KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Vairetti G., De Sena E., Catrysse M., Jensen S.H., Moonen M., van Waterschoot T. (2015), **Room acoustic system identification using orthonormal basis function models** Published in Proc. AES 60th International Conference on Dereverberation and Reverberation of Audio, Music and Speech, Leuven, Belgium, Feb. 2016.
Archived version	Author manuscript: the content is identical to the content of the accepted paper, but without the final typesetting by the publisher.
Published version	http://www.aes.org/e-lib/browse.cfm?elib=18086
Journal homepage	http://www.aes.org/conferences/60/
Author contact	giacomo.vairetti@esat.kuleuven.be *+ 32 (0)16 321817*
IR	url in Lirias https://lirias.kuleuven.be/handle/123456789/520688/2/15-121.pdf

(article begins on next page)

Room acoustic system identification using orthonormal basis function models

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

[1]*KU Leuven, Dept. of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, 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.*

Correspondence should be addressed to Giacomo Vairetti (giacomo.vairetti@kuleuven.be)

ABSTRACT

Parametric modeling is used in all those acoustic signal enhancement applications that require to model and identify a room impulse response (RIR) in a compact yet accurate way. Fixed-pole models based on orthonormal basis functions (OBFs) provide advantages over all-zero and pole-zero models. The parameters of an OBF model can be estimated from a measured target RIR by a scalable matching pursuit algorithm called OBF-MP. However, a measurement for the RIR is usually not available, and the model parameters should be estimated from input-output data. This paper introduces a block-based version of OBF-MP for the modeling and identification of room acoustic systems, which represents an intermediate step toward a sample-based recursive implementation of the algorithm. Simulation results show modeling capabilities comparable with the original OBF-MP.

1. INTRODUCTION

Room acoustic signal enhancement applications often require to accurately model and identify an acoustic system using a small number of model parameters. Parametric models aim at representing a room transfer function (RTF) with a rational expression in the z-domain, which can be implemented as a digital filter, under the assumption of a room being a causal, stable and linear system. The most widely used are the all-zero models, which define a finite impulse response (FIR) filter as a truncation of the sampled room impulse response (RIR). However, all-zero models usually require a large number of parameters, whose values depend strongly on the source and receiver positions [1]. Pole-zero models [2], which define an infinite impulse response (IIR) filter, can be used in the attempt of overcoming the drawbacks of all-zero models. However, pole-zero models are seldom used in acoustic signal enhancement applications, mainly due to the fact that a marginal improvement in terms of modeling capabilities over all-zero models is not justified by an increase in the complexity of the model and by problems of instability and convergence to local minima, which can arise from the nonlinear estimation of the model parameters [3].

An IIR filter can alternatively be expressed in a parallel form by a partial fraction expansion of the pole-zero transfer function [4], obtaining a model with a tap-transversal structure, but nonlinear in the pole parameters. Pairs of complex conjugate poles can be combined together obtaining a parallel of second-order resonators, whose impulse responses are sinusoids decaying exponentially in time. Models using a parallel IIR filter, sometimes called parallel second-order filter [5], are particularly relevant in the modeling of room acoustic systems due to their analogy to the Green's function of the acoustic wave equation [6, 7, 8], and the possibility of representing room resonances by fixing the poles in the model structure. Orthonormal basis function (OBF) models [9], which derive from an orthonormalization of such parallel structure, provide some desirable properties that can be exploited in the estimation of the model parameters. A scalable matching pursuit algo-

rithm named OBF-MP was proposed in [10], where the nonlinear problem of estimating the poles was avoided by defining a grid of candidate poles and iteratively selecting the poles that provide the best approximation of a measured target RIR. It has been shown that OBF models provide advantages compared to all-zero and pole-zero models in the approximation of both single- [8, 10] and multi-channel [11] RIRs. In practical applications, however, a measurement of the RIR is usually not available, and the poles should be estimated directly from input-output data. In [12], a recursive separable nonlinear least-squares (LS) method was proposed for the estimation of both poles and linear coefficients and applied to the identification of acoustic echo systems [13], but the method was limited to the case of OBF models with a single repeated pole.

In this paper, a block-based version of the OBF-MP algorithm is proposed for the estimation of the poles of an OBF model for the identification of an acoustic system from input-output data. In Section 2, a brief overview of OBF models using multiple poles is provided. In Section 3, the block-based OBF-MP algorithm, henceforth called BB-OBF-MP, is described. In Section 4, simulation results show the performance of the modified algorithm in comparison with OBF-MP in the identification of measured RIRs. In Section 5, results and future research are discussed.

2. OBF MODELS

Parametric modeling of room acoustics using OBF models consists in approximating a RIR as a linear combination of basis functions, which are orthonormalized versions of exponentially decaying sinusoids, describing resonances in the frequency domain. Thus, a RTF is modeled as a superposition of resonances, whose frequency and bandwidth are determined by the position of the poles. A resonance can be described in the z-domain by a second order resonator with transfer function

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

with $\boldsymbol{p}_i = \{p_i, p_i^*\} = \rho_i e^{\pm j\vartheta_i}$ a pair of complex-conjugate poles and * indicating complex conjugation. The radius $\rho_i = e^{-\zeta_i/f_s}$ defines the bandwidth of the resonance, or equivalently the rate of the exponential decay of the sinusoid, determined by the damping constant ζ_i (f_s is the sampling rate), while its angle $\vartheta_i = \omega_i/f_s$ determines the resonance frequency ω_i. Notice that the frequency of the magnitude peak in the frequency response is not exactly equal to ω_i, due to the influence of the complex conjugate pole (especially when the resonance has large bandwidth) [14, 15]. Multiple resonances can be modeled by arranging in parallel multiple filters having a transfer function as in (1) and poles can be placed arbitrarily over the unit disc, providing stability and flexibility in the allocation of spectral resolution.

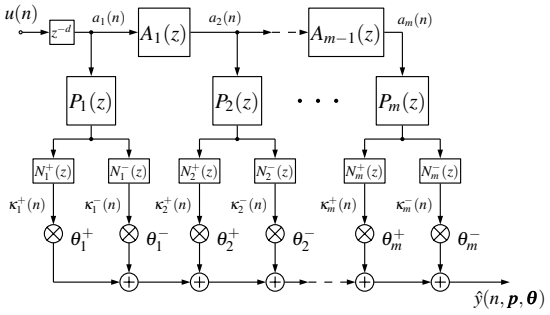

Fig. 1: The generalized OBF model structure.

OBF models originate from the orthonormalization of a parallel of second-order resonators, where orthogonality between any two consecutive basis functions is obtained by a second-order all-pass filter,

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

in which the zeros in $1/p_i$ and $1/p_i^*$ ensure that the basis functions defined by $\boldsymbol{p}_{i+1}$ are orthogonal to those generated by $\boldsymbol{p}_i$. Orthonormality between the two basis functions of each pole pair is enforced by a pair of orthonormalization filters $N_i^{\pm}(z)$. The generalized filter structure of OBF models is shown in Figure 1 for m pairs of complex-conjugate poles. Different models can be obtained through the choice of the orthonormalization filters, as explained in [16]. Here, the so-called Kautz model is used, where

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

A pair of OBFs have transfer functions given by $\Psi_i^{\pm}(z, \boldsymbol{p}_i) = N_i^{\pm}(z) P_i(z) \prod_{j=1}^{i-1} A_j(z)$ (cfr. Figure 1), so that the tap-outputs $\kappa_i^{\pm}(n, \boldsymbol{p}_i) = \{\kappa_i^+(n, \boldsymbol{p}_i), \kappa_i^-(n, \boldsymbol{p}_i)\}$ of each pair of OBFs are equal to the input signal $u(n)$ filtered by $\Psi_i^{\pm}(z, \boldsymbol{p}_i)$ (with $n = \iota/f_s$ the discrete time

variable). Also, the OBFs form a complete set in the Hardy space on the unit disc, so that any stable linear filter can be approximated with arbitrary accuracy by a linear combination of a finite number of OBFs [16]. OBF models are linear in the parameters $\boldsymbol{\theta}_i = \{\theta_i^+, \theta_i^-\}$ (where $i = 1,\ldots,m$), so that for a fixed set of pole pairs $\boldsymbol{p} = \{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_m\}$, the problem of estimating the values of the parameters in $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_m\}$ becomes linear and can be solved in closed form using linear regression. It follows that, for a given input-output data set $\{\boldsymbol{u},\boldsymbol{y}\} = \{u(n),y(n)\}_{n=1}^N$, the approximation of the output sequence $\boldsymbol{y}$ as a linear combination of the tap-outputs $\boldsymbol{\kappa}_i^{\pm}(\boldsymbol{p}_i) = \{\kappa_i^{\pm}(n,\boldsymbol{p}_i)\}_{n=1}^N$ (with $i = 1,\ldots,m$) of the basis transfer functions $\Psi_i^{\pm}(z,\boldsymbol{p}_i)$ for an input sequence $\boldsymbol{u}$, is given by $\hat{\boldsymbol{y}} = \boldsymbol{K}(\boldsymbol{p})\boldsymbol{\theta}$, where $\boldsymbol{K}(\boldsymbol{p})$ is a matrix whose columns are the tap-output sequences $\boldsymbol{\kappa}_i^{\pm}(\boldsymbol{p}_i)$. The values of the parameters in $\boldsymbol{\theta}$ can be estimated in a LS sense as $\hat{\boldsymbol{\theta}} = \boldsymbol{K}^{\dagger}\boldsymbol{y} = (\boldsymbol{K}^T\boldsymbol{K})^{-1}\boldsymbol{K}^T\boldsymbol{y}$ (with † indicating the Moore-Penrose pseudoinverse and ^{T} the transpose).

Orthogonality of the basis functions ensures that the LS estimation is numerically well-conditioned (the matrix $\boldsymbol{K}$ has a small condition number). When the input sequence $\boldsymbol{u}$ is white (with constant spectral density $S_{\boldsymbol{u}} = c$) and the number of samples of $\boldsymbol{u}$ is larger than the number of OBFs, the matrix $\boldsymbol{K}$ is orthogonal up to a scaling factor c. As a consequence, $\boldsymbol{K}$ is optimally conditioned and the autocorrelation matrix $\boldsymbol{R} = \boldsymbol{K}^T\boldsymbol{K}$ in the LS solution becomes a scaled identity matrix, i.e. $\boldsymbol{R} = c\boldsymbol{I}$ (if $\boldsymbol{u}$ is a unit-variance white noise signal, $\boldsymbol{R}$ reduces to an identity matrix [8, 10]). A finite-length random sequence is not exactly white. As a consequence, the columns of $\boldsymbol{K}$ are not perfectly orthogonal and $\boldsymbol{R}$ no longer identical, thus requiring the computation of $\boldsymbol{K}^{\dagger}$. However, OBF models are particularly robust in terms of numerical well-conditioning also when the input signal is not white [9], so that the entity of numerical errors remains small, also when a large number of basis functions is used.

For simplicity, the term *pole* will refer in the following to a pair of complex-conjugate poles ($\boldsymbol{p}_i = \{p_i, p_i^*\}$).

3. BLOCK-BASED OBF-MP ALGORITHM

The main problem with the modeling and identification of a RIR using OBF models consists in estimating the pole parameters, which appear in the denominator of the model transfer function, hence requiring in principle nonlinear optimization techniques (see e.g. the method in [12]). This nonlinear problem was avoided in [8, 10] by proposing a scalable matching pursuit algorithm, called OBF-MP, which estimates the poles of an OBF model used to approximate a measured target RIR by searching for the pair of basis functions which is most correlated with the target RIR. The search is performed on a dictionary of OBFs generated by a set of candidate poles $\boldsymbol{p}_g = \{\boldsymbol{p}_i\}$ (with $i = 1,\ldots,G$), distributed arbitrarily on the unit disc. The dictionary is constructed by the impulse responses $\varphi_i^{\pm}(n,\boldsymbol{p}_i)$ to the transfer functions $\Psi_i^{\pm}(z,\boldsymbol{p}_i)$ for each pole in $\boldsymbol{p}_g$ and updated at each iteration by generating a new set of basis functions, all orthonormal to the OBFs selected at the previous iterations, according to the model structure in Figure 1.

A block-based version of the OBF-MP algorithm, named BB-OBF-MP, is proposed here for estimating the pole parameters from input-output data $\{\boldsymbol{u}_f,\boldsymbol{y}_f\}$, with f the block index, which represents an intermediate step toward a sample-based recursive implementation of the algorithm. The aim is to estimate a new pole $\boldsymbol{p}_j$ to be included in the active pole set $\boldsymbol{p}_f^{\mathscr{A}}$, one pole per block. The search is performed as in OBF-MP, with the difference that, since the tap-outputs of the OBF structure are no longer orthogonal to each other, it should be done with respect to the residual data sequence and not to the output data sequence directly, as explained below. Another difference with OBF-MP is that the input-output data change at every block. Consequently, the tap-outputs related to the poles selected in the previous blocks have to be recomputed and the linear coefficients re-estimated for the new input data block. The BB-OBF-MP algorithm, listed below in detail and schematized in Figure 2, can be divided into two main steps: (i) the computation of the current residual data sequence $\boldsymbol{\varepsilon}_f$, with the estimation of the linear coefficients $\hat{\boldsymbol{\theta}}_f$ of the Kautz filter $K(z,\boldsymbol{p}_{f-1}^{\mathscr{A}})$ built from the poles in the current active set $\boldsymbol{p}_{f-1}^{\mathscr{A}}$, and (ii) the subsequent estimation of a new pole $\boldsymbol{p}_j$ to be included in $\boldsymbol{p}_f^{\mathscr{A}}$. The only exception is for the first block ($f = 1$), where only the second step is performed, the active pole set empty ($\boldsymbol{p}_0^{\mathscr{A}} = \emptyset$) and the current residual being set as $\boldsymbol{\varepsilon}_1 = \boldsymbol{y}_1$.

At each block f, an approximation $\hat{\boldsymbol{y}}_f$ of the current output data block $\boldsymbol{y}_f$, defined for the block length N_f as

$$\boldsymbol{y}_f = \{y(n)\}_{n=1+fN_f}^{(f+1)N_f}, \tag{4}$$

is obtained by filtering the current input data block $\boldsymbol{u}_f$ with a Kautz filter $K(z,\boldsymbol{p}_{f-1}^{\mathscr{A}})$ built from the poles in the current active set $\boldsymbol{p}_{f-1}^{\mathscr{A}}$ selected from previous blocks,

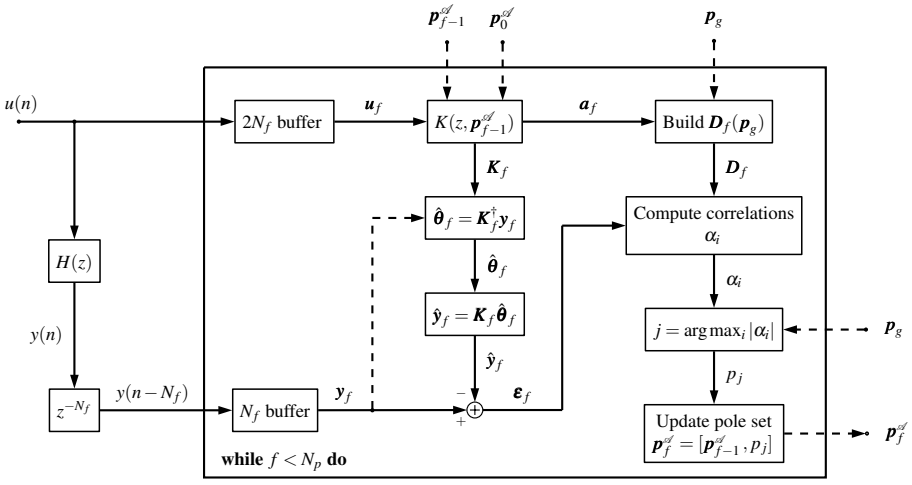

Fig. 2: The BB-OBF-MP algorithm in a schematic representation. Inbound dashed lines represent initial conditions and inputs, while outbound dashed lines represent outputs.

where the values of the coefficients $\hat{\boldsymbol{\theta}}_f$ are estimated using linear regression, as described in Section 2. In order to obtain the initial conditions of $K(z, \boldsymbol{p}_{f-1}^{\mathscr{A}})$, the previous N_f samples of the input signal $u(n)$ are fed to the filter. For this reason, the block length N_f should not be shorter than the length N_h of the RIR (or at least the estimated reverberation time of the room). This is equivalent to using a length $2N_f$ input data block (with $N_f > N_h$)

$$\boldsymbol{u}_f = \{u(n)\}_{n=1+(f-1)N_f}^{(f+1)N_f}. \tag{5}$$

The matrix $\boldsymbol{K}_f$, whose columns are the last N_f samples of the tap-outputs $\boldsymbol{\kappa}_l^{\pm}(\boldsymbol{p}_{f-1}^{\mathscr{A}})$ (with $l = 1, \ldots, f-1$) of the Kautz filter $K(z, \boldsymbol{p}_{f-1}^{\mathscr{A}})$, is then used to estimate the optimal values, in a LS sense, of the linear coefficients $\hat{\boldsymbol{\theta}}_f = \boldsymbol{K}_f^{\dagger} \boldsymbol{y}_f$. The approximation of the current output data block $\boldsymbol{y}_f$ is obtained as $\hat{\boldsymbol{y}}_f = \boldsymbol{K}_f \hat{\boldsymbol{\theta}}_f$ and the current residual data sequence as $\boldsymbol{\varepsilon}_f = \boldsymbol{y}_f - \hat{\boldsymbol{y}}_f$.

The second step of the algorithm consists in searching for the pole that produces the new pair of tap-output sequences which is most correlated with the current residual $\boldsymbol{\varepsilon}_f$. A grid $\boldsymbol{p}_g$ of G candidate poles has to be defined on the unit disc based on some prior knowledge of the acoustics of the room or some particular desired frequency resolution. For each pole $\boldsymbol{p}_i \in \boldsymbol{p}_g$ (with $i = 1, \ldots, G$), the sequences $\boldsymbol{\kappa}_{f,i}^{\pm}(\boldsymbol{p}_i) = \{\boldsymbol{\kappa}_{f,i}^{+}, \boldsymbol{\kappa}_{f,i}^{-}\}$ are obtained as the f-th tap-output sequences for an OBF model built from the poles in $\boldsymbol{p}_{f-1}^{\mathscr{A}}$ (cfr. Figure 1), with $\boldsymbol{\kappa}_{f,i}^{\pm}(\boldsymbol{p}_i)$ almost orthogonal, for the reasons explained in Section 2, to the tap-outputs $\boldsymbol{\kappa}_l^{\pm}(\boldsymbol{p}_{f-1}^{\mathscr{A}})$ computed in the first step. The tap-outputs $\boldsymbol{\kappa}_{f,i}^{\pm}(\boldsymbol{p}_i)$ are obtained by filtering the input data block $\boldsymbol{u}_f$ with the transfer functions $\Psi_i^{\pm}(z, \boldsymbol{p}_i) = N_i^{\pm}(z) P_i(z) \prod_{j=1}^{f-1} A_j(z)$, where the product corresponds to the series of all-pass filters defined by the poles in $\boldsymbol{p}_{f-1}^{\mathscr{A}}$ (cfr. Figure 1). Then, $\boldsymbol{\kappa}_{f,i}^{\pm}(\boldsymbol{p}_i)$ can be computed by filtering the output of the all-pass series $a_f(n) = \prod_{j=1}^{f-1} A_j(z) u(n)$ (where n is defined as in Eq. 5) with pairs of filters with transfer functions $\Gamma_i^{\pm}(z, \boldsymbol{p}_i) = N_i^{\pm}(z) P_i(z)$. The dictionary $\boldsymbol{D}_f$ is a matrix whose columns are the last N_f samples of the tap-output sequences $\boldsymbol{\kappa}_{f,i}^{\pm}(\boldsymbol{p}_i)$ built for each pole $\boldsymbol{p}_i \in \boldsymbol{p}_g$. The correlation of a pair of tap-output sequences $\boldsymbol{\kappa}_{f,i}^{\pm}(\boldsymbol{p}_i) \in \boldsymbol{D}_f$ of length N_f with the current residual $\boldsymbol{\varepsilon}_f$ is computed as

$$\alpha_{f,i} = \sqrt{\alpha_{i^+}^2 + \alpha_{i^-}^2} = \sqrt{(\boldsymbol{\kappa}_{f,i}^{+\,T} \boldsymbol{\varepsilon}_f)^2 + (\boldsymbol{\kappa}_{f,i}^{-\,T} \boldsymbol{\varepsilon}_f)^2}. \tag{6}$$

The pair of tap-output sequences in the dictionary having maximum correlation is selected according to $j = \arg\max_i \alpha_i$ and the corresponding pole $\boldsymbol{p}_j \in \boldsymbol{p}_g$ is added to the active pole set $\boldsymbol{p}_f^{\mathscr{A}}$. Finally, the algorithm moves to the next block ($f = f + 1$) and the two steps are repeated for a new input-output data set $\{\boldsymbol{u}_f, \boldsymbol{y}_f\}$, until a desired number of poles N_p has been estimated or the energy of the current residual sequence falls below a certain threshold.

4. SIMULATION RESULTS

The objective of both OBF-MP and its block-based version is to model and identify a room acoustic system

Algorithm 1 BB-OBF-MP

$\boldsymbol{p}_g = \{p_1, \ldots, p_G\}$ ▷ Define poles in the pole grid
$\boldsymbol{p}_0^{\mathscr{A}} = \emptyset$, $f = 1$ ▷ $p_f^{\mathscr{A}}$: active poles set, f: block index
while $f < N_p$ **do** ▷ N_p: desired number of poles
  $\boldsymbol{u}_f = \{u(n)\}_{n=1+N_f(f-1)}^{N_f(f+1)}$ ▷ current input block (length $2N_f$)
  $\boldsymbol{y}_f = \{y(n)\}_{n=1+N_f f}^{N_f(f+1)}$ ▷ current output block (length N_f)
  if $f > 1$ **then**
    Build $\boldsymbol{K}_f(\boldsymbol{p}_{f-1}^{\mathscr{A}}, \boldsymbol{u}_f)$ ▷ $\boldsymbol{K}_f$: matrix of tap-outputs $\boldsymbol{\kappa}_i^{\pm}$
    $\hat{\boldsymbol{\theta}}_f = \boldsymbol{K}_f^{\dagger} \boldsymbol{y}_f$ ▷ LS estimation of linear coefficients
    $\boldsymbol{\varepsilon}_f = \boldsymbol{y}_f - \boldsymbol{K}_f \hat{\boldsymbol{\theta}}_f$ ▷ current residual block (length N_f)
  end if
  Build $\boldsymbol{D}_f(\boldsymbol{p}_g, \boldsymbol{a}_f)$ ▷ $\boldsymbol{D}_f$: dictionary of tap-outputs $\boldsymbol{\kappa}_{f,i}^{\pm}$
  $j = \arg\max_i |\alpha_i|$ ▷ Max. correlation of $\boldsymbol{\kappa}_{f,i}^{\pm} \in \boldsymbol{D}_f$ with $\boldsymbol{\varepsilon}_f$
  $\boldsymbol{p}_f^{\mathscr{A}} = [\boldsymbol{p}_{f-1}^{\mathscr{A}}, \boldsymbol{p}_j]$ ▷ add selected pole to active pole set
  $f = f + 1$ ▷ Increase block index
end while

$H(z)$ using a Kautz model $K(z, \boldsymbol{p}, \boldsymbol{\theta})$ (cfr. Figure 2). Simulation results presented here aim at comparing the performances in the approximation of a target sampled RIR $\boldsymbol{h}$ of length N_h samples, provided by the poles $\boldsymbol{p}_{N_p}^{\mathscr{B}}$ estimated directly from $\boldsymbol{h}$ with the OBF-MP algorithm and by the poles $\boldsymbol{p}_{N_p}^{\mathscr{A}}$ obtained from input-output data with the BB-OBF-MP algorithm. When both pole sets are estimated, they are used in a Kautz filter fed with an impulsive input sequence $\boldsymbol{\delta} = \{\delta(n)\}_{n=0}^{N_h-1}$, where the linear coefficients of the filter are computed as $\hat{\boldsymbol{\theta}} = \boldsymbol{\Phi}^T \boldsymbol{h}$. The columns of $\boldsymbol{\Phi}$ are the tap-outputs $\boldsymbol{\varphi}_i^{\pm}(n, \boldsymbol{p}_i) = \Psi_i^{\pm}(z, \boldsymbol{p}_i)\delta(n)$ (with $n = 0, \ldots, N_h - 1$) built from the poles in the active pole set $\boldsymbol{p}_{N_p}^{\mathscr{X}}$ (with $\mathscr{X}$ being either $\mathscr{A}$ or $\mathscr{B}$), corresponding to the OBF impulse responses. The setup used for computing the approximation error sequence is given in Figure 3.

The input data set $\boldsymbol{u}$ is a zero mean white noise sequence, which is convolved with $R = 24$ RIRs taken from the SUBRIR database [11], obtaining R different output data sets $\boldsymbol{y}^r$. The database consists of 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 24 source-microphone positions. Each RIR was downsampled to $f_s = 800$ Hz

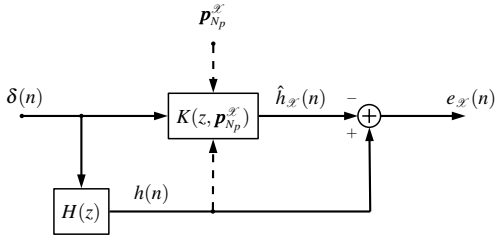

Fig. 3: Setup for the computation of the approximation error sequence for a given set of poles $\boldsymbol{p}_{N_p}^{\mathscr{X}}$.

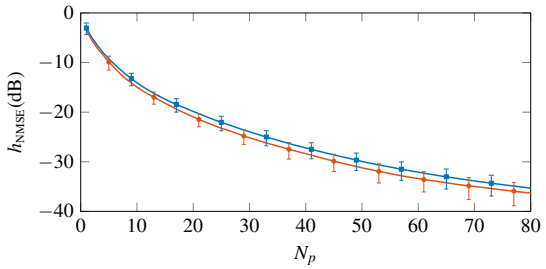

Fig. 4: The average NMSE in (7) vs. the number of poles N_p in the active pole set ($R = 24, M = 10$). Poles are estimated using OBF-MP (•) and using BB-OBF-MP (▪). The vertical lines represent standard deviations.

and truncated to $N_h = 1600$ samples from the direct path component, selected as its starting point. The pole grid $\boldsymbol{p}_g$ used in both algorithms has $G = 3000$ poles with 10 different radii distributed uniformly from 0.9 to 0.99 and with 300 different angles placed uniformly in the range $[1, f_s/2 - 1]$ Hz.

The BB-OBF-MP algorithm was run on each input-output data set $\{\boldsymbol{u}, \boldsymbol{y}^r\}$, for $M = 10$ different realizations of the input sequence $\boldsymbol{u}$. The block length N_f equals the length N_h of the RIRs, so that $\boldsymbol{u}$ has length $(N_p^{\max} + 1)N_h$ samples, with $N_p^{\max}$ the maximum number of desired poles.

The error measure used to compare the performance in the approximation of the target RIR using the poles estimated with the two algorithms is the Normalized Mean-Square-Error (NMSE) in the time domain, averaged over all RIRs and over all realizations. The NMSE is a measure of the energy of the approximation error sequence, normalized w.r.t. the energy of the RIR, and is given by

$$h_{\mathrm{NMSE}} = 10 \log_{10} \left[\frac{1}{MR} \sum_{m=1}^{M} \sum_{r=1}^{R} \frac{\|\boldsymbol{h}^r - \hat{\boldsymbol{h}}_{\mathscr{X}}^{r,m}\|_2^2}{\|\boldsymbol{h}^r\|_2^2} \right], \qquad (7)$$

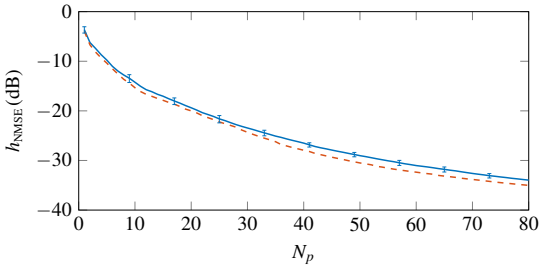

Fig. 5: The average NMSE in (7) vs. the number of poles N_p in the active pole set for a single RIR ($R = 1$, $M = 10$). Poles are estimated using OBF-MP (- - -) and using BB-OBF-MP (——). The vertical lines represent standard deviations.

with $\boldsymbol{h}^r$ indicating the r-th measured target RIR and $\hat{\boldsymbol{h}}^{r,m}_{\mathscr{X}}$ the approximated response of $\boldsymbol{h}^r$ obtained using the pole set $\boldsymbol{p}^{\mathscr{X}}_{N_p}$ for the m-th realization of the input sequence. The results of the NMSE w.r.t. the number of poles N_p in the active set (with $N_p^{\max} = 80$) for the SUBRIR database are shown in Figure 4. It can be seen that when the poles are estimated from input-output data using BB-OBF-MP, the average NMSE is comparable with the error obtained when the poles are estimated directly from the RIR using OBF-MP (for a given number of poles N_p, the difference is less than 1.5 dB). Figure 4 also reports the standard deviations for the NMSE (only for some values of N_p for a better visualization), showing that the results of the two algorithms are comparable also in terms of variability. This means that the entity of the approximation error does not depend strongly on the particular realization of the input sequence $\boldsymbol{u}$ and that the standard deviation of the results for both algorithms depends on the different characteristics of the RIRs in the database. The NMSE for a single RIR and for $M = 10$ realizations of $\boldsymbol{u}$ is depicted in Figure 5, showing much smaller standard deviations compared to the values in Figure 4.

The differences in the approximation error are a consequence of the fact that poles selected by the two algorithms are distributed similarly on the unit disc, but are not exactly the same poles. Figure 6 shows the active pole sets $\boldsymbol{p}^{\mathscr{A}}_{N_p}$ and $\boldsymbol{p}^{\mathscr{B}}_{N_p}$ for $N_p = 30$ estimated using BB-OBF-MP and OBF-MP, respectively, for the approximation of a single RIR (depicted at the top of Figure 7), and for one particular realization of $\boldsymbol{u}$. Figure 7 also shows the approximation error sequences obtained by these active pole sets, which correspond for both cases to a NMSE of about -23 dB (cfr. Figure 5). It can be

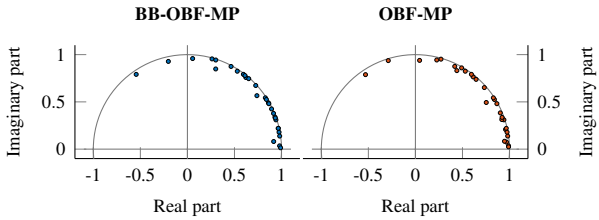

Fig. 6: The active pole sets $\boldsymbol{p}^{\mathscr{A}}_{30}$ and $\boldsymbol{p}^{\mathscr{B}}_{30}$ estimated using BB-OBF-MP (left) and OBF-MP (right) for the approximation of a target RIR (cfr. Figure 7) for one particular realization of $\boldsymbol{u}$. The complex-conjugate poles in the lower half disc are not shown.

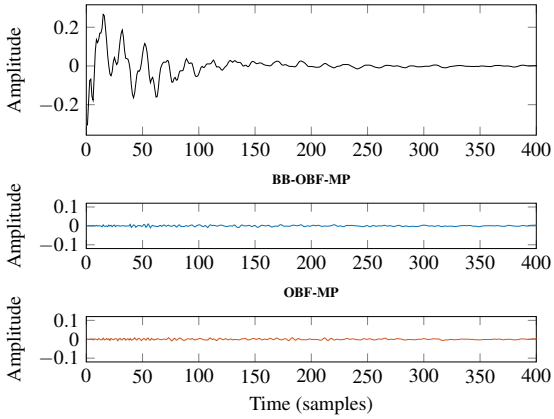

Fig. 7: The first 500 ms of the target RIR used in the example and the corresponding approximation error sequences obtained for the pole sets shown in Figure 6.

seen that the early reflections are well approximated, but also late reflections are reduced in energy. However, the residual error is due to the fact that only strong resonances are well modeled, while weaker resonances are only approximated. This can be seen in Figure 8, where the magnitude responses of the approximation sequences produced with the poles sets $\boldsymbol{p}^{\mathscr{A}}_{N_p}$ and $\boldsymbol{p}^{\mathscr{B}}_{N_p}$ for $N_p = 30$ are depicted, overlapped to the magnitude response of the target RIR. This is a consequence of the fact that only $N_p = 30$ poles have been used, which is not enough to model all the resonances of the target response individually. This is also noticeable from the magnitude response of the approximation error sequences, where the resonances above 150 Hz, which are not modeled individually by a pole, are visible. In order to model also these resonances, and to improve the approximation error over

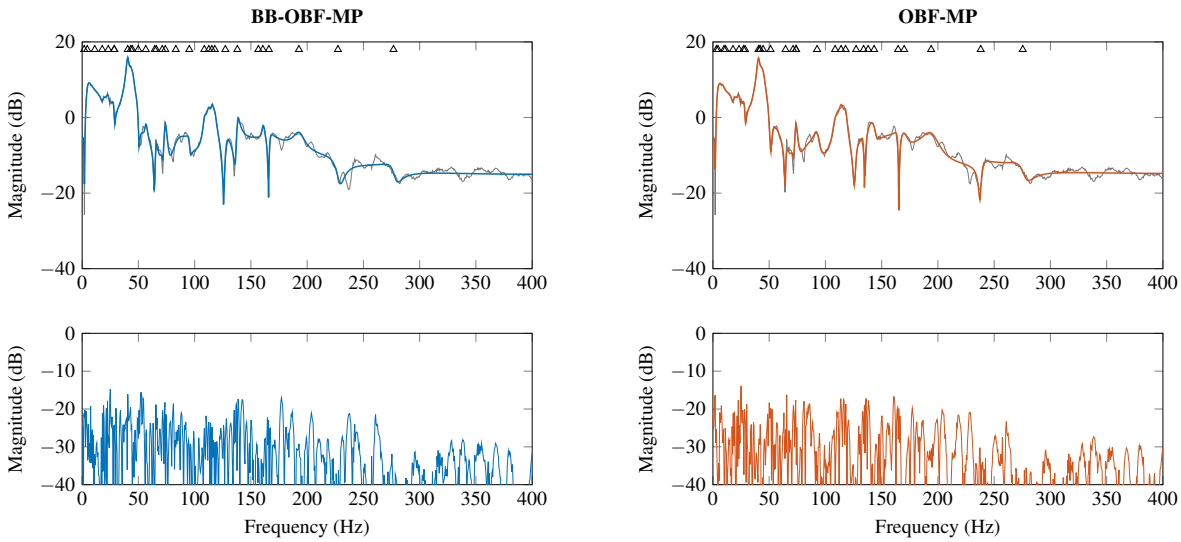

Fig. 8: The magnitude responses (top) of the target RIR (in gray) and of the approximation sequences obtained with BB-OBF-MP (left) and with OBF-MP (right). The frequency of the poles (△) in the active sets are reported above (cfr. Figure 6). The magnitude responses of the residual error sequences (cfr. Figure 7) are also shown (bottom).

the whole frequency range (since the same pole can be selected more than once, thus improving the approximation of resonances already identified), more poles should be added to the active set. In fact, as can be seen from Figure 4 and 5, the energy of the residual error sequences decreases as the number of poles in the active set increases. However, the NMSE decreases much slower when all the sinusoidal components of the RIR have been modeled and the energy of the residual sequence approaches the energy of the noise floor, which for the SUBRIR database is about -40 dB.

5. CONCLUSIONS AND FUTURE WORK

OBF models, which define an IIR filter, provide some advantages over all-zero and pole-zero models in the modeling and identification of a RTF, providing a more accurate and more compact representation of a measured RIR, and other desirable properties. The OBF-MP algorithm for the estimation of the pole parameters of an OBF model have been extended to the case for which a measured RIR is not available and the model parameters have to be estimated from input-output data. A block-based version of the algorithm, called BB-OBF-MP has been proposed, which estimates the pole parameters based on the correlation of the tap-outputs with the measured output signal, while the values of the linear coefficients are estimated with linear regression. Simulation results show that, provided that the block length is at least as long as the measured RIR, the poles estimated from input-output data provide an approximation of the target RIR almost as good as the approximation obtained from the poles estimated directly from the measured RIR using OBF-MP.

The BB-OBF-MP algorithm will be easily extended to the case in which a set of RTFs have to be modeled and identified jointly, in order to estimate a set of poles common to all RIRs, similarly to the algorithm presented in [11]. In this way, the total number of parameters to model a set of RIRs can be reduced and the parameter values would be less sensitive to changes in position of the source and the receiver, compared to the case in which the poles are estimated individually for each RIR (since not all the resonances of the room can be identified from a single input-output data set at a specific position of the source and the receiver). Future research will focus on the possibility of developing a sample-based recursive implementation of the algorithm, in the attempt to overcome the limitations imposed by the block length and make it applicable to acoustic signal enhancement applications.

6. ACKNOWLEDGMENT

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) was supported by a Postdoctoral Fellowship (F+/14/045) of the KU Leuven Research Fund. The scientific responsibility is assumed by its authors.

7. REFERENCES

[1] J. Mourjopoulos and M. A. Paraskevas, "Pole and zero modeling of room transfer functions," *J. Sound Vib.*, vol. 146, no. 2, pp. 281–302, 1991.

[2] G. Long, D. Shwed, and D. Falconer, "Study of a pole-zero adaptive echo canceller," *IEEE Trans. Circuits Syst.*, vol. 34, no. 7, pp. 765–769, 1987.

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

[4] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-time Signal Processing (2nd Ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1999.

[5] B. Bank, "Perceptually motivated audio equalization using fixed-pole parallel second-order filters," *IEEE Signal Process. Lett.*, vol. 15, pp. 477–480, 2008.

[6] H. Kuttruff, *Room acoustics*, 5th ed. Spon Press, 2009.

[7] Y. Haneda, Y. Kaneda, and N. Kitawaki, "Common-acoustical-pole and residue model and its application to spatial interpolation and extrapolation of a room transfer function," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 6, pp. 709–717, 1999.

[8] G. Vairetti, E. De Sena, M. Catrysse, S. H. Jensen, M. Moonen, and T. van Waterschoot, "A scalable algorithm for physically motivated and sparse approximation of room impulse responses with orthonormal basis functions," KU Leuven, Tech. Rep., 2015.

[9] P. Heuberger, P. van den Hof, and B. Wahlberg, *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer, 2005.

[10] G. Vairetti, T. van Waterschoot, M. Moonen, M. Catrysse, and S. H. Jensen, "An automatic model-building algorithm for sparse approximation of room impulse responses with orthonormal basis functions," in *Proc. 14th Int. Workshop Acoustic Signal Enhancement (IWAENC 2014)*, Antibes, France, Sep. 2014, pp. 249–253.

[11] G. Vairetti, E. De Sena, T. van Waterschoot, M. Moonen, M. Catrysse, N. Kaplanis, and S. H. Jensen, "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)*, Maastricht, The Netherlands, Jun 2015, pp. 149–154.

[12] L. S. 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.*, vol. 48, no. 6, pp. 562–579, 2001.

[13] ——, "Recursive identification of acoustic echo systems using orthonormal basis functions," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 3, pp. 278–293, 2003.

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

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

[16] B. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Trans. Automatic Control*, vol. 42, no. 4, pp. 515–521, 1997.