MULTICHANNEL IDENTIFICATION OF ROOM ACOUSTIC SYSTEMS WITH ADAPTIVE FILTERS BASED ON ORTHONORMAL BASIS FUNCTIONS

Giacomo Vairetti[1] *Enzo De Sena*[1] *Michael Catrysse*[2]
Søren Holdt Jensen[3] *Marc Moonen*[1] *Toon van Waterschoot*[1,4]

[1]KU Leuven, Dept. of Electrical Engineering (ESAT-STADIUS), Kasteelpark Arenberg 10, 3001 Leuven, Belgium
[2]Televic N.V., Leo Bekaertlaan 1, 8870 Izegem, Belgium
[3]Aalborg University, Dept. of Electronic Systems, Fredrik Bajers Vej 7B, 9220 Aalborg, Denmark
[4]KU Leuven, Dept. of Electrical Engineering (ESAT-ETC), AdvISe Lab, Kleinhoefstraat 4, 2440 Geel, Belgium

ABSTRACT

Many acoustic signal enhancement applications require adaptive filters with a long impulse response, but with a small number of filter parameters. Fixed-poles infinite impulse response (IIR) adaptive filters based on orthonormal basis functions (OBFs) present advantages over finite impulse response filters and other IIR filters, assuring stability and fast global convergence in the adaptation of the filter parameters. A scalable algorithm is introduced for the estimation of the poles of an adaptive OBF filter from multichannel input-output data. The set of poles, common to all the acoustic channels considered, is estimated in parallel to the adaptation of the linear filter parameters. It will be shown that the result of the identification with common poles is quite robust to variations in the room transfer function, suggesting the possibility that poles may be kept fixed after estimation.

***Index Terms*—** Orthonormal basis functions, parametric modeling, room acoustics, adaptive filtering, identification.

1. INTRODUCTION

Many acoustic signal enhancement applications require a compact yet accurate approximation of the room impulse response (RIR) at one or multiple locations of the source and the receiver inside a room. Parametric modeling and identification of room acoustic systems aim at representing a room transfer function (RTF) as a rational function in the z-domain that can be implemented using a digital filter, under the assumption of the room being a stable, causal and linear system. However, the RTF can be time-varying, due for instance to changes in the source or microphone position, thus requiring the filter parameters to adapt for tracking the variation. Filters having a finite impulse response (FIR) are widely used because of their simplicity, but they often require a large number of parameters. Filters having an infinite impulse response (IIR) can provide a reduction in the number of filter parameters, but they suffer from problems of instability and convergence to local minima, also introducing extra complexity in the estimation and adaptation of the parameters [1].

Fixed-poles IIR filters based on orthonormal basis functions (OBFs) [2] (henceforth called OBF filters) represent an appealing alternative to conventional IIR filters, especially in the modeling of room acoustic systems [3]. The filter structure is an orthonormalized parallel realization of second-order all-pole filters, each of these corresponding to a resonator. A RTF can then be modeled as a linear combination of resonances, whose frequencies and bandwidths are determined by the position of the poles. As for other fixed-poles IIR filters [4], OBF filters can increase the modeling accuracy dramatically compared to FIR filters by moving the poles away from the origin, thus reducing the distance between the true poles of the system and the poles of the filter [5]. Moreover, stability of the filter can be easily guaranteed by constraining the poles to be inside the unit disc, while preserving the global convergence properties of FIR filters [4]. The main advantage of OBF filters over other fixed-poles IIR filters is orthogonality, which provides numerical well-conditioning and fast convergence of the filter adaptation [6, 7]. Since the poles appear in the denominator of the filter transfer function (TF), nonlinear techniques are usually necessary to estimate the poles [4, 6, 8]. A scalable matching pursuit algorithm named OBF-MP was proposed in [9], where the nonlinear problem was avoided by defining a grid of candidate poles and iteratively selecting those providing the best approximation of a measured RIR. An extension of the algorithm, named OBF-GMP [10], was proposed for estimating a set of poles common to multiple RIRs measured at different source-receiver positions inside a room, thus obtaining a parametrization of a set of RIRs which is more compact and less sensitive to variations of the RTF. The poles are usually estimated off-line, starting from a set of RIR measurements [6]. In [11], a recursive separable nonlinear least-squares (LS) method was proposed for the on-line estimation of both poles and linear filter parameters from input-output data and applied to the identification of acoustic echo systems [12], but limited to the case of an OBF filter with a single repeated pole. However, while the adaptation of the linear filter parameters is straightforward [4], the gradient-based adaptation of a set of non-repeated poles is computationally expensive [13].

A block-based version of the OBF-MP algorithm, called BB-OBF-MP, was proposed in [14] for the iterative estimation of the poles from short input-output data segments, where the linear filter parameters are estimated at each block using linear regression. In this paper, the block-based algorithm is applied to an adaptive OBF filter using the least mean squares (LMS) algorithm. A common set

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of i) the FP7-PEOPLE Marie Curie Initial Training Network 'Dereverberation and Reverberation of Audio, Music, and Speech (DREAMS)', funded by the European Commission under Grant Agreement no. 316969, ii) KU Leuven Research Council CoE PFV/10/002 (OPTEC), iii) the Interuniversity Attractive Poles Programme initiated by the Belgian Science Policy Office: IUAP P7/19 'Dynamical Systems Control and Optimization' (DYSCO) 2012-2017, iv) KU Leuven Impulsfonds IMP/14/037, and v) supported by a Postdoctoral Fellowship (F+/14/045) of the KU Leuven Research Fund. The scientific responsibility is assumed by its authors.

of poles estimated from multichannel input-output data is then kept fixed and validated on input-output data corresponding to source and receiver positions in the room not used in the estimation. It will be shown that the common poles parametrization is quite robust to variations in the RTF, suggesting the possibility of keeping the poles fixed after estimation, thus removing the need for any adaptation. This paper is structured as follows. In Section 2, adaptive OBF filters in the context of room acoustics modeling are briefly introduced. In Section 3, the algorithm for the estimation of the poles of an adaptive OBF filter is described. In Section 4, simulation results are presented, and Section 5 concludes the paper.

2. ROOM ACOUSTICS MODELING USING OBF FILTERS

OBF filters are particularly appropriate for modeling room acoustic systems [3]. The filter structure is an orthonormalized parallel realization of second-order all-pole filters, each one defined by a pair of complex-conjugate poles. Each second-order all-pole filter, having a TF as in (1), acts as a resonator, so that a RTF can be modeled as a linear combination of resonances. The resonance frequency ω_i and bandwidth ζ_i are determined respectively by the angle $\vartheta_i = \omega_i/f_s$ and the radius $\rho_i = e^{-\zeta_i/f_s}$ of the pair of complex-conjugate poles $\boldsymbol{p}_i = [p_i, p_i^*] = \rho_i e^{\pm j\vartheta_i}$ (with f_s being the sampling rate and * indicating complex conjugation). Notice that, when the resonances have a large bandwidth, the actual central frequency of the resonance related to p_i deviates slightly from its theoretical value ω_i, due to the influence of p_i^* and of other poles as well [15, 16].

For each all-pole filter, two real-valued basis functions are generated. A second-order all-pass filter, having a TF as in (2), is used to orthogonalize the basis functions defined by $\boldsymbol{p}_{i+1}$ with respect to those generated by $\boldsymbol{p}_i$ (the poles p_i and p_i^* are canceled by the zeros in $1/p_i$ and $1/p_i^*$). A pair of orthonormalization filters $N_i^{\pm}(z)$ is then used to enforce orthonormality between the two basis functions of each pole pair. The resulting OBF filter structure is shown in Figure 1 for m pairs of complex-conjugate poles $\boldsymbol{p} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m]$. Orthonormalization filters with a TF as in (3) result in the so-called Kautz filter, but a different choice can be made, as explained in [17].

$$P_i(z) = \frac{1}{(1-p_i z^{-1})(1-p_i^* z^{-1})} \; , \tag{1}$$

$$A_i(z) = \frac{(z^{-1}-p_i)(z^{-1}-p_i^*)}{(1-p_i z^{-1})(1-p_i^* z^{-1})} \; , \tag{2}$$

$$N_i^{\pm}(z) = |1 \pm p_i| \sqrt{\frac{1-|p_i|^2}{2}} (z^{-1} \mp 1) \; . \tag{3}$$

As can be seen in Figure 1, an OBF filter has a transversal structure, linear in the filter parameters $\boldsymbol{\theta}_i^{\pm}(n) = [\theta_i^+(n), \theta_i^-(n)]$ (with $i = 1, \ldots, m$ and $n = t/f_s$ the discrete time variable). The intermediate signals $\boldsymbol{\kappa}_i^{\pm}(n) = [\kappa_i^+(n), \kappa_i^-(n)]$ generated by a pair of OBFs are filtered versions of the input signal $u(n)$, where the TF of a pair of OBFs is given by $\Psi_i^{\pm}(z) = N_i^{\pm}(z) P_i(z) \prod_{j=1}^{i-1} A_j(z)$, and hence $\boldsymbol{\kappa}_i^{\pm}(n) = \Psi_i^{\pm}(z) u(n)$. Since the OBFs form a complete set in the Hardy space on the unit disc under mild assumptions, any stable rational TF can be realized with arbitrary accuracy by a linear combination of a finite number of OBFs [17], with the output signal then given by

$$y(n, \boldsymbol{p}, \boldsymbol{\theta}) = \sum_{i=1}^{m} \kappa_i^+(n)\theta_i^+(n) + \sum_{i=1}^{m} \kappa_i^-(n)\theta_i^-(n) \; , \tag{4}$$

or in vector form as $y(n, \boldsymbol{p}, \boldsymbol{\theta}) = \boldsymbol{\kappa}^T(\boldsymbol{p}, n)\boldsymbol{\theta}(n)$, with $\boldsymbol{\kappa}(\boldsymbol{p}, n) = [\boldsymbol{\kappa}_1^{\pm}(n), \ldots, \boldsymbol{\kappa}_m^{\pm}(n)]^T$ and $\boldsymbol{\theta}(n) = [\boldsymbol{\theta}_1^{\pm}(n), \ldots, \boldsymbol{\theta}_m^{\pm}(n)]^T$, both with dimensions $2m \times 1$. The number of OBFs necessary to achieve a certain level of accuracy depends on the distance between the true poles of the systems and the fixed poles in the denominator of the filter TF [5]. It is then clear that a higher accuracy can be achieved by fixed-poles IIR filters, such as OBF filters, compared to FIR filters, where all poles are fixed at the origin. It follows that, although OBF filters introduce additional computational complexity per linear filter parameter compared to FIR filters, this extra complexity is compensated by a reduction in the number of linear filter parameters by an appropriate selection strategy for the position of the poles (such as the OBF-MP algorithm) [3].

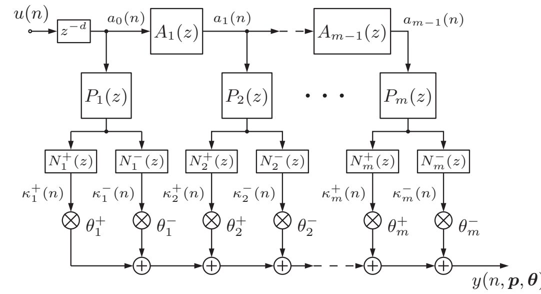

Fig. 1. The OBF filter for m pairs of complex-conjugate poles.

2.1. Adaptation of linear filter parameters

Fixed-pole adaptive filters (FPAFs) [4] have been proposed to overcome practical problems related to adaptive IIR filters, such as finite-precision effects, convergence and stability [1]. Linearity in the filter parameters induces globally convergent adaptation under the same conditions and with the same implementation complexity as for the adaptation scheme of FIR filters with the same number of adaptive parameters [4]. It follows that standard adaptive algorithms can be readily applied. Adaptive OBF filters share all the properties of FPAFs, with the additional property of orthogonality, which ensures better-behaved and faster convergence of the adaptation algorithm [5]. The adaptation rule for the recursive estimation of the linear filter parameter vector $\boldsymbol{\theta}(n)$ is given by (with $\boldsymbol{L}(n)$ a gain vector)

$$\hat{\boldsymbol{\theta}}(n+1) = \hat{\boldsymbol{\theta}}(n) + \boldsymbol{L}(n)\left(y(n) - \boldsymbol{\kappa}^T(\boldsymbol{p}, n)\hat{\boldsymbol{\theta}}(n)\right) \; . \tag{5}$$

The simplest choice for the gain vector is $\boldsymbol{L}(n) = \mu\boldsymbol{\kappa}(\boldsymbol{p}, n)$, with μ the step size, in which case (5) corresponds to the LMS algorithm. Different choices of the gain vector lead to different adaptation algorithms, such as the recursive least squares (RLS) algorithm or the Kalman filter [4, 5]. Here, only the LMS algorithm is considered for its simplicity. It should be noted that, as opposed to FIR filters where the regression vector is made of the last M samples of the input signal (with M the model order), the regression vector for the adaptive OBF filter is the vector of intermediate signal samples $\boldsymbol{\kappa}(\boldsymbol{p}, n)$. A known fact for the LMS algorithm is that the convergence is determined by the choice of the step size μ and by the eigenvalue spread of the correlation matrix of the intermediate signals $\boldsymbol{R} = \mathrm{E}\{\boldsymbol{\kappa}(\boldsymbol{p}, n)\boldsymbol{\kappa}^T(\boldsymbol{p}, n)\}$, with E denoting the expected value. The convergence speed is determined by the minimum eigenvalue of $\boldsymbol{R}$, $\lambda_{\min}$, accordingly to the exponential factor $(1 - \mu\lambda_{\min})^n$, with a larger value for $\lambda_{\min}$ yielding faster convergence [18]. A step size $\mu < 1/\lambda_{\max}$, with $\lambda_{\max}$ the maximum eigenvalue of $\boldsymbol{R}$, guarantees that the exponential factor decays to zero, so that the convergence rate in the mean for the LMS algorithm can be no faster than [5]

$$\left(1-\frac{\lambda_{\min}}{\lambda_{\max}}\right)^n=\left(1-\frac{1}{C(\boldsymbol{R})}\right)^n, \tag{6}$$

with $C(\boldsymbol{R})$ the condition number of $\boldsymbol{R}$. Therefore, a small (close to 1) condition number implies a faster convergence.

For OBF filters with a white input signal, with constant spectral density $\Phi_u(\omega)=c$, the convergence rate is optimal within the class of FPAFs, as the correlation matrix is optimally-conditioned ($\boldsymbol{R}=c\boldsymbol{I}$, with $\boldsymbol{I}$ the identity matrix). For nonwhite input signals, the optimal convergence rate is lost. However, OBF filters are particularly robust in terms of numerical well-conditioning also in this case [19], so that the condition number remains small, even when a large number of basis functions is used.

3. IDENTIFICATION ALGORITHM

The main issue with adaptive OBF filters is the adaptive estimation from input-output data of the poles, which appear in the denominator of the TFs $\Psi_i^{\pm}(z)$, thus requiring nonlinear estimation techniques (see e.g. the method in [11]). Gradient-based adaptive algorithms, such as the LMS, demand the computation of a sensitivity function (i.e. the gradient of the filter TF w.r.t. a parameter vector), which for an OBF filter with non-repeated poles is particularly complicated [13]. Since the adaptive estimation of the poles does not seem to be very practical, a multichannel identification algorithm is proposed in this paper, where the poles of an adaptive OBF filter are estimated from single-input multiple-output (SIMO) data sets. A set of common poles is estimated with a multichannel version of the BB-OBF-MP algorithm [14], modified to be applied to an adaptive OBF filter using the LMS algorithm as follows. In each block (i.e. every N_f samples), one pole pair is selected from a user-defined grid of candidate poles as the one that produces the pair of OBFs that is mostly correlated with the last N_f samples of the estimation error signal produced in each acoustic channel considered, similarly to the selection strategy used in the OBF-GMP algorithm [10]. The idea suggested here is that the set of poles estimated from SIMO data is common to all the acoustic channels considered in the estimation and is also robust to variations of the RTF, so that these poles can be kept fixed after estimation, without the need for adapting the position of the poles.

The proposed algorithm aims to build a SIMO adaptive OBF filter including one common pole pair at a time, so that the mean-square-error (MSE) of the acoustic channel r (with $r=1,\ldots,R$) is minimized,

$$\underset{\boldsymbol{p}_f^{\mathcal{A}},\hat{\boldsymbol{\Theta}}(n)}{\text{minimize}}\ \boldsymbol{e}^2(n)=(\boldsymbol{y}(n)-\hat{\boldsymbol{y}}(n))^2=\Big(\boldsymbol{y}(n)-\boldsymbol{\kappa}(\boldsymbol{p}_f^{\mathcal{A}},n)\hat{\boldsymbol{\Theta}}(n)\Big)^2, \tag{7}$$

where $\boldsymbol{e}(n)=[e_1(n),\ldots,e_R(n)]$ is the vector of the estimation errors for the R channels at time n, $\boldsymbol{y}(n)=[y_1(n),\ldots,y_R(n)]$ is the vector of the output signals and $\hat{\boldsymbol{y}}(n)=[\hat{y}_1(n),\ldots,\hat{y}_R(n)]$ the vector of the estimated outputs of the adaptive OBF filter. This vector is obtained by the linear combination of the intermediate signals $\boldsymbol{\kappa}(\boldsymbol{p}_f^{\mathcal{A}},n)=[\boldsymbol{\kappa}_1^{\pm}(n),\ldots,\boldsymbol{\kappa}_f^{\pm}(n)]$ ($\boldsymbol{\kappa}_i^{\pm}(n)=[\kappa_i^{+}(n),\kappa_i^{-}(n)]$), which are the input signal $u(n)$ filtered by the orthonormal basis TFs generated by the active pole set $\boldsymbol{p}_f^{\mathcal{A}}$ (with f the block index, corresponding to the number of common poles in the active set). The intermediate signals are weighted by the linear filter parameter vectors $\hat{\boldsymbol{\theta}}^r(n)=[\hat{\boldsymbol{\theta}}_1^{r\pm}(n),\ldots,\hat{\boldsymbol{\theta}}_f^{r\pm}(n)]$ (with $\hat{\boldsymbol{\theta}}_i^{r\pm}(n)=[\theta_i^{r+}(n),\theta_i^{r-}(n)]$) related to the R channels, which are

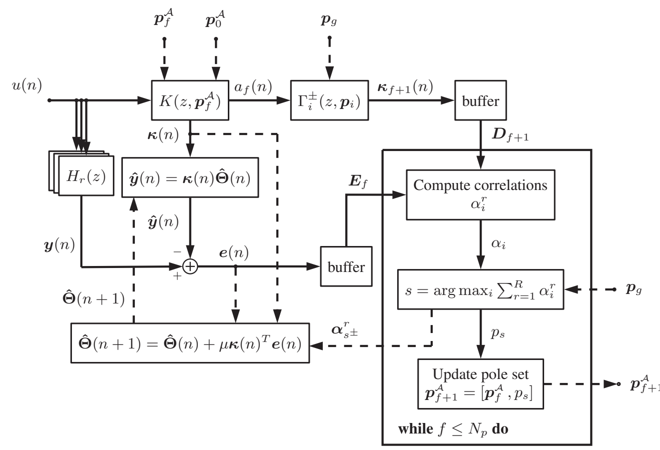

Fig. 2. The identification algorithm block diagram. Inbound dashed lines represent initial conditions and inputs, while outbound dashed lines represent outputs.

stacked in the matrix $\hat{\boldsymbol{\Theta}}(n)=[\hat{\boldsymbol{\theta}}^1(n),\ldots,\hat{\boldsymbol{\theta}}^R(n)]$ of dimensions $2f\times R$. The linear filter parameter matrix $\hat{\boldsymbol{\Theta}}(n)$ is updated using the LMS adaptation rule

$$\hat{\boldsymbol{\Theta}}(n+1)=\hat{\boldsymbol{\Theta}}(n)+\mu\,\boldsymbol{\kappa}(\boldsymbol{p}_f^{\mathcal{A}},n)\Big(\boldsymbol{y}(n)-\boldsymbol{\kappa}^T(\boldsymbol{p}_f^{\mathcal{A}},n)\hat{\boldsymbol{\Theta}}(n)\Big). \tag{8}$$

Initially ($f=0$), the active pole set $\boldsymbol{p}_f^{\mathcal{A}}$ is empty, so that no estimated output vector $\hat{\boldsymbol{y}}(n)$ is produced ($\boldsymbol{e}(n)=\boldsymbol{y}(n)$).

The poles of the adaptive OBF filter are estimated using a block-based matching pursuit algorithm, which is depicted in Figure 2 with a slightly simplified notation. First, a grid $\boldsymbol{p}_g$ of G candidate poles is defined on the unit disc based on some prior knowledge of the room acoustic system or some particular desired frequency resolution. For each pole pair $\boldsymbol{p}_i\in\boldsymbol{p}_g$ (with $i=1,\ldots,G$), the pair of intermediate signals $\boldsymbol{\kappa}_{f+1,i}^{\pm}(n)=[\kappa_{f+1,i}^{+}(n),\kappa_{f+1,i}^{-}(n)]$ is obtained as the $(f+1)$-th intermediate signals of an OBF filter built from the pole set $[\boldsymbol{p}_f^{\mathcal{A}},\boldsymbol{p}_i]$, i.e. by filtering the input $u(n)$ with the TFs $\Psi_i^{\pm}(z)=N_i^{\pm}(z)P_i(z)\prod_{j=1}^f A_j(z)$, where the product corresponds to the series of all-pass filters defined by the poles in $\boldsymbol{p}_f^{\mathcal{A}}$ (cfr. Figure 1). Equivalently, $\boldsymbol{\kappa}_{f+1,i}^{\pm}(n)$ can be computed by filtering the output of the all-pass series $a_f(n)=\prod_{j=1}^f A_j(z)u(n)$ with pairs of filters with TFs $\Gamma_i^{\pm}(z,\boldsymbol{p}_i)=N_i^{\pm}(z)P_i(z)$. The vector $\boldsymbol{\kappa}_{f+1}(n)=[\boldsymbol{\kappa}_{f+1,1}^{\pm}(n),\ldots,\boldsymbol{\kappa}_{f+1,G}^{\pm}(n)]$ of the intermediate signals computed for all the pole pairs in $\boldsymbol{p}_g$ are then collected for N_f samples and stacked to build the dictionary $\boldsymbol{D}_{f+1}$, which is a matrix whose columns $\boldsymbol{d}_i^{\pm}$ are the last N_f samples of the G pairs of intermediate signals $\boldsymbol{\kappa}_{f+1,i}^{\pm}(n)$. At each block, a pole pair is selected based on the correlation of the pairs of intermediate signal sequences $\boldsymbol{d}_i^{\pm}$ with the last N_f samples of the estimation error vector $\boldsymbol{e}(n)$, stacked to form a matrix $\boldsymbol{E}_f$, whose column $\boldsymbol{\epsilon}_f^r$ is a vector containing the last N_f samples of the estimation error signal $e_r(n)$. The correlation of each pair of intermediate signal sequences $\boldsymbol{d}_i^{\pm}$ with the $\boldsymbol{\epsilon}_f^r$ for the r-th channel is computed as

$$\alpha_{f+1,i}^r=\sqrt{{\alpha_{i^+}^r}^2+{\alpha_{i^-}^r}^2}=\sqrt{({\boldsymbol{d}_i^+}^T\boldsymbol{\epsilon}_f^r)^2+({\boldsymbol{d}_i^-}^T\boldsymbol{\epsilon}_f^r)^2}. \tag{9}$$

The pair of intermediate signal sequences in the dictionary having maximum correlation with the estimation error matrix $\boldsymbol{E}_f$ is selected according to $s=\arg\max_i\sum_{r=1}^R\alpha_{f+1,i}^r$ and the corresponding pole pair $\boldsymbol{p}_s \in \boldsymbol{p}_g$ is added to the active pole set $\boldsymbol{p}_{f+1}^{\mathcal{A}}$ and included in the adaptive OBF filter structure. The linear filter parameters $\hat{\boldsymbol{\theta}}_{f+1}^{r\pm}(n) = [\theta_{f+1}^{r+}(n), \theta_{f+1}^{r-}(n)]$ are set equal to the correlation coefficients $\boldsymbol{\alpha}_{s\pm}^{r} = [\alpha_{s+}^{r}, \alpha_{s-}^{r}]$ (with $r = 1, \dots, R$), normalized w.r.t. the norm of $\boldsymbol{d}_s^{\pm}$. In this way, the linear filter parameters are already close to their optimal value (assuming that the RTF of the acoustic channels is time-invariant during the estimation), so that a small value for μ can be used in order to achieve better accuracy with the LMS algorithm. Finally, the algorithm moves to the next block ($f = f + 1$) where another pole pair is estimated from the last N_f samples of the estimation error signals and of the candidate $(f+1)$-th intermediate signals as described above, until a desired number of poles N_p has been estimated or the MSE falls below a certain threshold.

4. SIMULATION RESULTS

The simulation results presented here aim to verify that OBF filters with common poles estimated with the algorithm described above increase the modeling accuracy compared to FIR filters with the same number of linear filter parameters. Another aim is to test the robustness of the estimated sets of common poles to variations of the RTF. To do so, the poles estimated from training data are fixed and validated on data related to different source and/or receiver positions in the room, different from those used during training. Simulations are performed on the SUBRIR database [10], consisting of 24 low-frequency RIRs measured in a rectangular listening room using a B&K 4939 1/4" microphone and a custom Genelec 1094A subwoofer (12-150 Hz, ±3 dB) for 4 source positions x_s and 6 microphone positions y_m. Each RIR is downsampled to $f_s = 800$ Hz and truncated to $N_h = 1600$ samples from the direct path component, selected as its starting point, and normalized in energy.

First, the poles are estimated from the training data. The training data are obtained from a SIMO room acoustic system, where the input signal $u(n)$ is a zero mean white noise sequence of length $N_p N_f$ (with block length $N_f = 2N_h$), which is convolved with $R = 6$ RIRs, corresponding to 2 source positions and 3 microphone positions, obtaining R different output sequences. There are $V = 120$ possible combinations of 2 source and 3 microphone positions. $M = 10$ different realizations of the input sequence for each combination give a training set containing $VMR = 7200$ input-output sequences. The BB-OBF-MP algorithm is run on the training data for each combination and each realization to obtain a set of $N_p = 20$ poles common to the 6 RIRs included in that specific combination. The pole grid $\boldsymbol{p}_g$ used in the BB-OBF-MP algorithm has $G = 3000$ poles with 10 different radii distributed logarithmically from 0.9 to 0.995 and with 300 different angles placed uniformly between 1Hz and 200Hz. The step size for the estimation is set to $\mu_e = 0.001$. The estimated set of poles for each combination and realization is then kept fixed for validation; two sets of validation data are considered: set A contains data related to RIRs measured for the same 2 source positions, but for 3 different microphone positions, w.r.t. the corresponding training data, so that, for instance, the poles estimated for the combination $C_t^1 = \{x_1, x_2, y_1, y_2, y_3\}$ are validated on the data corresponding to the combination $C_A^1 = \{x_1, x_2, y_4, y_5, y_6\}$ (and repeated for $M = 10$ input realizations), while set B contains data related to RIRs measured for 2 different source positions and 3 different microphone positions, w.r.t. the corresponding training data, so that the poles are validated on the combination $C_B^1 = \{x_3, x_4, y_4, y_5, y_6\}$. Validation is performed using adaptive OBF filters with step size $\mu_v = 0.003$ on the validation sets A and B, and also on the training data for comparison. In addition, the training data is used to identify the system with adaptive FIR filters with the same number N_θ of adaptive parameters, using LMS with the same step size. The linear filter parameters during validation are all initialized to zero.

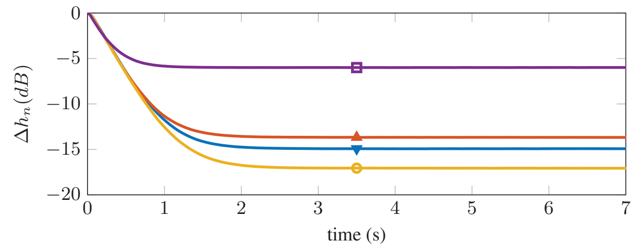

Fig. 3. The misadjustment in (10) for adaptive FIR filters on the training set (□) and for adaptive OBF filters on the training set (○) and on the validation set A (▽) and B (△) for $N_\theta = 40$.

The error measure used to compare the performance in the validation is the misadjustment Δh_n averaged over all different combinations, over all realizations and over the 6 RIRs in each combination, and is defined as

$$\Delta h_n = 10 \log_{10} \left[\frac{1}{VMR} \sum_{v=1}^{V} \sum_{m=1}^{M} \sum_{r=1}^{R} \frac{\|\boldsymbol{h}^r - \hat{\boldsymbol{h}}_{v,m}^{r,n}\|_2^2}{\|\boldsymbol{h}^r\|_2^2} \right], \quad (10)$$

with $\boldsymbol{h}^r$ indicating the r-th measured target RIR in each combination and $\hat{\boldsymbol{h}}_{v,m}^{r,n}$ the approximated RIR of $\boldsymbol{h}^r$ obtained using the set of pole estimated from the training data for the v-th combination and with the linear filter parameters estimated in the validation at sample n for the m-th realization of the input sequence. The approximated RIR $\hat{\boldsymbol{h}}_{v,m}^{r,n}$ is obtained as a linear combination of the length-N_h impulse responses of the OBFs built from the common poles estimated in the training, weighted by the linear filter parameters updated at sample n. The results of the misadjustment for $N_p = 20$ poles (corresponding to $N_\theta = 40$ linear filter parameters) are shown in Figure 3. It can be seen that the modeling accuracy is significantly increased using OBF filters compared to FIR filters with comparable convergence rate. The convergence rate for adaptive OBF models seems not to depend on the distance between the true poles and the poles of the system, since the same convergence rate results from all data sets. Moreover, common poles estimated on the training data provide good modeling accuracy when used to model RIRs for different source-microphone positions, also when the variations of the RTF are particularly significant (validation data set B). This fact suggests the possibility of estimating the poles for a finite number of source-microphone positions within a room in order to obtain a filter parametrization to be used anywhere in the same room, thus avoiding cumbersome computations for the adaptation of the poles.

5. CONCLUSIONS AND FUTURE WORK

In this paper, a room acoustic system identification algorithm for the estimation of the poles of adaptive OBF filters was proposed. A common set of poles estimated from multichannel data proved to reduce significantly the estimation error compared to FIR filters and to be quite robust to variations in the RTF, suggesting the possibility of keeping the poles fixed after estimation, thus avoiding adaptation of the pole position. Future research will focus on a better understanding of the common pole parametrization and on moving toward practical applications of the identification algorithm (e.g. by including nonwhite input signals).

6. REFERENCES

[1] J. J. Shynk. Adaptive IIR filtering. *ASSP Magazine, IEEE*, 6(2):4–21, 1989.

[2] P.S.C. Heuberger, P.M.J. van den Hof, and B. Wahlberg. *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer, 2005.

[3] G. Vairetti, E. De Sena, M. Catrysse, et al. A scalable algorithm for physically motivated and sparse approximation of room impulse responses with orthonormal basis functions. Technical report, KU Leuven, available at ftp://ftp.esat.kuleuven.be/pub/SISTA//gvairett/15-122.pdf, 2015.

[4] G. Williamson and S. Zimmermann. Globally convergent adaptive IIR filters based on fixed pole locations. *IEEE Trans. Signal Process.*, 44(6):1418–1427, 1996.

[5] B. Ninness and J.C. Gómez. Frequency domain analysis of tracking and noise performance of adaptive algorithms. *IEEE Treans. Signal Process.*, 46(5):1314–1332, 1998.

[6] A.N. Kaelin, A.G. Lindgren, and G.S. Moschytz. Simplified adaptive IIR filters based on optimized orthogonal prefiltering. *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, 42(5):326–333, 1995.

[7] B. Ninness, S. Gibson, and S. Weller. Practical aspects of using orthonormal system parameterisations in estimation problems. In *12th IFAC Symp. on Syst. Identification (SYSID 2000)*, 2000.

[8] T. Paatero and M. Karjalainen. Kautz filters and generalized frequency resolution: Theory and audio applications. *J. Audio Eng. Soc.*, 51(1/2):27–44, 2003.

[9] G. Vairetti, T. van Waterschoot, M. Moonen, et al. An automatic model-building algorithm for sparse approximation of room impulse responses with orthonormal basis functions. In *Proc. 14th Int. Workshop Acoust. Signal Enhancement (IWAENC 2014)*, pages 249–253, Antibes, France, September 2014.

[10] G. Vairetti, E. De Sena, M. Moonen, et al. A physically motivated parametric model for compact representation of room impulse responses based on orthonormal basis functions. In *Proc. of the 10th Eur. Congr. and Expo. on Noise Control Eng. (EURONOISE 2015)*, pages 149–154, Maastricht, The Netherlands, Jun 2015.

[11] L. SH Ngia. Separable nonlinear least-squares methods for efficient off-line and on-line modeling of systems using Kautz and Laguerre filters. *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, 48(6):562–579, 2001.

[12] L. SH Ngia. Recursive identification of acoustic echo systems using orthonormal basis functions. *IEEE Trans. Speech Audio Process.*, 11(3):278–293, 2003.

[13] G. Williamson, J.P. Ashley, and M. Nayeri. Structural issues in cascade-form adaptive IIR filters. In *Int. Conf. Acoust., Speech, Signal Process. (ICASSP-95)*, volume 2, pages 1436–1439, 1995.

[14] G. Vairetti, E. De Sena, M. Catrysse, et al. Room acoustic system identification using orthonormal basis function models. In *Proc. 60th Int. Conf. Audio Eng. Soc., Leuven, Belgium,*, 2016.

[15] J.O. Smith. *Introduction to Digital Filters with Audio Applications*. online book, 2007 edition, available at http://ccrma.stanford.edu/∼jos/filters/, accessed Jul. 2015.

[16] T. van Waterschoot and M. Moonen. A pole-zero placement technique for designing second-order IIR parametric equalizer filters. *IEEE Trans. Audio Speech Lang. Process.*, 15(8):2561–2565, 2007.

[17] B. Ninness and F. Gustafsson. A unifying construction of orthonormal bases for system identification. *IEEE Trans. Automatic Control*, 42(4):515–521, 1997.

[18] S. Haykin. *Adaptive filter theory*. Pearson Education India, 2008.

[19] B. Ninness and H. Hjalmarsson. Model structure and numerical properties of normal equations. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, 48(4):425–437, 2001.