A Room Impulse Response Measurement Method Robust Towards Nonlinearities Based on Orthogonal Periodic Sequences

A Room Impulse Response Measurement Method Robust towards Nonlinearities based on Orthogonal Periodic Sequences

Alberto Carini, *Senior Member, IEEE,* Stefania Cecchi, *Member, IEEE,* Alessandro Terenzi, and Simone Orcioni, *Senior Member, IEEE*

***Abstract*—An open problem in room impulse response (RIR) measurement is the effect of nonlinearities, especially those with memory, present in the measurement system, specifically in the power amplifier and in the loudspeaker. The nonlinearities can corrupt the measurement introducing artifacts. The paper discusses a RIR measurement method that is robust towards these nonlinearities. The proposed methodology allows measuring the RIR using the cross-correlation method, i.e., computing the cross-correlation between the output signal and an appropriate sequence. In contrast to other cross-correlation based methods, the proposed approach directly estimates the first-order kernel of the Volterra filter modeling the measurement systems, i.e., the system impulse response for small signals. The proposed approach exploits the concepts of orthogonal periodic sequences, recently proposed in the literature. The input signal can be any periodic persistently exciting sequence and can also be a quantized sequence. Measurements performed both on an emulated scenario and in real environments illustrate the validity of the approach and compare it with other competing RIR measurement methods.**

***Index Terms*—Room impulse response, orthogonal periodic sequences, Volterra filters, nonlinear systems.**

I. Introduction

ONE of the most common operations in acoustics and audio processing is the measurement of the room impulse response (RIR). The measurement is used to estimate the characteristics of the acoustic environment [1], [2] and to compensate, improve, or exploit these characteristics [3]–[17]. Many different approaches have been proposed over the years. The first methodologies used for RIR measurement are based on impulsive signals [18], directly estimating the impulse response of a linear system. Periodic pulses [19], [20] were proposed to contrast the effect of noise. These approaches present a limitation in the amplitude of the pulses. When the pulses are electronically generated, a too high pulse could damage the loudspeaker or could activate the protection circuit of the amplifier, invalidating the measurement. To overcome this limitation, time-stretched pulses [21], [22], consisting of an expanded pulse excitation, were proposed. An approach that had great success and is still popular today uses as input signal a maximal length sequence (MLS) [23], which is a periodic binary pseudo-noise sequence of period $2^k - 1$ with $k \in \mathbb{N}$ [24]. Since the MLSs have an almost perfect autocorrelation function, it is possible to estimate the room response with the cross-correlation method, computing the cross-correlation between the output sequence and the input MLS. Also in noiseless conditions, there is an estimation error due to the not perfect autocorrelation, but this error is usually negligible and tends to 0 increasing the period of the MLS. The perfect periodic sequences (PPSs) for linear filters [25]–[27] overcome this error by guaranteeing a perfect autocorrelation function, a train of unit pulses. In noiseless conditions, these PPSs allow the perfect identification of the impulse response of any linear system using the cross-correlation method. The main limit of MLSs and also of PPSs for linear systems is the sensitivity to the nonlinearities of the measurement system. To contrast the effect of noise, it is common to increase the reproduction volume beyond the limit of linearity of the power amplifier or the loudspeaker and nonlinear effects arise in these components of the measurement system. These nonlinear effects are particularly evident in the identification with MLSs, where they cause the appearance of artifacts in the form of impulsive noises [28]. These pulses are in reality replicas of the same impulse response at different delays and with different amplitudes. The replicas are originated by the fact that the product of an MLS $b(n)$ with the same delayed MLS $b(n-i)$ is again the same MLS delayed by another quantity $b(n-j)$ [29].

The problems of the MLS approach highlighted the need for room response measurement methods robust not only towards the noise but also towards the nonlinearities that inevitably arise in the measurement system at high reproduction volumes. Sweeps signals have been largely employed for RIR estimation due to their robustness towards noise and nonlinearities. The first sweeps used for impulse response measurement were the linear sweeps of the time delay spectrometry technique [30], [31]. Nowadays, the most popular RIR measurement technique is based on exponential sweeps, independently developed by different researchers [32], [33]. The approach has been improved in [34], where a synchronized exponential sweep technique is presented. The popularity of the exponential sweeps derives from their robustness towards the nonlinear-

A. Carini is with the Department of Engineering and Architecture, University of Trieste, Trieste, Italy e-mail: {acarini@units.it}.

S. Cecchi, A. Terenzi, and S. Orcioni are with the Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy. E-mail: {s.cecchi, a.terenzi, s.orcioni}@staff.univpm.it.

Parts of this work have been published in A. Carini, S. Orcioni, and S. Cecchi, "On Room Impulse Response Measurement Using Orthogonal Periodic Sequences," in EUSIPCO 2019 - 27th European Signal Processing Conference (EUSIPCO), 2019, pp. 5486–5490.

This work was supported in part by a Department of Engineering and Architecture Research Grant.

ities of the measurement system. In the paper introducing the exponential sweep technique [32], it was shown that if the measurement system is an Hammerstein system, i.e., is a memoryless nonlinearity with in cascade a linear system [35], the artifacts originated by the nonlinear terms can be segregated at negative times and windowed out. In reality, it was shown recently in [36] that the measures with exponential sweeps performed according to [32]–[34] are still affected by the nonlinear kernels of the Hammerstein model, unless the measurement is corrected accounting for the higher-order kernels. As in other papers (e.g., [37]–[40]) the approach of [36] uses the exponential sweeps to identify Hammerstein models for nonlinear system emulation but the results are applicable also to room response estimation. We must point out that almost all measurement systems based on exponential sweeps do not implement the correction of [36]. Moreover, the correction is effective only for memoryless nonlinearities, which rarely occur. Many papers [41], [42] have shown that nonlinear distortions with memory can affect the RIR measurement based on exponential sweep technique. The same weakness appears also in another technique we have to mention: the Perfect Sweeps introduced in [43], which are periodic sequences with the sinusoidal characteristics of sweeps and with a perfect autocorrelation function.

Since the robust measurement of the RIR in presence of nonlinearities with memory is an open and challenging problem, the authors and colleagues started researching possible solutions. A first approach was proposed in [44]–[47]. The approach is based on modeling the entire measurement system with an orthogonal nonlinear filter, a Legendre nonlinear (LN) filter in [44], [45] or a Wiener nonlinear (WN) filter in [46], [47]. The first-order kernel of this system is then identified using a PPS for nonlinear filters using again the cross-correlation method. In the context of nonlinear filters, PPSs are periodic sequences that guarantee the perfect orthogonality of the basis functions over a sequence period. PPSs for LN and WN filters differ in their input sample distribution, which is uniform in LN filters and Gaussian in WN filters. The first-order kernel of the measurement system is given by the convolution between the RIR and the first-order kernel of the power amplifier and loudspeaker system. This is already a good estimate of the RIR and can be further improved by compensating the power amplifier and loudspeaker system inverting their frequency response with the Kirkeby algorithm, as done in [48]. It should be noted that the first-order kernel of LN or WN filters does not coincide with the impulse response for a small amplitude of the measurement system. Changing the amplitude of the input signal will change the measured first-order kernel.

A novel approach for nonlinear system identification was proposed in [49], [50], where the concept of orthogonal periodic sequence (OPS) was introduced. OPS can be used for the identification of any Functional Link Polynomial filter (FLiP), which is a broad class of nonlinear filters comprising also LN, WN, and Volterra filters [51]. Given a periodic input signal, by definition, an OPS is a sequence that cross-correlated with the output of the system provides one of the diagonals of the nonlinear model. The input signal can be any persistently exciting sequence and can also be a quantized sequence.

The OPSs are interesting candidates also for coping with the problem of the robust RIR measurement in presence of nonlinearities with memory. RIR measurement using OPSs was first proposed by the authors in the conference paper of [52]. In the present paper, we fully detail and study the RIR measurement based on OPSs, modeling the entire measurement system as a Volterra filter. The approach measures the first-order kernel of the Volterra filter using the cross-correlation method, i.e., computing the cross-correlation between the system output and an appropriate OPS. It should be noted that the first-order kernel of the Volterra model is also the impulse response for small signals of the model. The proposed approach is the first that allows to directly and robustly estimate the impulse response for small signals of the measurement system with the cross-correlation method. The input signal can be generated as any persistently exciting periodic sequence with samples that can have uniform, Gaussian, or pink distribution [53].

The main original contributions of this paper are the following: i) Full detailed presentation and study of a novel RIR measurement approach robust towards nonlinearities based on OPS. ii) Detailed description of how OPS suitable for RIR measurement can be obtained. iii) Comprehensive analysis of the harmful effects that can corrupt RIR measurement, i.e., of an underestimation of the memory or the nonlinearity of the measurement system, including an original study of the effect of Gaussian and non-Gaussian noises. iv) Experimental results that include emulated measurements and real measurements taken in a room, comparing the proposed approach with competing approaches (MLSs, exponential sweeps, and PPSs for LN and WN filters).

It has to be pointed out that the approaches modeling the measurement with a Volterra, LN, or WN filter, as well as the proposed approach, are ineffective against the sub-harmonic distortions that sometimes may affect loudspeakers [54]–[57]. It was shown in [58] that polynomial filters cannot describe the sub-harmonic distortions.

There are many related papers, mainly about nonlinear system identification, that for space limitations could not be mentioned in the review of the state of the art of this Introduction. For some recent contributions about the identification or control of nonlinear systems the interested reader is referred to [59]–[65].

The rest of the paper is organized as follows: Section II reviews Volterra filters and provides a Volterra representation of the measurement system. Section III discusses OPSs and how OPSs, suitable for the first-order kernel measurement, can be derived. Section IV studies RIR measurement using OPSs and all detrimental effects that can influence this measurement. Section V provides experimental results that compare the proposed approach with competing approaches. Section VI reports concluding remarks.

The following notation is used throughout the paper: $\mathbb{N}$ is the set of natural numbers and $\mathbb{R}$ that of real numbers, $\mathbb{R}_1$ is the interval $[-1, +1]$, $< a(n) >_L$ is the sum of $a(n)$ over a period of L consecutive samples, $E[\cdot]$ indicates expectation and $*$ convolution, $\delta(n)$ is the unit pulse sequence, and $\lceil \cdot \rceil$ is the ceil operator.

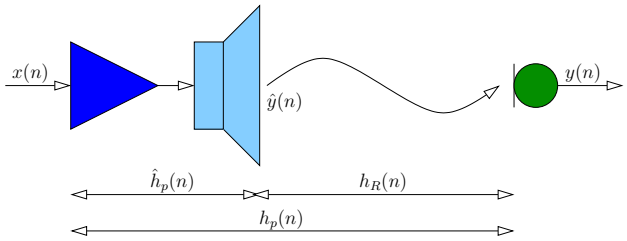

Fig. 1. A typical RIR measurement system.

II. A Volterra model of the measurement system

A. The measurement system

Figure 1 represents a typical scheme of the RIR measurement system. The system is composed of a power amplifier, a loudspeaker, the room acoustic path, and a microphone. At the sound pressure levels normally used for the measurement, the acoustic path is a linear system[1] and the objective of the measurement is to estimate its impulse response $h_{\mathrm{R}}(n)$. This estimate is performed by observing the room response indirectly via the measurement system and, thus, is affected by the characteristics of the power amplifier, the loudspeaker, and the microphone. Very often at the high reproduction volumes used in the measurement to contrast the effect of noise, the power amplifier and the loudspeaker behave as mildly nonlinear systems, making the entire measurement chain nonlinear. Due to the low levels of the acquired signals, the microphone can still be considered as a linear system and in what follows its response will be neglected or, better, incorporated in the amplifier and loudspeaker response.

To account for their nonlinearities, the power amplifier and loudspeaker cascade will be modeled as a Volterra filter.

B. Volterra filters

Volterra filters are polynomial filters that derive from the double truncation, with respect to the order and memory of the Volterra series [66]. According to the Stone-Weierstrass theorem [67], they can arbitrarily well approximate any discrete time invariant, finite memory N, continuous nonlinear system,

$$y(n) = f[x(n), x(n-1), ..., x(n-N+1)], \tag{1}$$

with $x(n)$ the input signal and $x(n) \in [-1, +1]$, $y(n)$ the output signal, and $f[\cdot, ..., \cdot]$ a continuous function from $\mathbb{R}^N$ to $\mathbb{R}$.

In triangular form, a discrete time Volterra filter of order K and memory N has the following input-output relationship

$$y(n) = h_0 + \sum_{r=1}^{K} \sum_{n_1=0}^{N-1} \sum_{n_2=n_1}^{N-1} \cdots \sum_{n_r=n_{r-1}}^{N-1} h_{r,n_1,...,n_r} \cdot \cdot x(n-n_1)x(n-n_2)...x(n-n_r). \tag{2}$$

It is useful to interpret the Volterra filters as a linear combination of basis functions

$$x(n-n_1)x(n-n_2)...x(n-n_r).$$

[1]Nonlinearities can be observed in the acoustic paths only for sound pressure levels larger than the pain threshold.

Each of these basis function is a product of delayed input samples and, without loss of generality, we can assume $n_1 \leq n_2 \leq ... \leq n_r$ and $r \in \mathbb{N}$. By definition, r is the *order* of the basis function, and the maximum time difference between the involved input samples, $n_r - n_1$, is its *diagonal number*. For $r = 0$, the basis function reduces to the constant 1.

By considering a change of variables, the input-output relationship of the Volterra filter can also be expressed in *diagonal form* [68], i.e.,

$$y(n) = h_0 + \sum_{r=1}^{K} \sum_{s_1=0}^{D} \sum_{s_2=s_1}^{D} \cdots \sum_{s_{r-1}=s_{r-2}}^{D} \sum_{m=0}^{N-1-s_{r-1}} h_{m,m+s_1,...,m+s_{r-1}} x(n-m)x(n-m-s_1)...x(n-m-s_{r-1}), \tag{3}$$

where D is the maximum diagonal number, and $D = N - 1$ for the Volterra filter in (2). Since natural systems typically have the most relevant coefficients at low diagonal numbers, the maximum diagonal number is conveniently limited considering $D \ll N - 1$.

Equation (3) highlights the filter-bank nature of Volterra filters: for each tupla $(s_1, s_2, ..., s_{r-1})$, the zero-lag basis function $f_p(n) = x(n)x(n-s_1)...x(n-s_{r-1})$ is processed by a linear filter having impulse response

$$h_p(n) = \left\{ h_{m,m+s_1,...,m+s_{r-1}} \text{ for } 0 \leq m \leq N - s_{r-1} - 1 \right\}$$

of length $N_p = N - s_{r-1}$. The $h_p(n)$ are the *diagonals* of the Volterra filter. For compactness, in what follows we will rewrite (3) in the following equivalent forms

$$y(n) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p-1} h_p(m) f_p(n-m), \tag{4}$$

$$y(n) = \sum_{p=0}^{R-1} h_p(n) * f_p(n), \tag{5}$$

where the zero lag basis functions $f_p(n)$ are

$$\begin{aligned} f_0(n) &= 1, \\ f_1(n) &= x(n), \\ f_2(n) &= x^2(n), \\ f_3(n) &= x(n)x(n-1), \\ &\vdots \\ f_{D+2} &= x(n)x(n-D) \\ f_{D+3} &= x^3(n), ...; \end{aligned}$$

$R = \binom{D+K}{D+1} + 1$ is the number of zero lag basis functions; N_p is the memory length of the basis function $f_p(n)$ with $N_p = N - D_p$ where D_p is the diagonal number of $f_p(n)$.

C. Volterra representation of the measurement system

Let us assume that the system composed of the power amplifier and the loudspeaker can be modeled as a Volterra filter of order K, memory $\hat{N}$, and diagonal number D and that the acoustic path has impulse response $h_{\mathrm{R}}(n)$ of length M. In this condition, the entire measurement system of Figure 1 can

also be modeled as a Volterra filter with the same order K and diagonal number D, but with memory length $N = \hat{N}+M-1$. This can be easily proved by modeling the power amplifier and loudspeaker system following (5) as

$$\hat{y}(n) = \sum_{p=0}^{R-1} \hat{h}_p(n) * f_p(n), \tag{6}$$

with $\hat{h}_p(n)$, the p-th Volterra diagonal of the power amplifier and loudspeaker system, having length $\hat{N}_p = \hat{N} - D_p$. The measurement system has the following input-output relationship

$$\begin{aligned} y(n) &= h_{\mathrm{R}}(n) * \hat{y}(n) \\ &= h_{\mathrm{R}}(n) * \sum_{p=0}^{R-1} \hat{h}_p(n) * f_p(n) \\ &= \sum_{p=0}^{R-1} h_{\mathrm{R}}(n) * \hat{h}_p(n) * f_p(n) \\ &= \sum_{p=0}^{R-1} h_p(n) * f_p(n), \end{aligned} \tag{7}$$

where $h_p(n) = h_{\mathrm{R}}(n) * \hat{h}_p(n)$ is the p-th diagonal of the entire measurement system. The system is composed of the same zero lag basis functions, and thus has the same order and diagonal number. Each basis function is convolved with $h_p(n)$ of length $\hat{N}_p + M - 1$. In particular, the linear basis $f_1(n) = x(n)$ is convolved with $h_1(n) = h_{\mathrm{R}}(n) * \hat{h}_1(n)$, which is the first-order kernel of the measurement system and has length $\hat{N} + M - 1 = N$.

We will measure $h_1(n)$ using an appropriate OPS. $h_1(n)$ approximates the RIR $h_{\mathrm{R}}(n)$ but the measure is affected by convolution with $\hat{h}_1(n)$, the first-order kernel of the amplifier and loudspeaker system. $\hat{h}_1(n)$ can be measured and characterized in an anechoic chamber, maybe using the same input-OPS pair used for measuring $h_1(n)$. As was proposed for linear systems [48], $h_{\mathrm{R}}(n)$ can be obtained by equalizing $h_1(n)$ with the inverse response of $\hat{h}_1(n)$, exploiting the Kirkeby algorithm, as follows [69]:

$$h_{\mathrm{R}}(n) = \mathrm{IFFT}\left[\frac{\mathrm{FFT}[h_1(n)] \cdot \mathrm{FFT}[\hat{h}_1(n)]^*}{\mathrm{FFT}[\hat{h}_1(n)] \cdot \mathrm{FFT}[\hat{h}_1(n)]^* + \epsilon(\omega)}\right], \tag{8}$$

where $\mathrm{FFT}[\cdot]$ and $\mathrm{IFFT}[\cdot]$ are direct and inverse FFT operators, respectively, $\epsilon(\omega)$ is a frequency-dependent regularization parameter, and all operations are performed per single frequency bin. In reality, the Kirkeby algorithm estimates the RIR $h_{\mathrm{R}}(n)$ only when the loudspeaker has a uniform response at all directions in the frequency range of interest. This is never the case. Therefore, researchers and technicians performing the measure are usually satisfied with the knowledge that amplifier and loudspeaker affects the measurement in a known, mild manner, and $h_{\mathrm{R}}(n)$ is usually approximated with $h_1(n)$.

III. Orthogonal Periodic Sequences

We will present now a methodology for obtaining the first-order kernel of a Volterra filter of order K, diagonal number D, and memory N based on the cross-correlation method. By definition, an OPS is a periodic sequence that cross-correlated with the filter output provides one of the diagonals $h_p(m)$ of the Volterra filter in (4). In RIR measurement the interest is focused on the estimation of $h_1(m)$ for all $m \in [0, N-1]$. Let us consider a periodic input sequence $x(n)$ of a sufficiently large period L, better specified later in this Section. To guarantee the invertibility of the data matrices introduced in the following, the input sequence is assumed to persistently excite the Volterra filter. The condition is easily satisfied and is guaranteed if the input samples have a Gaussian distribution, a white uniform distribution, a pink distribution, or any other sufficiently rich random distribution. The input sequence could also be quantized, provided that a sufficiently large number of levels are used. In fact, an independent, identically distributed sequence must have at least $K+1$ distinct values to persistently excite an order K Volterra filter [70].

We want to develop the OPS $z(n)$ of period L such that

$$h_1(j) =< y(n)z(n-j) >_L, \tag{9}$$

for $0 \le j \le N-1$ and $<.>_L$ is the sum of argument over a period of L consecutive samples. Inserting (4) in (9), we have

$$h_1(j) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p-1} h_p(m) < f_p(n-m)z(n-j) >_L . \tag{10}$$

For $j = 0$, to meet (9) it must be

$$< f_0(n)z(n) >_L =< z(n) >_L = 0, \tag{11}$$

$$< f_1(n)z(n) >_L =< x(n)z(n) >_L = 1, \tag{12}$$

$$< f_1(n-m_1)z(n) >_L =< x(n-m_1)z(n) >_L = 0, \tag{13}$$

$$< f_p(n-m_p)z(n) >_L = 0, \tag{14}$$

for all $1 < m_1 \le N-1$, $0 \le m_p \le N_p-1$ and $1 < p \le R-1$.

For $j > 0$, to meet (9) together with (11)–(14) it must also be

$$< f_p(n)z(n-j) >_L =< f_p(n+j)z(n) >_L = 0. \tag{15}$$

for all $0 < j \le N-1$ and for all $0 < p \le R-1$.

Thus, the OPS $z(n)$ must satisfy the linear equation system

$$< z(n) >_L = 0, \tag{16}$$

$$< x(n)z(n) >_L = 1, \tag{17}$$

$$< x(n-m_1)z(n) >_L = 0, \tag{18}$$

$$< f_p(n-m_p)z(n) >_L = 0, \tag{19}$$

for all $-(N-1) < m_1 \le N-1$ and $m_1 \neq 0$, $-(N-1) \le m_p \le N_p - 1$ and $1 < p \le R-1$. The equation system has Q equations and L variables ($z(n)$ for $n \in [0, L-1]$), with

$$Q = N_D + (R-1)(N-1). \tag{20}$$

The equation system is critically determined for $L = Q$ and under-determined for $L > Q$. If the periodic input is persistently exciting, the system always admits a solution. In matrix form, it can be written as follows,

$$\mathbf{Sz} = \mathbf{d} \tag{21}$$

where $\mathbf{z}$ is a vector collecting the samples of $z(n)$, $\mathbf{d}$ is the vector $[0, 1, 0, ..., 0]^T$ with just one nonzero element, and $\mathbf{S}$ is

a square or fat matrix. Each row of $\mathbf{S}$ is formed by the samples of a basis function $f_p(n-m_p)$, with n ranging along the row from 0 to $L-1$, and m_p and p changing along the columns with $-(N-1) \leq m_p \leq N_p - 1$ and $0 \leq p \leq R-1$. The first row, corresponding to $p=0$, is formed by all ones. The other rows come in groups of N_p+N-1 rows formed by the same rotated elements, for the periodicity of $x(n)$. The system has minimum norm solution given by

$$\mathbf{z} = \mathbf{S}(\mathbf{S}\mathbf{S}^T)^{-1}\mathbf{d}. \tag{22}$$

The matrix $\mathbf{S}\mathbf{S}^T$ has a peculiar form. Its elements are cross-correlations between basis functions with different time delays, i.e., are

$$< f_{p_1}(n-m_{p_1}) f_{p_2}(n-m_{p_2}) >_L \tag{23}$$

where $0 \leq p_1, p_2 \leq R-1$, $-(N-1) \leq m_{p_1} \leq N_{p_1}-1$, and $-(N-1) \leq m_{p_2} \leq N_{p_2}-1$. $\mathbf{S}\mathbf{S}^T$ is a block matrix whose entries are Toeplitz matrices with different sizes. For each couple of basis functions f_{p_1} and f_{p_2} with $1 \leq p_1, p_2 \leq R-1$, the corresponding block matrix has size $(N_{p_1}+N-1)\times(N_{p_2}+N-1)$. Efficient algorithms exist for the inversion of the matrix $\mathbf{S}\mathbf{S}^T$ or for computing the product $(\mathbf{S}\mathbf{S}^T)^{-1}\mathbf{d}$ [71]. Working with the polynomial basis functions of Volterra filters, $\mathbf{S}\mathbf{S}^T$ could have a bad conditioning. In all our experiments, working with double precision arithmetic, for a sufficiently large L, we have always been able to find a solution with good accuracy.

In summary, to develop an OPS suitable for the identification of the first-order kernel of a Volterra filter, we must choose the filter parameter N, K, D, a periodic input sequence with period equal to or grater than Q given in (20), and then compute (22) with some efficient algorithm as that of [71].

IV. RIR MEASUREMENT USING OPSs

Measuring the RIR using an OPS involves first deciding the memory length of the room response N, the order of nonlinearity K, and the diagonal number D of the measurement system. An input-OPS pair suitable for the identification of the first-order kernel of a Volterra model of memory N, order K, and diagonal number D must be chosen. The input-OPS pair can be precomputed, and many of these pairs have been prepared and made available in [53]. The periodic input sequence has to be applied for at least a period L plus $N-1$ samples, to guarantee a stationary response over a period L. The RIR is then measured from the cross-correlation between the microphone signal and the OPS.

Let us assume that the measurement system can be represented as a Volterra filter of memory length N_{Sys}, order K_{Sys}, and diagonal number D_{Sys}. According to (4), its input-output relationship can be written as

$$y(n) = \sum_{p=1}^{R_{\mathrm{Sys}}-1} \sum_{m=0}^{N_{\mathrm{Sys}}-D_p-1} \tilde{h}_p(m) f_p(n-m) + \overline{\nu}(n), \tag{24}$$

where we have neglected the constant term $\tilde{h}_0$, $R_{\mathrm{Sys}} = \binom{D_{\mathrm{Sys}}+K_{\mathrm{Sys}}}{D_{\mathrm{Sys}}+1} + 1$ is the number of zero lag basis functions, and $\overline{\nu}(n)$ is the measurement noise.
In noise absence, i.e., when $\overline{\nu}(n) = 0$, and if $N_{\mathrm{Sys}} \leq N$, $K_{\mathrm{Sys}} \leq K$, and $D_{\mathrm{Sys}} \leq D$, the theory of Section III guarantees a perfect estimation of the measurement system first-order kernel.

In this Section we analyze the effect of an underestimation of the memory length, order, or diagonal number of the measurement system, and study the effect of noise. Moreover, the computational complexity of the RIR measurement with the proposed approach is also discussed.

A. Memory underestimation

We separately analyze the effect of memory underestimation and of the nonlinearity underestimation. The case of an underestimation of the memory of the measurement system is first considered. Let us assume $N_{\mathrm{Sys}} = N + \Delta_N$, with $\Delta_N > 0$, $K_{\mathrm{Sys}} \leq K$, $D_{\mathrm{Sys}} \leq D$, and let us neglect the noise effect $\overline{\nu}(n) = 0$. In this case, $R_{\mathrm{Sys}} = R$ and the system in (24) can be written as

$$y(n) = \sum_{p=0}^{R-1} \sum_{m=0}^{N+\Delta_N-D_p-1} \tilde{h}_p(m) f_p(n-m). \tag{25}$$

The coefficients $\tilde{h}_p(n)$ are estimated with (9) using an OPS for a Volterra filter of memory N, order K, diagonal number D. The memory underestimation causes an aliasing error, which influences only the first Δ_N terms of $h_1(n)$. This can be understood considering that, for $j \in [0, \Delta_N - 1]$, the estimation of $\tilde{h}_1(j)$ with $< y(n)z(n-j) >_L$ is affected by the cross-correlations $< x(n-k)z(n-j) >_L$ with $k \in [N+j, N+\Delta_N-1]$. In contrast, for $j \in [\Delta_N, N-1]$ and $k \in [N, N+\Delta_N-1]$, the cross-correlations

$$< x(n-k)z(n-j) >_L = < x(n-k+j)z(n) >_L = 0$$

for the orthogonality conditions imposed by (16)–(19).

B. Underestimation of the nonlinearity

We turn our attention to the case of an underestimation of the order or diagonal number of the Volterra filter. Let assume $K_{\mathrm{Sys}} > K$ or $D_{\mathrm{Sys}} > D$, $N_{\mathrm{Sys}} \leq N$, and $\overline{\nu}(n) = 0$. The system in (24) can be written as

$$y(n) = \sum_{p=0}^{R-1} \sum_{m=0}^{N-D_p-1} \tilde{h}_p(m) f_p(n-m) + \Delta_{K,D}(n), \tag{26}$$

where $\Delta_{K,D}(n)$ is a term formed by the linear combination of all basis functions of order greater than K and diagonal number greater than D. The coefficients $\tilde{h}_p(n)$ are again estimated with (9). The underestimation of the nonlinearity causes an aliasing error which affects all coefficients of RIR. In fact, the estimation of $\tilde{h}_1(j)$ with (9) is affected by an error equal to $< \Delta_{K,D}(n) z(n-j) >_L$. The error is deterministic. But, if we assume to work with a large period L and a large number of basis functions in $\Delta_{K,D}(n)$, for the law of large numbers, the error can be assumed stochastic and Gaussian distributed and its effect can be deemed similar to a measurement noise. The larger the order K or the diagonal number D, the smaller is the error $< \Delta_{K,D}(n) z(n-j) >_L$. Thus, we can increase the protection against nonlinearities

of the RIR measurement method increasing the order K or diagonal number D. The improved protection is paid with an increase of the period L of the input sequence. Due to the "curse of dimensionality", i.e., the exponential increase in the number of coefficients with the order K, only small values of K can be used in practice.

C. The measurement noise effect

We now study the effect of a measurement noise $\overline{\nu}(n)$ on the room response identification. Let us assume $N_{\text{Sys}} \leq N$, $K_{\text{Sys}} \leq K$, and $D_{\text{Sys}} \leq D$, while $\overline{\nu}(n) \neq 0$. The measurement system has input-output relationship

$$y(n) = \sum_{p=1}^{R-1} \sum_{m=0}^{N_p-1} \tilde{h}_p(m) f_p(n-m) + \overline{\nu}(n). \tag{27}$$

We consider separately the case of a Gaussian and non-Gaussian noise.

1) Gaussian noise: We assume $\overline{\nu}(n)$ to be a colored Gaussian noise,

$$\overline{\nu}(n) = h_\nu(n) * \nu(n), \tag{28}$$

where $\nu(n)$ is a zero-mean, variance σ_ν^2, white Gaussian noise. $h_\nu(n)$ is the causal, finite memory forming filter with memory N_ν. Without loss of generality we assume

$$\sum_{n=0}^{N_\nu-1} h_\nu(n)^2 = 1. \tag{29}$$

Note that for $h_\nu(n) = \delta(n)$ the measurement noise is zero-mean white Gaussian with variance σ_ν^2.

Applying (9), for the properties of OPSs we have

$$h_1(j) = \tilde{h}_1(j) + < (h_\nu(n) * \nu(n)) z(n-j) >_L, \tag{30}$$

which shows that the identification in this conditions is unbiased, since $\nu(n)$ is zero mean and $E[h_i(j)] = \tilde{h}_i(j)$.

It is interesting to estimate the mean square deviation (MSD) of the j-th coefficient of the impulse response, which is defined as

$$\text{MSD}_j = E[(h_1(j) - \tilde{h}_1(j))^2], \tag{31}$$

and according to (9) is

$$\text{MSD}_j = E[(< (h_\nu(n) * \nu(n)) z(n-j) >_L)^2]. \tag{32}$$

It was shown in the appendix of [50] that

$$\text{MSD}_j = \sigma_\nu^2 \sum_{m=-N_\nu+1}^{L-1} < h_\nu(n-m) z(n-j) >_L^2 . \tag{33}$$

When the measurement noise is white Gaussian ($h_\nu(n) = \delta(n)$), the mean square deviation simplifies to

$$\text{MSD}_j = \sigma_\nu^2 < z(n)^2 >_L, \tag{34}$$

and is independent of the delay j.

As can be expected, equations (33) and (34) show that MSD_j is proportional to the noise power σ_ν^2. Moreover, MSD_j is proportional to $< z(n)^2 >_L$ and, thus, inversely proportional to $< x^2(n) >_L$, i.e., the energy of the input sequence over a period, as it can be seen from (17). For a constant power of $x(n)$, $< x^2(n) >_L$ is proportional to L. Consequently, MSD_j is also inversely proportional to the period L. Similar to all other cross-correlation methods, the accuracy of the estimation can be improved by increasing the period L or by computing the cross-correlation over multiple periods.

To compare different OPSs on equal terms, the noise gain G_ν was introduced in [49]. For the measurement of the first-order kernel of a Volterra filter, the noise gain is

$$G_\nu = \frac{\text{MSD}_j}{\sigma_\nu^2} < x^2(n) >_L, \tag{35}$$

and evaluated for a white measurement noise is equal to

$$G_\nu = < z^2(n) >_L \cdot < x^2(n) >_L . \tag{36}$$

G_ν depends on the distribution of the input samples and on the period L. For a specific input sample distribution, G_ν can vary greatly with L, which influences the power of the OPS $z(n)$. In this regard, the minimum period of the OPS $L = Q$ has been found to be a bad choice. For $L = Q$, G_ν assumes very large values that often make the identification using OPSs useless. In contrast, for $L \gg Q$ it is possible to obtain reasonable values of G_ν, which take to a robust estimate of the RIR.

2) Non Gaussian noise: It is also interesting to study the RIR measurement in presence of a non-Gaussian measurement noise. Applying (9) and taking into account the properties of OPSs,

$$h_1(j) = \tilde{h}_1(j) + < \overline{\nu}(n) z(n-j) >_L . \tag{37}$$

If $\overline{\nu}(n)$ is zero mean, also in this case $E[h_i(j)] = \tilde{h}_i(j)$ and the measurement method is unbiased. Moreover, for the central limit theorem the error $h_i(j) - \tilde{h}_i(j)$ has a Gaussian distribution, with

$$\text{MSD}_j = E[(< \overline{\nu}(n) z(n-j) >_L)^2]. \tag{38}$$

MSD_j depends on the auto-correlation function of the noise and on the OPS $z(n)$. The average over a large number of samples L provides also a certain protection towards outliers, even if the identification does not resort to any algorithmic protection as in [72]–[75] to guarantee robustness towards outliers.

D. Computational cost

The proposed approach can be implemented in time domain or in frequency domain. In time domain, the computational cost of the cross-correlation in (9) is of LN operations, i.e., multiplications and additions. If the cross-correlation is computed in the frequency domain, it is of $(\log_2(L)+1)L$ complex operations, assuming a FFT costs $L \log_2(L)$ operations and that the FFT of $z(n)$ is precomputed. These computational costs are identical to those of RIR measurement using MLSs.

V. Experimental results

Two sets of experiments are presented. The first set considers an emulated scenario, where the input signals are applied to a nonlinear pre-amplifier and the recorded outputs are

convolved with a known RIR. The second set of experiments considers the identification of a real RIR. The measurements performed with the proposed approach based on OPSs with input samples having Gaussian, uniform, and pink distribution, are compared with measurements based on PPSs for LN and WN filters, MLSs, and exponential sweeps. In both experiments, we consider the same set of input signals. The OPSs have memory length $N = 8192$ samples, order $K = 3$ (the maximum order allowed by the curse of dimensionality), diagonal number D ranging from 0 to 5, period

$$L = \epsilon 2^{\lceil \log_2(Q) \rceil}, \tag{39}$$

and $\epsilon = 4$. The OPS input samples are quantized in the set $[-512 : +512]/512$. Figure 2 shows the noise gain of the OPSs for $N = 8192$, $K = 3$, $D = 2$, and for different periods L. For $L = Q = 163,820$ the noise gain is unacceptably high, but it reduces to acceptable values increasing L. In our experience the best compromise is obtained when in (39) $\epsilon = 4$ or $\epsilon = 8$, i.e., $L = 2^{20}$ or 2^{21} in this case. In the paper we present the results for $\epsilon = 4$ but comment also the case of $\epsilon = 8$. The PPSs for WN and LN filters have $N = 8192$, $K = 3$, D ranging from 0 to 4, and L ranging from 2^{17} till 2^{21}. The MLSs have period ranging from $2^{17} - 1$ till $2^{21} - 1$. The exponential sweeps have lengths ranging from 2^{17} till 2^{21} and sweep from 20 Hz till 20,000 Hz. They have been constructed as in [34] and the beginning and end of the exponential sweeps have been multiplied by a half Hamming window of length $L/100$ to avoid fringe effects. The sampling frequency is 44 100 Hz. All input sequences have the same power, which means that they have different peak amplitude. Table I shows the ratio between the peak amplitude of different sequences and the amplitude of the MLSs for the same signal power. Working with negligible or mild nonlinearities, audio engineers and technicians typically choose the power of the input signal in order to contrast the noise. This suggests to use the same power for all input sequences. It should be observed that working with strong nonlinearities the choice is typically different. The maximum amplitude of the input signal is chosen to limit the maximum nonlinear distortion, which depends on the amplitude of the input signal. In this paper, we prefer to work with the same power for all input sequences in order to stress the robustness of the proposed method. According to Table I, the OPSs with input samples having Gaussian or uniform distributions have peak amplitudes much larger than the other signals and thus will be affected by larger nonlinear distortions. The interested reader is referred to [52] for a comparison of the different measurement methods considering the same peak amplitude for all signals.

A. Experiment 1

In the first experiment, the input signals were applied to a Behringer MIC 500 vacuum tube preamplifier (set in "valve" mode) and the corresponding output, recorded with a Focusrite Scarlett 2i2 audio interface, was convolved with a previously measured RIR of 8192 samples. A white Gaussian noise has also been added to the resulting signals to have a 40 dB signal to noise ratio. The pre-amplifier has a potentiometer,

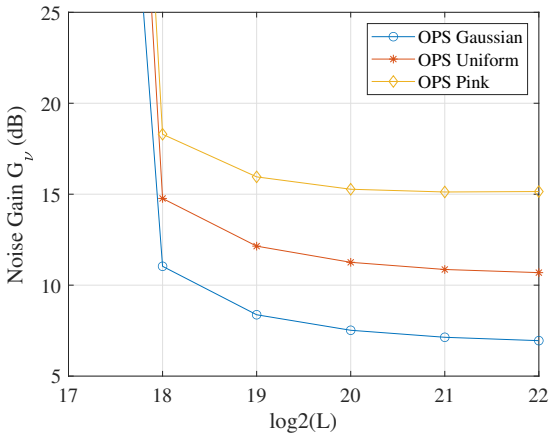

Fig. 2. Noise gain of OPSs with Gaussian, uniform or pink input distribution for $N = 8192$, $K = 3$, $D = 2$ versus period L.

TABLE I
RATIO BETWEEN THE PEAK AMPLITUDE OF DIFFERENT INPUTS AND THE AMPLITUDE OF THE MLSS FOR THE SAME SIGNAL POWER.

Sequence	Ratio
OPS Gaussian	$\sqrt{12}$
OPS Uniform	$\sqrt{3}$
OPS Pink	$\sqrt{24}$
PPS WN	$\sqrt{12}$
PPS LN	$\sqrt{3}$
Exponential Sweep	$\sqrt{2}$
MLS	1

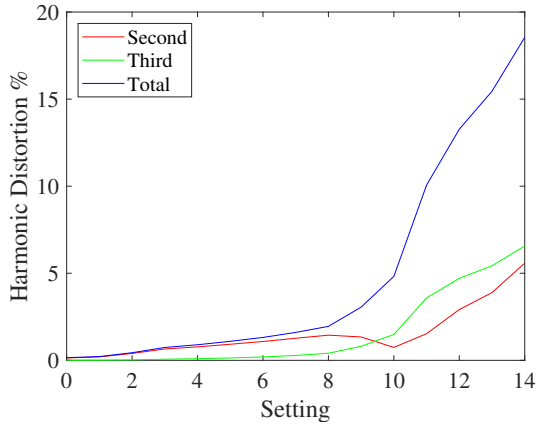

Fig. 3. Experiment 1: Second, third, and total harmonic distortion of the MIC-500 preamplifier at the different settings.

which allows to change the amount of nonlinear distortion. Fifteen different potentiometer settings have been considered and Figure 3 shows the second, third, and total harmonic distortion on a 1 kHz tone at the maximum amplitude of the sequences for the different potentiometer settings (briefly indicated as Setting 0 to 14 in the Figures). The harmonic distortion is the ratio in percent between the power of an harmonic (or all harmonics in case of total distortion) and that of the fundamental frequency. Most of the harmonic distortions of Figure 3 are much greater than those envisaged in the

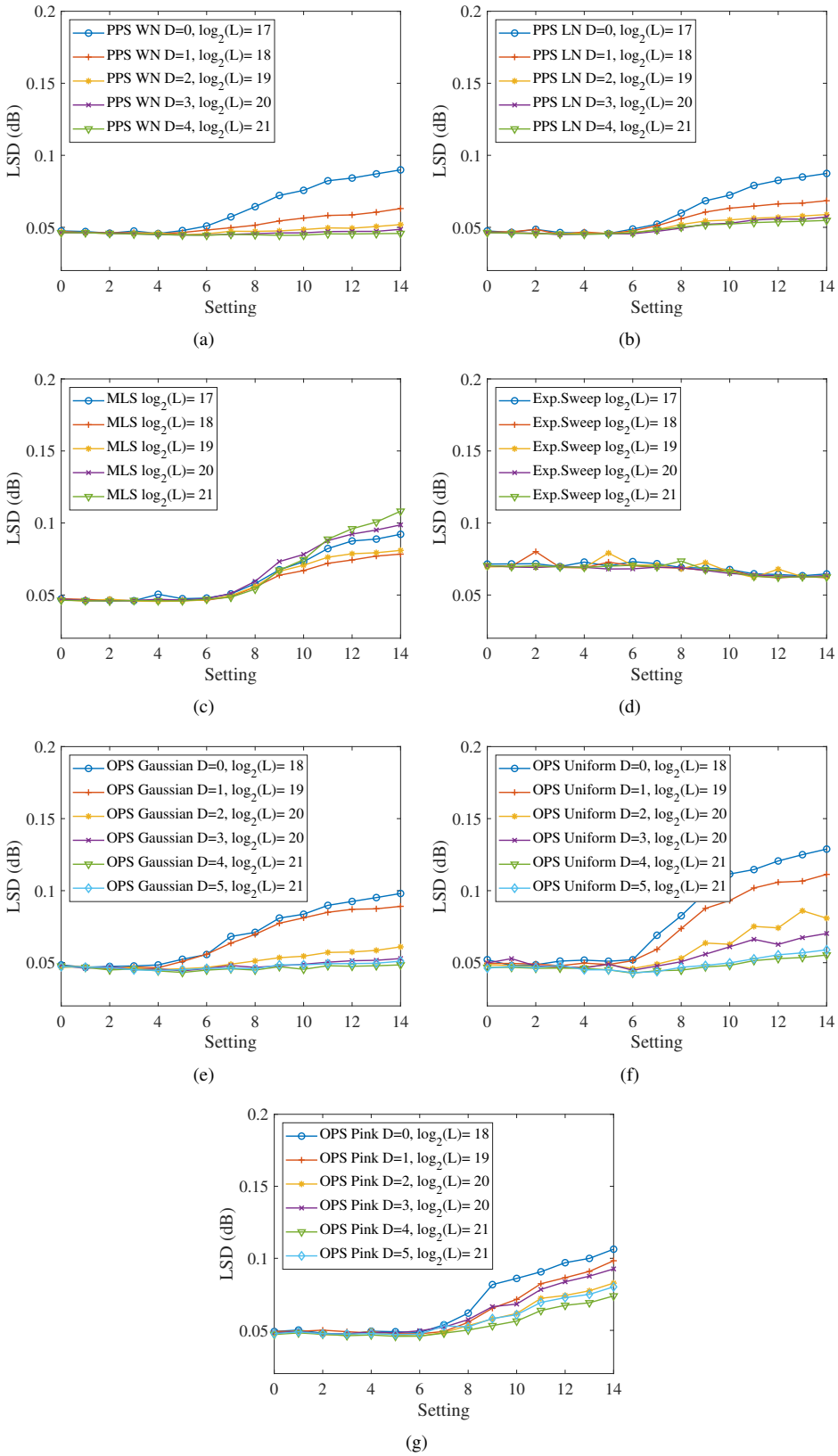

Fig. 4. Experiment 1: Log-spectral distance in the band $[100, 18000]$ Hz at the different settings for (a) PPSs for WN filters, (b) PPSs for LN filters, (c) MLSs, (d) exponential sweeps, (e) OPS with Gaussian distribution, (f) OPS with uniform distribution, (g) OPS with pink distribution, without compensation of the MIC 500.

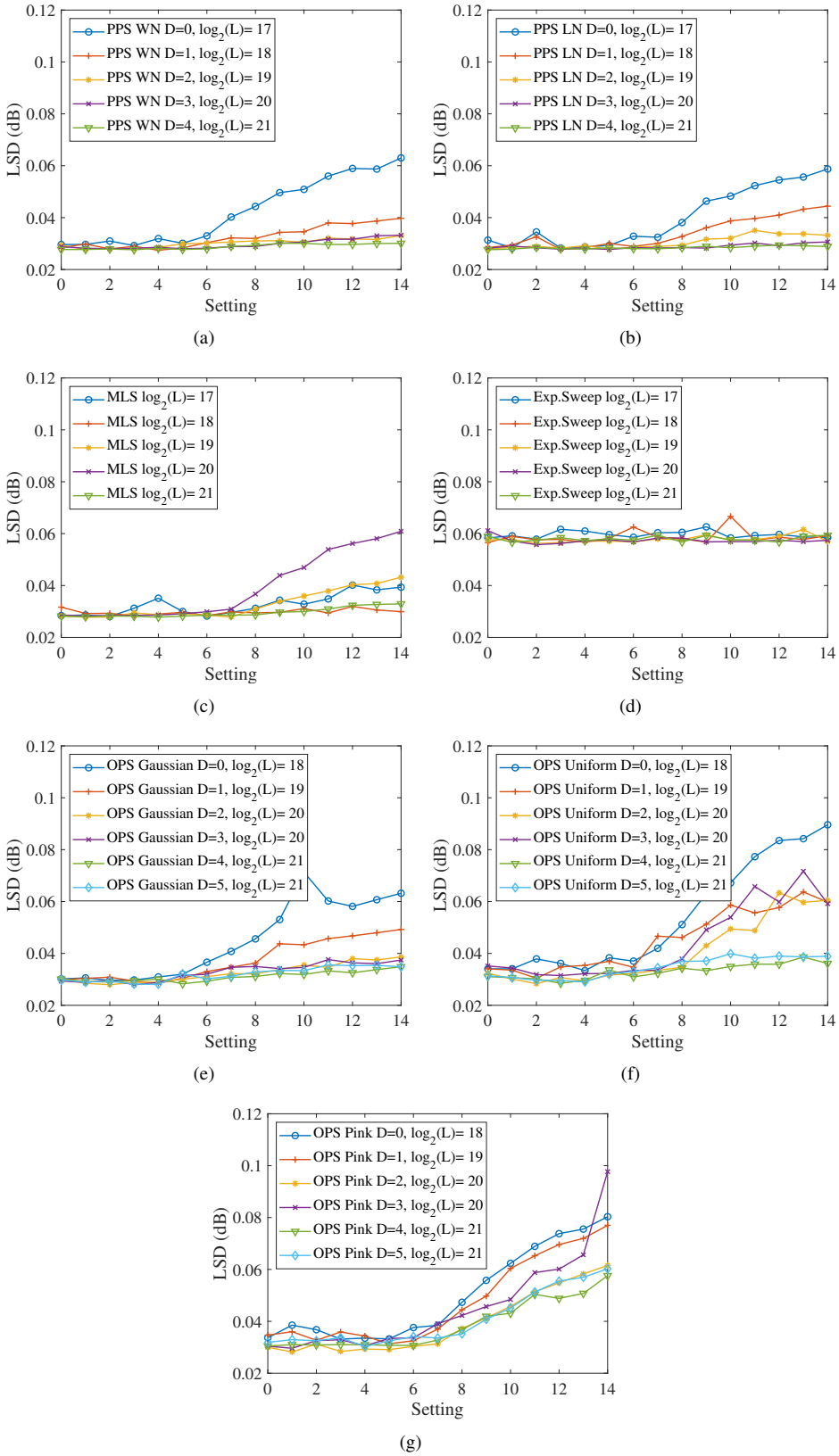

Fig. 5. Experiment 1: Log-spectral distance in the band $[100, 18000]$ Hz at the different settings for (a) PPSs for WN filters, (b) PPSs for LN filters, (c) MLSs, (d) exponential sweeps, (e) OPS with Gaussian distribution, (f) OPS with uniform distribution, (g) OPS with pink distribution, with compensation of the MIC-500,

RIR measurement system. They have been specially selected in such a way to underline the robustness of the proposed approach and the differences between the compared methods. Many parameters could be used to compare the different measurement approaches. In what follows the different methods are compared in terms of log-spectral distance (LSD) [3], [76] between the measured room magnitude response and its actual value. The LSD is a well-known distance measure used in audio processing to compare different magnitude spectra. Its success is motivated by its simplicity and effectiveness in compressing the spectral differences in a single parameter in dB. The LSD is defined in the band $B = [k_1 \frac{F_S}{T}, k_2 \frac{F_S}{T}]$, with k_1 and $k_2 \in \mathbb{N}$, F_S the sampling frequency and T the number of samples of the discrete Fourier transform (DFT), as follows:

$$\text{LSD} = \sqrt{\frac{1}{k_2 - k_1 + 1} \sum_{k=k_1}^{k_2} \left[10 \log_{10} \frac{|H_R(k)|^2}{|\hat{H}_R(k)|^2} \right]^2}, \quad (40)$$

where $|H_R(k)|$ is the actual room magnitude response and $|\hat{H}_R(k)|$ is the measured room magnitude response.

Figure 4 shows the LSD in the band $[100, 18\,000]$ Hz, falling strictly inside the passband, measured with the different methods without compensation of the MIC 500 preamplifier. Figure 5 shows the same LSD plots with the preamplifier compensation using the Kirkeby algorithm in (8). The reader should first notice the improvement of performance that can be obtained in all PPSs and OPSs by increasing the diagonal number D. The protection against nonlinearities can be increased by augmenting the diagonal number or the order of nonlineary, at the cost of a larger period of the sequence. In Figure 4 and 5, the best performance is offered by the PPSs for WN filters (where the input samples have a Gaussian distribution) and the OPSs for Gaussian distribution for diagonal number D at least 3. These methods combine the robustness towards nonlinearity of the considered approaches with a input sample distribution that concentrate most of the power on low amplitude samples and less excites the nonlinearities of the MIC 500. The PPSs for LN systems and the OPSs with input samples having uniform distribution, provide also good results for the larger diagonal numbers ($D = 4, 5$), i.e., for larger protection against nonlinearities. In reality, they tend to excite more the nonlinearities of the system, as is apparent for larger distortion settings, i.e., setting 8 to 14, and D lower than 4. The OPSs with input samples having Pink distribution are the most affected by the nonlinearities. Many of their samples are much larger than those of the other sequences and are thus much more affected by nonlinear distortion. We must point out that, if the same peak amplitude is imposed to all sequences, the OPSs for Pink distribution provide the best performances, with perfornance similar to WN filters and the OPS for Gaussian distribution. The OPSs with samples having Gaussian or uniform distribution give here worse results than the PPSs for WN and LN systems, respectively, because of the worse noise gain. Similar results have been obtained also doubling the period of the OPS, i.e., for $\epsilon = 8$ in (39), with a small improvement visible only on the OPSs with smaller D. The MLSs are well known to suffer the effect of nonlinearities and the results of Figure 4 and 5 confirm this general knowledge. Different MLSs are affected in a different manner by the nonlinearity artifacts and the MLS with period $2^{20} - 1$ here gives worse results than the shorter period MLSs. The exponential sweeps provide robust behavior with a uniform LSD at all settings, albeit larger than with the other methods. In Figure 4 the LSD of the exponential sweeps tends to reduce slightly with the settings since the spectral variations introduced by the MIC 500 preamplifier tend in this case to reduce with the setting. As a matter of fact, with the preamplifier compensation the LSD of the exponential sweep measurement becomes almost constant.

The preamplifier compensation significantly reduces the LSD. The reader should note the different vertical scale of Figures 4 and 5. Nevertheless, the different methods are affected in the same way by the nonlinearites and by the noise with and without compensation of the preamplifier.

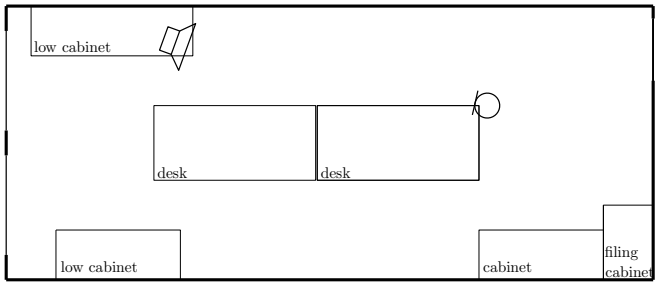

Fig. 6. Experiment 2: Floor plan of the room used in the measurements.

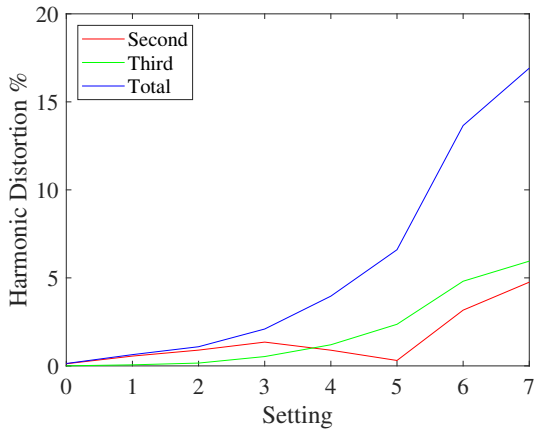

Fig. 7. Experiment 2: Second, third, and total harmonic distortion of the MIC-500 preamplifier at the different settings.

B. Experiment 2

The second set of experimental results considers real room impulse response measurements performed in a room of 2.7 m $\times$ 3.7 m $\times$ 4.3 m. Figure 6 shows the floor plan of the room used in the measurement. The loudspeaker and the microphone are positioned 2.5 m apart at an height of 1.2 m. The measurement system is composed of a Focusrite Scarlett 2i2 audio interface, a RCF Arya Pro5 loudspeaker, a Behringer EMC 8000 microphone. Since previous mesurement shows the room impulse response at 44.1 kHz sampling frequency has less than 8 000 samples, the same input signals of the

previous experiment are applied also in this case. To stress the difference between the different methods, the signals are again passed through a MIC 500 preamplifier in "valve" mode before being applied to the loudspeaker. The output of the MIC 500 preamplifier has been recorded simultaneously to the microphone to allow compensation of the preamplifier. Eight different settings (Setting 0 to 7) of the MIC 500 potentiometer controlling the nonlinearity are considered in this case. Figure 7 shows the second, third, and total harmonic distortion on a 1 kHz tone at the maximum amplitude of the sequences for the different settings. The output SNR is at around 20 dB for most settings, apart from Setting 0 and 1 where it is 6 dB and 15 dB respectively.

In this case the actual room impulse response is unknown. To compare the different measurement in terms of LSD, we assume as reference for each setting the result obtained with the PPS of WN filters of order $k = 3$, memory $N = 8192$, diagonal number $D = 4$, that gave the best results in the previous experiment. Figure 8 shows the LSD in the band $[100, 18\,000]$ Hz measured with the different methods without compensation of the MIC 500 preamplifier. The results obtained are for the most similar to those of the previous experiment. The best performance is obtained with the PPSs for WN filters and with the OPSs for Gaussian inputs for diagonal number D at least 3, but also the PPSs for LN filters and the OPSs for uniform inputs provide here similar good results for sufficiently large D. In the OPSs with smaller D, a small improvement of performance can be obtained by doubling the period of the OPS, i.e., for $\epsilon = 8$ in (39). The OPSs with input samples having Pink distribution are the most affected by the nonlinearities, for the same reason explained in first experiment related to the larger amplitude of the input samples. Also in this case, if the same peak amplitude would be considered for all sequences, the OPS for Pink distribution are among the best performing sequences with performance similar to those of the PPSs for WN filters and with the OPSs for Gaussian inputs. The MLSs are affected by the larger nonlinearities, but provide anyhow acceptable results. The exponential sweeps provide a uniform LSD at all settings, albeit larger than that achievable with the other methods.

The LSD in the band $[100, 18\,000]$ Hz measured when the MIC 500 preamplifier, compensated with the inverted filter, has also been computed. Since in this case also the reference magnitude response is compensated with the inverted filter, the results are very similar to those of Figure 8 and have not been included for space limitations.

VI. Conclusion

The paper has presented a novel RIR measurement method robust towards the nonlinearities that may affect the power amplifier or the loudspeaker. The proposed approach allows measuring the RIR with the cross-correlation method. A periodic input sequence is applied to the measurement system and the corresponding output is measured. The room impulse response is then obtained by computing the cross-correlation between the output signal and an appropriate OPS sequence over a period. The output signal should be at steady state in this period, which means that an input sequence sufficiently longer than one period should be used in the measurement to guarantee this condition. The robustness towards measurements is achieved taking into account the nonlinearities in the development of the OPS sequence. The input sequence can be any periodic persistently exciting sequence and can also be quantized, but the best results have been obtained with Gaussian sequences. Using OPSs, protection against nonlinearities can be increased augmenting the order or the diagonal number of the considered nonlinear model, at the cost of an increased period of the input sequence. In contrast to other cross-correlation methods discussed in the literature, the proposed approach is the only one that directly estimates the first-order kernel of the Volterra filter modeling the measurement systems, i.e., the system impulse response for small signals. Measurements performed both on an emulated scenario and in a real environment have illustrated the validity of the approach also in comparison with other competing approaches, e.g., MLSs, exponential sweeps, and PPSs for WN and LN systems. While the proposed approach is particularly suitable for coping with nonlinearities affecting the measurement system, it can be effectively used also when the measurement system has negligible nonlinearities obtaining an identification performance similar to those of other methods currently adopted in practice, i.e., MLS, exponential sweeps, PPSs. Many OPSs with different order and diagonal numbers have been developed for measuring room impulse responses as long as 262 144 samples and are freely available for download from [53].

References

[1] T. Rossing, Ed., *Springer Handbook of Acoustics*. New York, NY: Springer-Verlag, 2007.

[2] M. Kleiner, *Acoustics and Audio Technology*. Fort Lauderdale, FL: J. Ross Publishing, 2011.

[3] S. Bharitkar and C. Kyriakakis, *Immersive Audio Signal Processing*. New York, NY: Springer-Verlag, 2006.

[4] W. G. Gardner, *3-D Audio Using Loudspeakers*. New York, NY: Kluwer Academic Publishers, 1998.

[5] J. Garas, *Adaptive 3D Sound Systems*. New York, NY: Kluwer Academic Publishers, 2000.

[6] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems – Algorithms and DSP Implementation*. New York, NY: Wiley, 1996.

[7] B. Gunel, "Room shape and size estimation using directional impulse response measurements," in *Proc. Forum Acusticum Sevilla*, 2002.

[8] M. Kuster, "Reliability of estimating the room volume from a single room impulse response," *The Journal of the Acoustical Society of America*, vol. 124, no. 2, pp. 982–993, 2008.

[9] F. Antonacci, J. Filos, M. Thomas, E. Habets, A. Sarti, P. Naylor, and S. Tubaro, "Inference of room geometry from acoustic impulse responses," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 10, pp. 2683–2695, Dec. 2012.

[10] S. Tervo and T. Tossavainen, "3d room geometry estimation from measured impulse responses," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012, pp. 513–516.

[11] Y. E. Baba, A. Walther, and E. A. P. Habets, "3d room geometry inference based on room impulse response stacks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 5, pp. 857–872, 2018.

[12] M. Crocco, A. Trucco, and A. Del Bue, "Uncalibrated 3d room geometry estimation from sound impulse responses," *Journal of the Franklin Institute*, vol. 354, no. 18, pp. 8678–8709, 2017.

[13] J. Zhu and D. Rao, "Performance analysis of room impulse response reshaping using binaural method," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 261, no. 2. Institute of Noise Control Engineering, 2020, pp. 4156–4162.

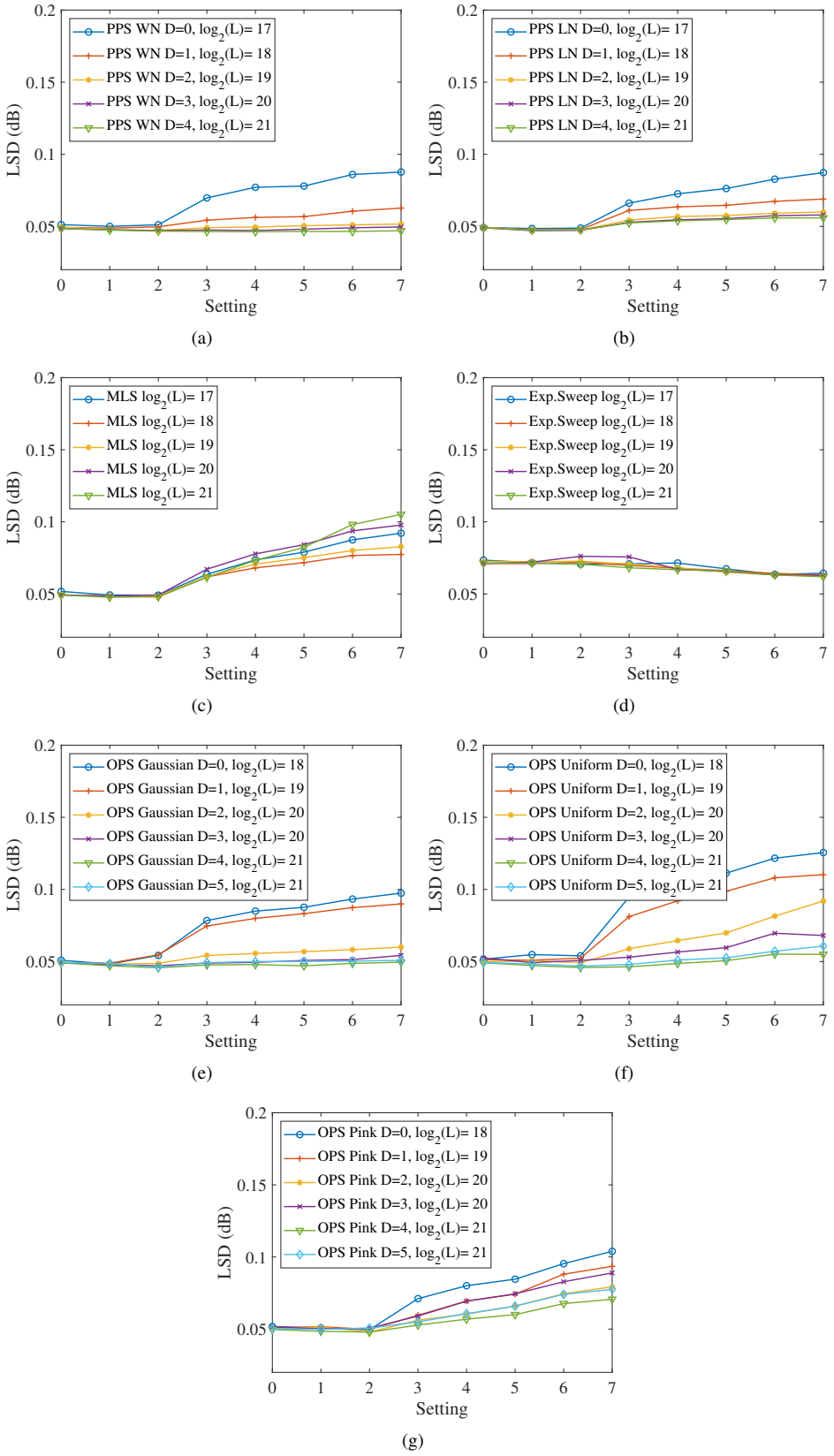

Fig. 8. Experiment 2: Log-spectral distance in the band $[100, 18000]$ Hz at the different settings for (a) PPSs for WN filters, (b) PPSs for LN filters, (c) MLSs, (d) exponential sweeps, (e) OPS with Gaussian distribution, (f) OPS with uniform distribution, (g) OPS with pink distribution, without compensation of the MIC-500.

[14] F. Xu, X. Feng, Y. Shen, and Y. Shen, "A novel closed-form solution for acoustic reflector localization based on room impulse response," in *2020 International Conference on Culture-oriented Science & Technology (ICCST)*. IEEE, 2020, pp. 235–239.

[15] V. Bruschi, S. Nobili, S. Cecchi, and F. Piazza, "An innovative method for binaural room impulse responses interpolation," in *Audio Engineering Society Convention 148*. Audio Engineering Society, 2020.

[16] A. Raikar, K. Nathwani, A. Panda, and S. K. Kopparapu, "Effect of microphone position measurement error on rir and its impact on speech intelligibility and quality," *Proc. Interspeech 2020*, pp. 5056–5060, 2020.

[17] C. Tuna, A. Canclini, F. Borra, P. Götz, F. Antonacci, A. Walther, A. Sarti, and E. A. Habets, "3d room geometry inference using a linear loudspeaker array and a single microphone," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1729–1744, 2020.

[18] A. Schaumberger, "Impulse measurement techniques for quality determination in Hi-Fi equipment, with special emphasis on loudspeakers," *Journal of the Audio Engineering Society*, vol. 19, no. 2, pp. 101–107, 1971.

[19] J. M. Berman and L. R. Fincham, "The application of digital techniques to the measurement of loudspeakers," *Journal of the Audio Engineering Society*, vol. 25, pp. 370–384, Jun. 1977.

[20] L. R. Fincham, "Refinements in the impulse testing of loudspeakers," *Journal of the Audio Engineering Society*, vol. 33, no. 3, pp. 133–140, Mar. 1985.

[21] N. Aoshima, "Computer-generated pulse signal applied for sound measurement," *The Journal of the Acoustical Society of America*, vol. 69, no. 5, pp. 1484–1488, May 1981.

[22] Y. Suzuki, F. Asano, H. Y. Kim, and T. Sone, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses," *The Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1119–1123, Feb. 1995.

[23] M. R. Schroeder, "Integrated-impulse method measuring sound decay without using impulses," *The Journal of the Acoustical Society of America*, vol. 66, no. 2, pp. 497–500, Aug. 1979.

[24] F. J. MacWilliams and N. J. A. Sloane, "Psuedo-random sequences and arrays," *Proceedings of the IEEE*, vol. 64, no. 12, pp. 1715–1729, Dec. 1976.

[25] C. Antweiler and M. Dörbecker, "Perfect sequence excitation of the NLMS algorithm and its application to acoustic echo control," *Annales des Telecommunications*, vol. 49, no. 7-8, pp. 386–397, Jul. 1994.

[26] C. Antweiler and M. Antweiler, "System identification with perfect sequences based on the NLMS algorithm," *AEU-Archiv fur Elektronik und Ubertragungstechnik*, vol. 49, no. 3, pp. 129–134, May 1995.

[27] C. Antweiler, *Multi-Channel System Identification with Perfect Sequences–Theory and Applications–*. John Wiley & Sons, Ltd, 2008, ch. 7, pp. 171–198.

[28] J. Vanderkooy, "Aspects of MLS measuring systems," *Journal of the Audio Engineering Society*, vol. 42, no. 4, pp. 219–231, 1994.

[29] M. Wright and J. Vanderkooy, "Comments on -Aspects of MLS measuring systems- and author's reply," *Journal of the Audio Engineering Society*, vol. 43, no. 1/2, pp. 48, 49, Feb. 1995.

[30] R. C. Heyser, "Acoustical measurements by time delay spectrometry," *Journal of the Audio Engineering Society*, vol. 15, pp. 370–382, Oct. 1967.

[31] H. Biering and O. Z. Pedersen, "System analysis and time delay spectrometry (part I)," *Brüel & Kjær Technical Review*, no. 1, pp. 3–51, 1984.

[32] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio Engineering Society Convention 108*, Paris, France, Feb 2000.

[33] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, Jun. 2001.

[34] A. Novak, P. Lotton, and L. Simon, "Synchronized swept-sine: Theory, application, and implementation," *Journal of the Audio Engineering Society*, vol. 63, no. 10, pp. 786–798, 2015.

[35] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York, NY: Wiley, 2000.

[36] T. Schmitz and J.-J. Embrechts, "Hammerstein kernels identification by means of a sine sweep technique applied to nonlinear audio devices emulation," *Journal of the Audio Engineering Society*, vol. 65, no. 9, pp. 696–710, 2017.

[37] E. Armelloni, A. Bellini, and A. Farina, "Not-linear convolution: A new approach for the auralization of distorting systems," in *Audio Engineering Society Convention 110*, May 2001.

[38] L. Tronchin, "The emulation of nonlinear time-invariant audio systems with memory by means of Volterra series," *Journal of the Audio Engineering Society*, vol. 60, no. 12, pp. 984–996, 2013.

[39] L. Tronchin and V. L. Coli, "Further investigations in the emulation of nonlinear systems with Volterra series," *Journal of the Audio Engineering Society*, vol. 63, no. 9, pp. 671–683, 2015.

[40] P. Burrascano and M. Ciuffetti, "Noise Reduction in the Swept Sine Identification Procedure of Nonlinear Systems," *Applied Sciences*, vol. 11, no. 16, p. 7273, aug 2021.

[41] A. Torras-Rosell and F. Jacobsen, "A new interpretation of distortion artifacts in sweep measurements," *Journal of the Audio Engineering Society*, vol. 59, no. 5, pp. 283–289, May 2011.

[42] D. G. Ćirić, M. Marković, M. Mijić, and D. Šumarac Pavlović, "On the effects of nonlinearities in room impulse response measurements with exponential sweeps," *Applied Acoustics*, vol. 74, no. 3, pp. 375 – 382, Mar. 2013.

[43] C. Antweiler and G. Enzner, "Perfect sequence lms for rapid acquisition of continuous-azimuth head related impulse responses," in *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2009, pp. 281–284.

[44] A. Carini, S. Cecchi, and L. Romoli, "Room impulse response estimation using perfect sequences for legendre nonlinear filters," in *2015 23rd European Signal Processing Conference (EUSIPCO)*, 2015, pp. 2541–2545.

[45] A. Carini, S. Cecchi, and L. Romoli, "Robust room impulse response measurement using perfect sequences for legendre nonlinear filters," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 11, pp. 1969–1982, 2016.

[46] A. Carini, S. Cecchi, A. Terenzi, and S. Orcioni, "On room impulse response measurement using perfect sequences for Wiener nonlinear filters," in *2018 26th European Signal Processing Conference (EUSIPCO)*, 2018, pp. 982–986.

[47] A. Carini, S. Cecchi, and S. Orcioni, "Robust room impulse response measurement using perfect periodic sequences for Wiener nonlinear filters," *Electronics*, vol. 9, no. 11, 2020.

[48] A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Audio Engineering Society Convention 122*, Vienna, Austria, May 2007.

[49] A. Carini, S. Orcioni, and S. Cecchi, "Introducing the Orthogonal Periodic Sequences for the Identification of Functional Link Polynomial Filters," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 5486–5490.

[50] A. Carini, S. Orcioni, A. Terenzi, and S. Cecchi, "Orthogonal periodic sequences for the identification of functional link polynomial filters," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5308–5321, 2020.

[51] A. Carini, S. Cecchi, and S. Orcioni, *Orthogonal LIP Nonlinear Filters*. Oxford, UK: Butterworth-Heinemann, 2018, ch. 2, pp. 15–46.

[52] A. Carini, S. Orcioni, and S. Cecchi, "On room impulse response measurement using orthogonal periodic sequences," in *2019 27th European Signal Processing Conference (EUSIPCO)*. IEEE, 2019, pp. 1–5.

[53] ——, "Orthogonal periodic sequences," `http://www2.units.it/ipl/res_OPSeqs.htm`, 2019, [Online; accessed 2019-11-15].

[54] P. Pedersen, "Subharmonics in forced oscillations in dissipative systems. Part I," *The Journal of the Acoustical Society of America*, vol. 6, no. 4, pp. 227–238, 1935.

[55] J. K. Hubbard, "Subharmonic and nonharmonic distortions generated by high-frequency compression drivers," in *Audio Engineering Society Conference: 6th International Conference: Sound Reinforcement*, May 1988.

[56] J. Stewart, "Detection and diagnoses of subharmonic tones generated in woofers," in *Audio Engineering Society Convention 114*. Audio Engineering Society, Mar 2003.

[57] F. Bolaños, "Measurement and analysis of subharmonics and other distortions in compression drivers," in *Audio Engineering Society Convention 118*. Audio Engineering Society, May 2005.

[58] S. A. Billings, *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. John Wiley & Sons, Chichester, West Sussex, UK, 2013.

[59] D. Comminiello and J. C. Príncipe, *Adaptive learning methods for nonlinear system modeling*. Butterworth-Heinemann, 2018.

[60] H. Tao, X. Li, W. Paszke, V. Stojanovic, and H. Yang, "Robust PD-type iterative learning control for discrete systems with multiple time-delays subjected to polytopic uncertainty and restricted frequency-domain," *Multidimensional Systems and Signal Processing*, vol. 32, no. 2, pp. 671–692, 2021.

[61] L. Zhou, H. Tao, W. Paszke, V. Stojanovic, and H. Yang, "PD-type iterative learning control for uncertain spatially interconnected systems," *Mathematics*, vol. 8, no. 9, pp. 1–18, 2020.
[62] X. Zhang, H. Wang, V. Stojanovic, P. Cheng, S. He, X. Luan, and F. Liu, "Asynchronous Fault Detection for Interval Type-2 Fuzzy Nonhomogeneous Higher-level Markov Jump Systems with Uncertain Transition Probabilities," *IEEE Transactions on Fuzzy Systems*, vol. 6706, no. c, pp. 1–1, 2021.
[63] V. Stojanovic and D. Prsic, "Robust identification for fault detection in the presence of non-Gaussian noises: application to hydraulic servo drives," *Nonlinear Dynamics*, vol. 100, no. 3, pp. 2299–2313, 2020.
[64] M. Scarpiniti, D. Comminiello, and A. Uncini, "Convex combination of spline adaptive filters," in *2019 27th European Signal Processing Conference (EUSIPCO)*, 2019, pp. 1–5.
[65] S. S. Bhattacharjee and N. V. George, "Nearest Kronecker product decomposition based linear-in-the-parameters nonlinear filters," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2111–2122, 2021.
[66] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York: John Wiley and Sons, 2000.
[67] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1976.
[68] G. M. Raz and B. D. V. Veen, "Baseband volterra filters for implementing carrier based nonlinearities," *IEEE Trans. Signal Processing*, vol. 46, no. 1, pp. 103–114, Jan. 1998.
[69] O. Kirkeby, P. Rubak, and A. Farina, "Analysis of ill-conditioning of multi-channel deconvolution problems," in *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99 (Cat. No. 99TH8452)*. IEEE, 1999, pp. 155–158.
[70] R. D. Nowak and B. D. Van Veen, "Random and pseudorandom inputs for Volterra filter identification," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2124–2135, 1994.
[71] G.-O. Glentis and N. Kalouptsidis, "Efficient algorithms for the solution of block linear systems with Toeplitz entries," *Linear Algebra and its Applications*, vol. 179, pp. 85 – 104, 1993.
[72] V. Stojanovic and N. Nedic, "Robust identification of oe model with constrained output using optimal input design," *Journal of the Franklin Institute*, vol. 353, no. 2, pp. 576–593, 2016.
[73] ——, "Robust Kalman filtering for nonlinear multivariable stochastic systems in the presence of non-Gaussian noise," *International Journal of Robust and Nonlinear Control*, vol. 26, no. 3, pp. 445–460, 2016.
[74] V. Stojanovic and V. Filipovic, "Adaptive input design for identification of output error model with constrained output," *Circuits, Systems, and Signal Processing*, vol. 33, no. 1, pp. 97–113, 2014.
[75] V. Filipovic, N. Nedic, and V. Stojanovic, "Robust identification of pneumatic servo actuators in the real situations," *Forschung im Ingenieurwesen*, vol. 75, no. 4, pp. 183–196, 2011.
[76] Y. Huang, J. Benesty, and J. Chen, *Acoustic MIMO Signal Processing*. Springer, Berlin, Germany, 2006.

Alberto Carini (Senior Member, IEEE) received the Laurea degree (summa cum laude) in electronic engineering and the Dottorato di Ricerca (Ph.D.) degree in information engineering from the University of Trieste, Trieste, Italy, in 1994 and 1998, respectively. In 1996 and 1997, during his doctoral studies, he was a Visiting Scholar with the University of Utah, Salt Lake City. From 1997 to 2003, he was a DSP Engineer with Telit Mobile Terminals SpA, Trieste. In 2003, he was with Neonseven srl, Trieste, as an audio and DSP expert. From 2001 to 2004, he collaborated with the University of Trieste as a Contract Professor of digital signal processing. From 2004 to 2018, he was an Associate Professor with the University of Urbino, Urbino, Italy. From 2018, he has been an Associate Professor with the University of Trieste, Trieste, Italy. His research interests include system identification, adaptive filtering, nonlinear filtering, nonlinear equalization, acoustic echo cancellation, active noise control, and room response equalization. He is a Member of EURASIP. His professional activities are as follows: Member-at-Large, SPS Conference Board (2009–2010); Member, SPS Signal Processing Theory and Methods Technical Committee (2006–2011 and 2015-2020); and Program Co-Chair signal processing area, 2009 International Symposium on Image and Signal Processing and Analysis (ISPA) and 2013 International Symposium on ISPA. He is currently an Editorial Board Member of Elsevier Signal Processing. He received the Zoldan Award for the best Laurea degree.

Stefania Cecchi (Member, IEEE) was born in Amandola, Italy, in 1979. She received the Laurea degree (with honors) in electronic engineering from the University of Ancona (now University Politecnica delle Marche, Italy) in 2004 and the Ph.D. degree in electronic engineering from the University Politecnica delle Marche (Ancona, Italy) in 2007. She was a Post Doc Researcher at DII (Department of Information Engineering) at the same university above, from February 2008 to October 2015, and Assistant Professor from November 2015 to October 2018. She is an Associate Professor at the same department since November 2018. She is the author or coauthor of several international papers. Her current research interests are in the area of digital signal processing, including adaptive DSP algorithms and circuits, speech, and audio processing. Dr. Cecchi is a member of the Acoustic Engineering Society (AES), Institute of Electrical and Electronics Engineers (IEEE), Italian Acoustical Association (AIA).

Alessandro Terenzi was born in Senigallia, Italy, in 1991. He received the Laurea degree (with honors) in electronic engineering in July 2016 at the Polytechnic University of Marche (Italy), and the Ph.D degree in Electronic engineering on March 2021. He is now a Post Doc Researcher at DII (Department of Information Engineering) at the same university. His current research interests are in the area of digital signal processing, including nonlinear audio system and audio processing. Dr. Terenzi is a member of the Acoustic Engineering Society (AES).

Simone Orcioni (Senior Member, IEEE) received the Laurea degree in Electronics Engineering from Università degli Studi di Ancona, Italy, in 1992 and the Ph.D. degree in 1995. From 1997 to 1999 he was Post-Doctoral Fellow at the same university. In 2000 he became Assistant Professor, teaching courses in analog and digital electronics, and publishing a text book. In 2017 he was Guest Professor at Ubiquitous Computing Lab (UC-Lab), HTWG Konstanz - University of Applied Sciences and lecturer of the course "Digital Signal Processing". From 2021 he is Associate Professor at Department of Information Engineering of Università Politecnica delle Marche. He has published more than fifty papers in journals, and more than a hundred in proceeding of international conference and chapters of international books. He was Guest Editor for EURASIP Journal on Embedded Systems, Frontiers in Energy Research and MDPI Sensors, Reviewer for many International Journals, Program Chair of three International Conferences, and Editor of four International Books. He has been working in statistical device modeling and simulation, analog circuit design, cyber-physical system simulation and linear and nonlinear system identification. His current research interests include dsp and power electronics for renewable energies.