF algorithm, in comparison with the PEM-PBFDAF [22], the FDAF [32], as well as the time-domain algorithms NLMS and PEM-AFROW [17], counting the number of per-output-sample real multiplications [32]. The following assumptions are made: a real multiplication and a real division have equal complexity; each length-M FFT/IFFT operation has a complexity of $D = M\log_2(M)$ multiplications [32]; the Levinson-Durbin algorithm on a length-M signal vector has a complexity of $n_J^2 + (5+M)n_J + M$ multiplications. Table II lists the per-output-sample computational complexity. The normalization by R in the frequency-domain algorithms is only included to simplify the comparison; in reality, the system implementing the algorithms has a time equivalent to R samples to carry out a whole algorithm iteration.

The complexity of the different algorithms as a function of the prefilter order n_J is shown at the top of Fig. 3. The results are obtained using $n_F = L = V = R = M/2 = 80/P$, $P = 1$ for the two non-partitioned algorithms (PEM-FDAF and PEM-FDKF), $P = \{2, 4\}$ for the two partitioned algorithms (PEM-PBFDAF and PEM-PBFDKF), and fixing the overlapping factor to $\gamma = 1$. A subscript is used to indicate the number of partitions and overlapping factor, respectively, when $P > 1$, e. g. PEM-PBFDAF$_{21}$ refers to the case with $P = 2$ and $\gamma = 1$.[1] For the PEM-AFROW, M and R are the window size and the hop size used to estimate the source signal model. The grayed part of the plots corresponds to values of n_J between 10 and 20, being common order values when using an AR source signal model for speech signals [22], [23], [40]. The value $n_J = 15$ is highlighted in the plot since it is used in the simulations presented in the following section. Using these values, the number

[1]For the sake of simplicity, the PB entries in Table II include also the non-partitioned algorithms, i.e., PEM-FDAF and PEM-FDKF, using $P = 1$.

TABLE II
PER-OUTPUT-SAMPLE COMPUTATIONAL COMPLEXITY AND MEMORY REQUIREMENTS OF THE COMPARED ALGORITHMS

Algorithm	Computational complexity	#	Memory requirements	#
NLMS	$3R+2$	322	$4R+7$	328
PEM-AFROW	$\frac{n_J^2+(M+3R+3)n_J+(1+R)M+2R(2R+3)}{R}$	568	$5R+n_H+15$	430
FDAF	$\frac{5D+13M}{R}$	99	$20M+R+4$	3284
PEM-PBFDAF	$\frac{(4P+3)D+3n_J^2+(2M+R+7)n_J+(2+17P)M+R}{R}$	226 | 307 | 457	$8M+3R+n_J+2(P-1)L+19PM+4$	4579 | 3899 | 3559
PEM-PBFDKF	$\frac{(4P+3)D+3n_J^2+(2M+R+7)n_J+(5+25P)M+R}{R}$	248 | 345 | 527	$10M+3R+n_J+2(P-1)L+31PM+5$	6820 | 5980 | 5560

A numerical value is given in both cases for $n_F = L = V = R = M/2 = 80/P$, $n_J = 15$, and $P = \{1, 2, 4\}$.

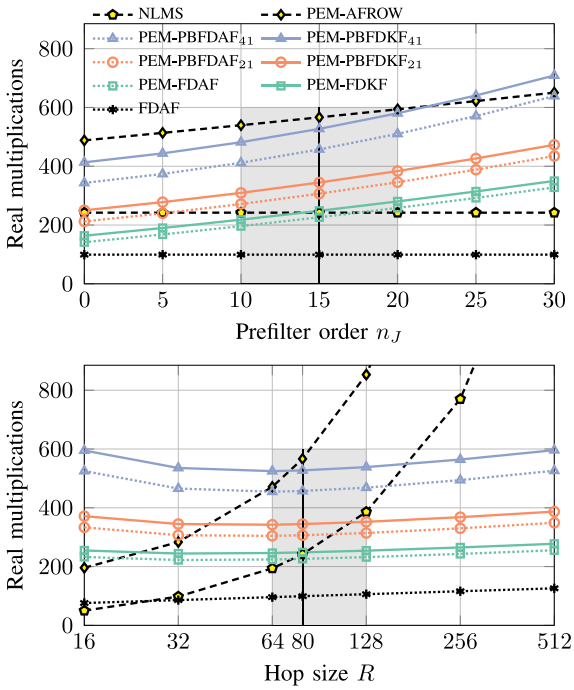

Fig. 3. Per-output-sample computational complexity of the existing and proposed algorithms as a function of the prefilter order n_J and the hop size R (top and bottom, respectively).

of real multiplications for the different algorithms is also indicated in Table II, showing that the FDAF is the cheapest, while the PEM-AFROW is the most computationally expensive (even more expensive than the PB algorithms with four partitions). For the PEM-AFROW, a window length of $M = 160$ (10 ms at 16 kHz) might be too short as common values are in the order of 40–60 ms [17], but it was kept to simplify the comparison; by using longer windows, the complexity of the algorithm would be even higher.

The complexity of the different algorithms as a function of the hop size R is shown at the bottom of Fig. 3. In this case, $n_J = 15$ and the other parameters $n_F = L = V = R = M/2$ vary with R. The grayed part of the plots corresponds to values of R yielding tolerable latency values in a HA scenario [54]. As expected, one can see how the complexity of time-domain algorithms increases far quicker than the complexity of the frequency-domain algorithms, even within the grayed part.

We conclude the section presenting a worst-case estimate of the memory requirements for the considered algorithms (i.e., requiring the allocation of each variable, including temporary ones), which is also given in Table II in terms of floating-point values to be allocated. Using $n_F = L = V = R = M/2 = 80/P$ and $n_J = 15$, the exact number of floating-point values to be allocated is also given. This shows that the requirements of frequency-domain algorithms are more than ten-fold those of time-domain algorithms and, as expected for the PB algorithms, the memory requirements decrease as P increases. Compared to the PEM-PBFDAF, the PEM-PBFDKF requires roughly 50% more memory.

VII. SIMULATIONS

In this section we provide simulation results in order to assess the performance of the proposed algorithms. The algorithms (proposed and baseline) are compared in terms of three measures, assessing the estimation error, the achievable amplification, and the sound quality.

The first measure is the misadjustment (Mis), defined as the dB level of the normalized distance between the true and estimated feedback path:

$$\mathrm{Mis}(\kappa) = 20\log_{10}\frac{\|\mathbf{f}_{\mathrm{r}}(\kappa)\|}{\|\mathbf{f}_{\mathrm{t}}(\kappa)\|} \tag{69}$$

where $\mathbf{f}_{\mathrm{r}}(\kappa) = \mathbf{f}_{\mathrm{t}}(\kappa) - \hat{\mathbf{f}}(\kappa)$. Throughout the section, the vector containing the true feedback path coefficients will be referred to as acoustic impulse response (AIR).

The second measure is based on the so-called maximum stable gain (MSG), i.e., the maximum gain achievable at a given time without compromising the system stability; if the forward path $G(q,\kappa)$ is spectrally flat, the MSG is given as:

$$\mathrm{MSG}(\kappa) = -20\log_{10}\left[\max_{l\in\mathcal{P}(\kappa)}|F_{\mathrm{r}}(\kappa,l)|\right] \tag{70}$$

where $\mathcal{P}(\kappa)$ is the set of frequencies satisfying the phase condition of the Nyquist stability criterion [1], and $F_{\mathrm{r}}(\kappa,l)$ is the lth element of $\mathbf{F}_{\mathrm{r}}(\kappa) = \mathbf{G}^{10}_{M\times R}\mathbf{F}_R\mathbf{f}_{\mathrm{r}}(\kappa)$. The MSG is then normalized according to the maximum stable gain of the system when no feedback canceller is included, i.e., $K_{\mathrm{MSG}}(\kappa)$,

$$K_{\mathrm{MSG}}(\kappa) = -20\log_{10}\left[\max_{l\in\mathcal{P}(\kappa)}|F(\kappa,l)|\right] \tag{71}$$

resulting in the added stable gain (ASG):

$$\mathrm{ASG}(\kappa) = \mathrm{MSG}(\kappa) - K_{\mathrm{MSG}}(\kappa). \tag{72}$$

Finally, the sound quality is assessed by means of an objective measure called frequency-weighted log-spectral signal distortion (SD), a distance measure proven to correlate well with subjective evaluation of feedback cancellation algorithms [16]. The SD was calculated using the clean source signal $v(t)$ and the processed signal $d(t)$.

A. Stationary Feedback Path

For the first set of simulations, a stationary feedback path was employed. Such a scenario represents only a simplified situation, since a time variation in the feedback path can be caused by a multitude of different situations in real life, e. g. bringing a hand, or a phone close to the ear in a HA application. However, this simplified situation can be used to test the best-case scenario for the algorithms.

The following simulation results compare the performance of the PEM-FDKF and the PEM-PBFDKF for two different numbers of partitions, $P = \{2, 4\}$, and $\gamma = 1$ (PEM-PBFDKF$_{21}$ and PEM-PBFDKF$_{41}$, respectively), with three baseline algorithms. The chosen baseline algorithms are the PEM-FDAF [23] and the PEM-PBFDAF, based on the work by Spriet *et al.* [22], for two different numbers of partitions, $P = \{2, 4\}$, and $\gamma = 1$ (PEM-PBFDAF$_{21}$ and PEM-PBFDAF$_{41}$, respectively). The adaptive stepsize was calculated in a different way from what has been proposed in [22] and was defined as:

$$\mu_p(\kappa, l) = \mu_{\mathrm{fix}} \left[\delta + |U_{\mathrm{j},p}(\kappa, l)|^2\right]^{-1}, \tag{73}$$

where μ_{fix} is a real, constant, and frequency-independent value, and δ is a regularization parameter. Some algorithm-specific parameters were kept fixed in the whole set of simulations as follows: the values of μ_{fix}, δ and $\mathbf{P}(0)$ are set as to guarantee a smooth convergence in the Mis curves in the first simulated case (cf. Fig. 5), i.e., $\mu_{\mathrm{fix}} = 0.01$, $\delta = 5e-5$ and $\mathbf{P}(0) = 1.6e-2\,\mathbf{I}_{M\times M}$, respectively; although, usually, when comparing adaptive filters performance one either tunes the parameters yielding similar lowest Mis or similar speed of convergence, in this case, the proposed algorithms yield better performance in both ways.

Additionally, the transition factor is set to $\alpha = 0.99999$ [18], while the DFT length and the hop size are set to $M = 160/P$ and $L = R = V = 80/P$, respectively. The algorithms are tested in a simulated environment with the following characteristics: the transfer function $G(q, \kappa)$ is set to have constant gain $K_1 = K_{\mathrm{MSG}} - 3$ dB and a delay of $d_G = 80 samples$; $F_{\mathrm{t}}(q, \kappa) = F_{\mathrm{t}}(q)$ is a measured feedback path from a commercial HA [22] and its normalized magnitude response is shown by the thick black line (AIR1) in Fig. 4.

Fig. 5 compares the algorithm performance, by means of Mis and ASG, when varying the input source signal $v(t)$: each one of 10 overlapping 20-s segments, taken from Track 05 of the Music for Archimedes database [55], of a recorded clean male speech signal resampled at $f_{\mathrm{s}} = 16$ kHz corrupted by a 30 dB SNR additive white Gaussian noise are used as source signal $v(t)$. In

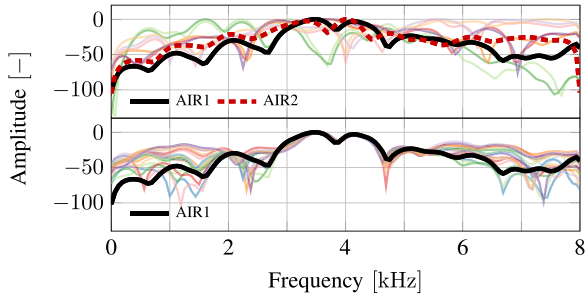

Fig. 4. Top: Magnitude responses of the 10 different AIR used in the simulations (cf. Figs. 5, 6 and 8). Bottom: Magnitude responses of 20 AIR obtained with (23a) every 1s, starting from AIR1 (cf. Fig. 7).

Fig. 5, and the following figures, the curve denotes the mean result of each algorithm while the shading corresponds to ± 1 standard deviations.

The results confirm the expected performance increase of the PEM-FDKF over PEM-FDAF [18]. Specifically, in Fig. 5, similar convergence speeds and steady-state values are are observed in the Mis curves for the PEM-FDAF and PEM-PBFDAF, with a slight improvement in both the metrics as the number of partition increases. A similar situation characterizes the ASG values. For the PEM-FDKF and PEM-PBFDKF, similar Mis and ASG values are obtained as the number of partitions is increased; a higher ASG is obtained for the PEM-PBFDAF$_{41}$, even though a greater variance is observed. The performance obtained with the FDKF-based algorithms is better, both in terms of Mis and ASG, than the performance obtained with the FDAF-based algorithms, with improvements in the range from 5 to 10 dB.

Another parameter that might change the performance of the tested algorithms is the true feedback path included in the simulation. To simulate this variability we set up a similar simulation scenario as in the previously discussed case: 10 different feedback path AIR (cf. top of Fig. 4) were employed, while the source signal $v(t)$ was fixed to a single segment of the track used in the previous simulation; the other parameters were kept fixed. The results are shown in Fig. 6 and provide a similar trend as in the previously discussed case regarding the performance improvements when using FDKF-based algorithms. The convergence curves are less smooth than in the previously discussed case and the variance is lower in the single estimates, suggesting that both the smoothness and the variance of the curves is mainly depending on the source signal content, rather than on the AIR. Nevertheless, the FDKF-based algorithms yield much less erratic results than the FDAF-based algorithms; this is due to the need of a finer tuning of the FDAF-based algorithms parameters, e. g. μ_{fix} in (73), when dealing with different feedback paths, while the FDKF-based algorithms handle these differences more robustly.

The sound quality results, in terms of mean and maximum SD, for the first three simulation scenarios are shown in Table III and labeled with the figure number showing the corresponding case. In both the scenarios the FDKF-based algorithms yield better performance and the increase of the number of partitions does not affect the results remarkably.

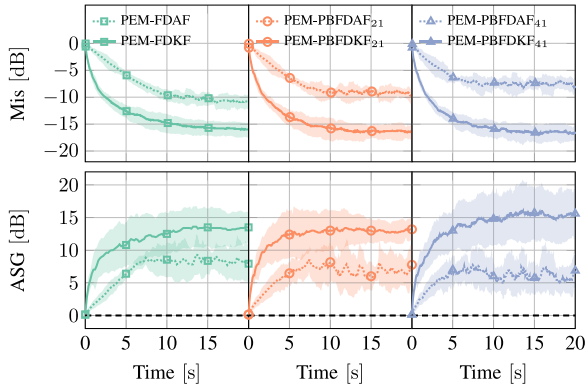

Fig. 5. Mis and ASG performance over time of PEM-FDAF, PEM-FDKF, PEM-PBFDAF and PEM-PBFDKF using the constant forward path gain K_1, AIR1, and 10 different source signals.

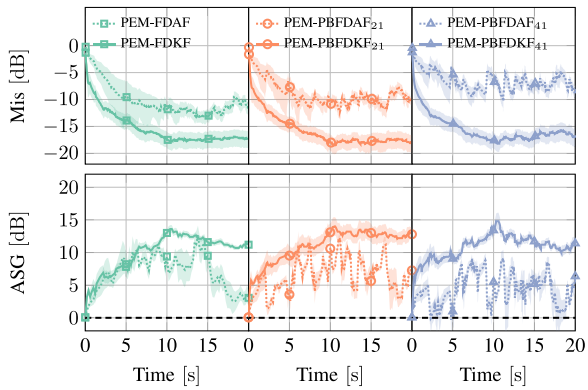

Fig. 6. Mis and ASG performance over time of PEM-FDAF, PEM-FDKF, PEM-PBFDAF and PEM-PBFDKF using the constant forward path gain K_1 and 10 different AIR.

TABLE III
MEAN AND MAXIMUM SD VALUES FOR THE DIFFERENT ALGORITHMS IN THE FIVE DIFFERENT SIMULATED SCENARIOS

	mean(SD) | max(SD)							
	Fig. 5		Fig. 6		Fig. 7		Fig. 8	
PEM-FDAF	0.93	2.77	0.85	2.79	1.05	2.97	1.26	3.67
PEM-PBFDAF$_{21}$	1.06	3.39	1.20	4.51	1.13	3.62	1.30	3.91
PEM-PBFDAF$_{41}$	1.26	3.81	1.67	5.57	1.27	3.92	1.44	4.40
PEM-FDKF	0.54	2.60	0.61	2.99	0.74	2.74	1.10	3.87
PEM-PBFDKF$_{21}$	0.48	2.47	0.54	2.80	0.66	2.26	0.88	3.69
PEM-PBFDKF$_{41}$	0.47	2.35	0.54	2.20	0.67	2.39	0.68	3.36

B. *Non-Stationary Feedback Path*

For the second set of simulations, a non-stationary feedback path was employed. Both a smooth time evolution and an abrupt change were simulated. The first type of non-stationarity is given by the model described in (23a), where $\mathbf{F}(0)$ is AIR1, and the "true" simulated transition factor is $\alpha_t = 0.99999$. The process noise vector is generated as a complex normally-distributed random process with zero mean and variance set equal to a scaled version of the frequency-dependent variance of the 10 recorded AIR shown at the top of Fig. 4; this was done to avoid high variability at the beginning and at the end of the AIR, in order to match more realistic AIR profiles. In this simulation, for the FDKF-based algorithms, both the transition factor and the frequency-dependent variance of the process noise were matching the "true" ones used in (23a). The magnitude responses of 20 AIR obtained with (23a) every after 1s of simulation are shown at the bottom of Fig. 4. The results of the simulation are shown in Fig. 7. The performance are worse than in the stationary case, but it should be noted that (23a) yields a strongly nonstationary behavior, i.e., a true feedback path changing every single frame. Despite the very fast dynamics of the "true" feedback path, the FDKF-based algorithms manage to provide a performance gap in the Mis between 2 to 5 dB. Similar conclusions hold for the ASG, with improvements of the FDKF-based algorithms over the FDAF-based algorithms being lower than in the stationary case.

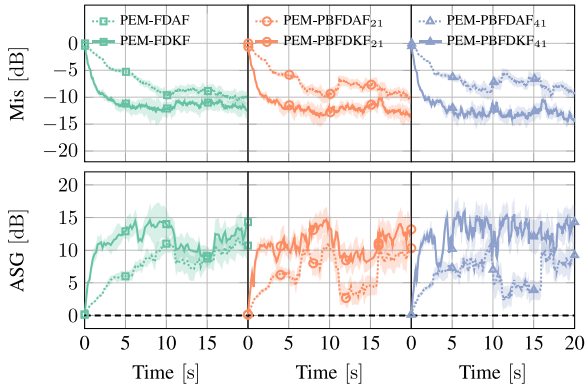

Fig. 7. Mis and ASG performance over time of PEM-FDAF, PEM-FDKF, PEM-PBFDAF and PEM-PBFDKF using the constant forward path gain K_1, a smooth AIR transition (cf. bottom of Fig. 4), simulated 10 times.

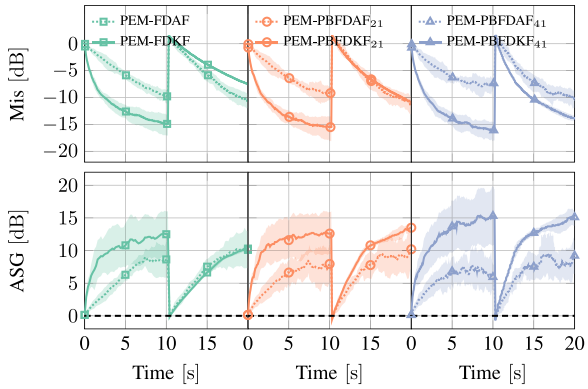

Fig. 8. Mis and ASG performance over time of PEM-FDAF, PEM-FDKF, PEM-PBFDAF and PEM-PBFDKF using the constant forward path gain K_1, an abrupt AIR transition from AIR1 to AIR2, and 10 different source signals.

The final simulation scenario considered in this section involves an abrupt feedback path change, from AIR1 to AIR2 (cf. Fig. 4), at half of the simulation period. The results are shown in Fig. 8. It can be noticed how the abrupt change causes a performance drop, which is naturally more evident for the

FDKF-based algorithms since, after 10 s, these have already reached almost −15 dB Mis and 12 to 14 dB ASG, roughly 5 dB better than the performance measures of the FDAF-based algorithms at the same time. However, unlike the FDAF-based algorithms reconverging to similar levels, the FDKF-based algorithms behave differently after the abrupt change, showing an improved convergence speed as the number of partitions is increased. While the effect on the PEM-FDKF algorithm is expected, since the abrupt change causes a model mismatch in the state equation, the better performance of the PBFDKF algorithms might be due to the better tracking behavior due to the smaller block size of the PB implementations.

The sound quality results for the last two simulation scenarios, shown in Table III, confirm the findings from the first two. Also in this case, an increased number of partitions provides lower SD values than for the non-partitioned case, possibly due to the improved tracking behavior of the algorithm in case of a smaller block size.

VIII. Conclusion

In this paper, we have presented a new AFC algorithm employing a PEM-based signal prewhitening in combination with a FDKF; additionally, we have derived the ICs to successfully apply such algorithm in a closed-loop system scenario.

We have also proposed an extension of the algorithm by means of a PB implementation, which makes the algorithm more appealing for use in systems where large algorithmic delays are not tolerated, such as in HA applications.

A study of the computational complexity of the new algorithms has shown that, compared to existing frequency-domain algorithms, both the complexity increase and the memory requirements increase are limited.

Simulation results for different scenarios, including both stationary and non-stationary feedback paths, have shown the improved performance of the proposed algorithms in terms of different objective measures.

Acknowledgment

The authors would like to thank Prof. P. Sommen and Dr. F. Kuech for their helpful discussions on PBFDAF and Bang & Olufsen for providing the speech material. The scientific responsibility is assumed by its authors.

References

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

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

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

[4] M. Guo, S. Jensen, and J. Jensen, "Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancement," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2549–2563, Nov. 2012.

[5] C. R. C. Nakagawa, S. Nordholm, and W. Y. Yan, "Feedback cancellation with probe shaping compensation," *IEEE Signal Process. Lett.*, vol. 21, no. 3, pp. 365–369, Mar. 2014.

[6] G. Schmidt and T. Haulick, "Signal processing for in-car communication systems," *Signal Process.*, vol. 86, no. 6, pp. 1307–1326, 2006.

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

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

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

[10] C. R. C. Nakagawa, S. Nordholm, and W. Y. Yan, "Analysis of two microphone method for feedback cancellation," *IEEE Signal Process. Lett.*, vol. 22, no. 1, pp. 35–39, Jan. 2015.

[11] 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.

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

[13] 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.

[14] F. Strasser and H. Puder, "Adaptive feedback cancellation for realistic hearing aid applications," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 12, pp. 2322–2333, Dec. 2015.

[15] H. Schepker, L. T. T. Tran, S. Nordholm, and S. Doclo, "Improving adaptive feedback cancellation in hearing aids using an affine combination of filters," in *Proc. 2016 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar 2016, pp. 231–235.

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

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

[18] G. Bernardi, T. van Waterschoot, J. Wouters, M. Hillbratt, and M. Moonen, "A PEM-based frequency-domain Kalman filter for adaptive feedback cancellation," in *Proc. 23rd Eur. Signal Process. Conf.*, Nice, France, Aug. 2015, pp. 270–274.

[19] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Wiener variable step size and gradient spectral variance smoothing for double-talk-robust acoustic echo cancellation and acoustic feedback cancellation," *Signal Process.*, vol. 104, pp. 1–14, Nov. 2014.

[20] K. Ngo, T. van Waterschoot, M. G. Christensen, M. Moonen, and S. H. Jensen, "Improved prediction error filters for adaptive feedback cancellation in hearing aids," *Signal Process.*, vol. 93, no. 11, pp. 3062–3075, 2013.

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

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

[23] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "A frequency-domain adaptive filter (FDAF) prediction error method (PEM) framework for double-talk-robust acoustic echo cancellation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 12, pp. 2074–2086, Dec. 2014.

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

[25] M. Rotaru, F. Albu, and H. Coanda, "A variable step size modified decorrelated NLMS algorithm for adaptive feedback cancellation in hearing aids," in *Proc. 10th Int. Symp. Electron. Telecommun.*, Nov. 2012, pp. 263–266.

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

[27] J. Benesty and D. Morgan, "Frequency-domain adaptive filtering revisited, generalization to the multi-channel case, and application to acoustic echo cancellation," in *Proc. 2000 IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2000, vol. 2, pp. 789–792.
[28] G. Enzner, "Bayesian inference model for applications of time-varying acoustic system identification," in *Proc. 18th Eur. Signal Process. Conf.*, Aalborg, Denmark, Aug. 2010, pp. 2126–2130.
[29] A. Jain, S. Goel, K. Nathwani, and R. M. Hegde, "Robust acoustic echo cancellation using Kalman filter in double talk scenario," *Speech Commun.*, vol. 70, pp. 65–75, 2015.
[30] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. Jensen, "Transform domain prediction error method for improved acoustic echo and feedback cancellation," in *Proc. 20th Eur. Signal Process. Conf.*, Bucharest, Romania, Aug. 2012, pp. 2422–2426.
[31] G. Bernardi, T. van Waterschoot, J. Wouters, and M. Moonen, "An all-frequency-domain adaptive filter with PEM-based decorrelation for acoustic feedback control," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, New Paltz, NY, USA, Oct. 2015, pp. 1–5.
[32] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.
[33] E. Moulines, O. Ait Amrane, and Y. Grenier, "The generalized multidelay adaptive filter: Structure and convergence analysis," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 14–28, Jan. 1995.
[34] B. H. Nitsch, "A frequency-selective stepfactor control for an adaptive filter algorithm working in the frequency domain," *Signal Process.*, vol. 80, no. 9, pp. 1733–1745, Sep. 2000.
[35] G. W. Enzner, "A model based optimum filtering approach to acoustic echo control: Theory and practice," Ph.D. dissertation, Inst. für Nachrichtengeräte und Datenverarbeitung (IND), RWTH, Aachen, Germany, 2006.
[36] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. New York, NY, USA: Springer, 2001.
[37] J.-S. Soo and K. Pang, "Multidelay block frequency domain adaptive filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 2, pp. 373–376, Feb. 1990.
[38] M. L. Valero, E. Mabande, and E. A. Habets, "A state-space partitioned-block adaptivefilter for echo cancellation using inter-band correlations in the Kalman gain computation," in *Proc. 2015 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2015, pp. 599–603.
[39] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[40] 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.
[41] 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.
[42] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.
[43] F. Kuech, E. Mabande, and G. Enzner, "State-space architecture of the partitioned-block-based acoustic echo controller," in *Proc. 2014 IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2014, pp. 1295–1299.
[44] S. Haykin, A. Sayed, J. Zeidler, P. Yee, and P. Wei, "Adaptive tracking of linear time-variant systems by extended RLS algorithms," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1118–1128, May 1997.
[45] H. W. Sorenson, "Least-squares estimation: From Gauss to Kalman," *IEEE Spectrum*, vol. 7, no. 7, pp. 63–68, Jul. 1970.
[46] J. B. Rawlings and D. Q. Mayne, *Model Predictive Control: Theory and Design*. Madison, WI: Nob Hill Publishing, 2009.
[47] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Process. Mag.*, vol. 11, no. 3, pp. 18–60, Jul. 1994.
[48] P. Sommen, P. van Gerwen, H. Kotmans, and A. Janssen, "Convergence analysis of a frequency-domain adaptive filter with exponential power averaging and generalized window function," *IEEE Trans. Circuits Syst.*, vol. 34, no. 7, pp. 788–798, Jul. 1987.
[49] F. Faccenda, S. Squartini, E. Principi, L. Gabrielli, and F. Piazza, "A real-time dual-channel speech reinforcement system for intra-cabin communication," *J. Audio Eng. Soc*, vol. 61, no. 11, pp. 889–910, Nov. 2013.
[50] G. Rombouts, T. van Waterschoot, and M. Moonen, "Robust and efficient implementation of the PEMAFROW algorithm for acousic feedback cancellation," *J. Audio Eng. Soc.*, vol. 55, no. 11, pp. 955–966, 2007.
[51] P. Sommen, "On the convergence properties of a partitioned block frequency domain adaptive filter (PBFDAF)," in *Proc. 5th Eur. Signal Process. Conf.*, Barcelona, Spain, Sep. 1990, pp. 201–204.
[52] H. Buchner, J. Benesty, and W. Kellermann, "Generalized multichannel frequency-domain adaptive filtering: Efficient realization and application to hands-free speech communication," *Signal Process.*, vol. 85, no. 3, pp. 549–570, 2005.
[53] H. Buchner, J. Benesty, T. Gansler, and W. Kellermann, "Robust extended multidelay filter and double-talk detector for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 5, pp. 1633–1644, Sep. 2006.
[54] M. A. Stone and B. C. Moore, "Tolerable hearing aid delays. I. Estimation of limits imposed by the auditory path alone using simulated hearing losses," *Ear Hearing*, vol. 20, no. 3, pp. 182–191, 1999.
[55] Bang and Olufsen, "Music for archimedes," CD B&O 101, 1992.

Giuliano Bernardi (S'12) was born in Asolo, Italy, in 1987. He received the M.Sc. degree in engineering acoustics from Denmark Technical University, Lyngby, Denmark, in 2011, and the M.Eng. degree in bioengineering from the University of Padua, Padua, Italy, in 2012. He is currently working toward the Ph.D. degree at the STADIUS research division within the Department of Electrical Engineering at KU Leuven, Leuven, Belgium.

His research focuses on acoustic feedback control, specifically for hearing-aid applications.

Toon van Waterschoot (S'04–M'12) received the M.Sc. and Ph.D. degrees (2009) in electrical engineering, both from KU Leuven, Leuven, Belgium, in 2001 and 2009, respectively

He is currently a tenure-track Assistant Professor at KU Leuven, Leuven, Belgium. He has previously held teaching and research positions with the Antwerp Maritime Academy, the Institute for the Promotion of Innovation through Science and Technology in Flanders, and the Research Foundation - Flanders in Belgium, with Delft University of Technology in The Netherlands, and with the University of Lugano in Switzerland. His research interests include signal processing, machine learning, and numerical optimization, applied to acoustic signal enhancement, acoustic modeling, audio analysis, and audio reproduction.

Dr. Waterschoot has been serving as an Associate Editor of the *Journal of the Audio Engineering Society* and the *EURASIP Journal on Audio, Music, and Speech Processing*, and as a Guest Editor of Elsevier *Signal Processing*. He is a Member of the Board of Directors of the European Association for Signal Processing and a Member of the IEEE Audio and Acoustic Signal Processing Technical Committee. He was the General Chair of the 60th AES International Conference in Leuven, Belgium (2016), and has been serving on the Organizing Committee of the European Conference on Computational Optimization and the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. He is a member of EURASIP, ASA, and AES.

Jan Wouters was born in 1960. He received the Master's and Ph.D. degrees in physics from the University of Leuven, Katholieke Universiteit (KU) Leuven, Leuven, Belgium, in 1982 and 1989, respectively, with intermission for officer military service.

From 1989 to 1992, he was a Postdoctoral Research Fellow with the Belgian National Fund for Scientific Research, Institute of Nuclear Physics, Catholic University of Louvain, Louvain-la-Neuve and at NASA Goddard Space Flight Center, USA. Since 1993, he has been a Professor at the Neurosciences Department, KU Leuven, and has been a Full Professor since 2005. His research interests include audiology and the auditory system, signal processing for cochlear implants, and hearing aids.

Dr. Wouters is on the Editorial Board of the *International Journal of Audiology*, the *Journal of Communication Disorders*, and the *Journal B-ENT*. He is the President of the European Federation of Audiological Societies and the Belgian Audiological Society. He is a member of the International Collegium for ORL (CORLAS), and a Board Member of the International Collegium for Rehabilitative Audiology.

Marc Moonen (M'94–SM'06–F'07) is a Full Professor at the Electrical Engineering Department, KU Leuven, Leuven, Belgium, 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.

Prof. Moonen has served as an Editor-in-Chief for the EURASIP *Journal on Applied Signal Processing* (2003–2005), Area Editor for Feature Articles in *IEEE Signal Processing Magazine* (2012–2014), and has been a member of the editorial board of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, *IEEE Signal Processing Magazine*, Integration-the *VLSI Journal*, the EURASIP *Journal on Wireless Communications and Networking*, the EURASIP *Journal on Applied Signal Processing,* and EURASIP *Signal Processing*. He received the 1994 KU 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 journal best paper awards from the IEEE TRANSACTIONS ON SIGNAL PROCESSING (with Geert Leus and Daniele Giacobello) and from Elsevier *Signal Processing* (with Simon Doclo). He was chairman of the IEEE Benelux Signal Processing Chapter (1998–2002), a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications, and President of EURASIP (European Association for Signal Processing, 2007–2008 and 2011–2012).