# Using Periodic Sequences for HRTFs Measurement Robust Towards Nonlinearities in Automotive Audio Applications


S. Cecchi, V. Bruschi, S. Nobili, A. Terenzi
Dept. of Information Engineering,
Università Politecnica delle Marche, Italy
Email:s.cecchi@univpm.it

A. Carini
Dept. of Engineering and Architecture
University of Trieste, Italy
Email:acarini@units.it



***Abstract*—Head Related Transfer Functions (HRTFs) are a mathematical representation of the acoustic transfer functions between a sound source and the listener's ears. They play an important role for both creating immersive audio scenarios and evaluating the perception of a sound system from human being point of view. These applications are also important for automotive scenario, since the car is nowadays the most used audio listening environment. Automotive audio has been attracting a great deal of attention in recent years as well as the use of HRTFs inside the car for creating these immersive scenarios. In this context, we proposed a technique for HRTFs measurement based on perfect or orthogonal periodic sequences that allows an effective measurement robust towards system nonlinearities. Experiments with an emulated scenario and measurements performed in real environments illustrate the validity of the approach in comparison with other competing HRTFs measurement methods of the state of the art.**



***Index Terms*—Head Related Transfer Functions measurement, Perfect Periodic Sequences, Orthogonal Periodic Sequences, Automotive audio applications**


## I. Introduction

Head related transfer functions (HRTFs) are a mathematical representation of the acoustic transfer functions between a sound source and the listener's ears. They are measured using in-ear measurement performed by artificial heads or miniature microphones positioned inside the listener's ears. In this way, HRTFs are capable to contain all the cues important for the sound localization, such as interaural level differences (ILD), interaural time differences (ITD) and spectral cues used by the human brain to localize a sound source in the space.

The HRTFs are used in several audio applications [1]. They are used for creating immersive audio scenarios such as in binaural rendering over headphones, where the virtual source is created by filtering the audio signal with the selected HRTFs or in the crosstalk cancellation systems, where the HRTFs are used to cancel the crosstalk contributions due to the loudspeakers superposition in binaural reproduction.

These applications are also important for automotive scenario, since the car is nowadays the most used audio listening environment. Automotive audio has been attracting a great deal of attention in recent years as well as the application of immersive audio in the car cockpit, as reported in [2], [3]. The HRTFs can be used inside the car for creating immersive scenarios through binaural techniques [4] or they can be used for evaluating the perception of a sound system from human being point of view recreating a virtual environment [5].

Focusing on the measurement procedure, the main methodologies used in the state of the art are the pseudo random sequences (maximum length sequence (MLS)/inverse repeated sequence (IRS) and Golay code) [6], [7] or sweep signals (linear and exponential sweeps) [8], [9]. A detailed description of these methodologies can be found in the review paper [1]. Several problems can occur during the HRTF measurement causing errors, such as nonlinear distortions of the electro-acoustic systems, environmental noises, reflections from measurement environments, characteristics of sound sources, and temperature changes [1], [10]. While most of these problems can be solved measuring the HRTFs in a controlled environment, such as an anechoic chamber, the problem of nonlinear distortions can be solved through the use of an appropriate measurement procedure and stimuli [11]–[13].

In this context, we proposed a technique for HRTFs measurement based on perfect and orthogonal periodic sequences that allows an effective measurement robust towards system nonlinearities, as those originated by the power amplifier or the loudspeaker in presence of high reproduction volumes. Starting from the results obtained in [14] and [15], perfect and orthogonal periodic sequences are applied for the first time to the measurement of the HRTFs. Considering a real car environment and an emulated scenario, some results are reported to illustrate the validity of the approach in comparison with other competing HRTFs measurement methods of the state of the art.

The paper is organized as follows. Section II describes the proposed methodology and introduces more in detail the periodic perfect sequences (PPS) and the orthogonal perfect sequences (OPS). The measurement system has been tested taking into comparison other measurements methods of the state of the art. In particular, the results obtained in a real scenario are reported in Section III-A while the results obtained in an emulated nonlinear scenario are reported in Section III-B. Finally, conclusions are drawn in Section IV.

## II. FLIP FILTERS, PPSS, AND OPSS

In the proposed methodology, the measurement system is modelled as a mildly nonlinear system, in particular, as a functional link polynomial (FLiP) filter. The objective of the measurement is the estimation of the transfer function of the linear part of the FLiP filter, the so called first order kernel. Neglecting the effect of the loudspeaker and microphone, the measured transfer functions coincides with the HRTF. Directly modelling the measurement system as a nonlinear filter, allow us to perform a robust measurement also when nonlinearties due to the amplifier or the loudspeaker affect the measure. In what follows, we briefly review the FLiP filters and the first order kernel measurement methodologies based on perfect periodic sequences (PPSs) and orthogonal periodic sequences (OPSs).

### A. FLiP filters

FLiP filters [16] are a broad class of nonlinear filters, which includes the Volterra filters, the Wiener nonlinear (WN) filters deriving from the truncation of the Wiener series, the Legendre nonlinear (LN) filters, and many others.

A FLiP filter of memory $N$, order $K$, diagonal number $D$ has the following input-output relationship

$$\begin{aligned} y(n) \; = \; & h_0 + \sum_{m=0}^{N-1} h_{1,m} f_1[x(n-m)] + \\ & \sum_{r=1}^{K} \sum_{i_1=0}^{D} \sum_{i_2=i_1}^{D} \cdots \sum_{i_{r-1}=i_{r-2}}^{D} \sum_{m=0}^{N-i_{r-1}} h_{r,i_1,i_2,\ldots,i_{r-1}} \cdot \\ & f_{r,i_1,i_2,\ldots,i_{r-1}}[x(n-m), x(n-m-i_1), \\ & \quad x(n-m-i_2) \ldots, x(n-m-i_{r-1})] \end{aligned} \quad (1)$$

where $x(n)$ are the input samples, $h_0$ is a constant term usually neglected in audio applications, $h_{1,m}$ is the first order kernel, $f_1[\cdot]$ is the first order basis function with $f_1[x(n)] = x(n)$, $h_{r,i_1,\ldots,i_{r-1}}$ is the $r$-th order kernel, and $f_{r,i_1,\ldots,i_{r-1}}[\cdot]$ are the basis functions of order $r$, i.e., polynomials of order $r$ that are product of "generating" polynomials in the arguments. Note that the diagonal number $D$ is the maximum time difference between the samples involved in the filter and in real systems is often limited to low values. Moreover, it should be noted that the filter is composed by a linear filter

$$\sum_{m=0}^{N-1} h_{1,m} x(n-m)$$

plus a combination of higher order polynomial basis functions. By choosing different sets of generating polynomials, different families of nonlinear filters can be obtained. In Volterra filters the generating polynomials are the monomials $x$, $x^2$, $x^3$, ...; in LN filters they are Legendre polynomials; in WN filters they are Hermite polynomials. LN filters have orthogonal basis function for a white uniform distribution of the input samples. WN filters have orthogonal basis function for a Gaussian distribution of the input samples. Thus, LN and WN filters are orthogonal FLiP filters and for this reason they admit PPSs, i.e., periodic sequences that guarantee the same orthogonality of the basis functions on a finite period [17]. In contrast, Volterra filters are non orthogonal filters for any input sample distribution and do not admit PPSs, but still they admit OPSs as detailed in the following.

### B. Perfect periodic sequences

PPSs for LN and WN filters have been developed in [18], [17] by considering a set of variables representing the samples of the periodic sequence and imposing the orthogonality of all basis functions of the filter. In this way, an underdetermined system of nonlinear equations has been obtained and solved using iterative methods.

Since for a PPS input all basis functions are mutually orthogonal and $x(n-m)$ is a basis function, the linear kernel $h_{1,m}$ can be measured with the cross-correlation method, i.e., computing the cross-correlation between the system output and the PPS input sequence:

$$h_{1,m} = \frac{< y(n)x(n-m) >_P}{< x^2(n) >_P}, \quad (2)$$

where $< \cdot >_P$ indicates the sum of the terms between angular brackets for $n$ ranging over a period $P$ of the sequence.

### C. Orthogonal periodic sequences

Volterra filters do not admit PPSs, but they can still be identified with the cross-correlation method using OPSs. Given any persistently exiting periodic input sequence $x(n)$ of sufficiently large period $P$, it has been shown in [19] that an orthogonal periodic sequence $z(n)$ of period $P$ can be developed such that

$$h_{1,m} =< y(n)z(n-m) >_P . \quad (3)$$

The OPS has been developed by imposing $< x(n)z(n) >_P = 1$ and at the same time the orthogonality of $z(n)$ with all other Volterra basis functions. In this way, an underdetermined linear equation system in the variables $z(n)$, for $n$ ranging over a period $P$, has been derived and solved obtaining the OPS.

### D. OPSs vs PPSs

For the same memory length $N$, order $K$, diagonal number $D$, and period $P$, the estimation of an OPS is much simpler and faster than a PPS, since a linear equation system rather than a nonlinear one has to be solved. Nevertheless, OPSs and PPSs have a different behavior in presence of noise. To compare sequences of different periods on equal terms, the noise gain has been introduce in [19] and is defined as

$$G_\nu = \frac{\text{MSD}}{E[\nu(n)^2]} < x^2(n) >_P, \quad (4)$$

where MSD is the mean square deviation in the coefficient estimate, i.e., MSD $= E[(\tilde{h}_{1,m} - h_{1,m})^2]$, with $\tilde{h}_{1,m}$ the true value and $h_{1,m}$ the estimated one, and $E[\nu(n)^2]$ is the noise variance. It can be proved that PPSs have always $G_\nu = 1$. On the contrary, for OPSs when the period $P$ is small, i.e., close to the minimum value allowed by the conditions of the linear equation system, the noise gain $G_\nu$ assumes very large values that make the identification with OPSs useless. Nevertheless, for large periods $G_\nu$ assumes reasonable values that make the identification with OPSs feasible and useful. We have found experimentally that for the same values of $N$, $K$ and $D$, the period of the OPS should be twice that of a PPS to obtain reasonable values of $G_\nu$.

Eventually, we must point out that PPSs can be applied to the identification of the first order kernel only of orthogonal FLiP filters, e.g., LN and WN filters, while the first order kernel of Volterra filters can be estimated with the cross-correlation method only using OPSs. Moreover, the first order kernel of the Volterra model coincides with the impulse response of the system when the input signal amplitude tends to zeros, which is not the case for LN and WN filters.

## III. Experimental results

Experimental results have been divided in two phases. In the first one, the validation of the proposed approach in a real scenario has been proposed. Since the considered car environment has shown a low level of nonlinearity, a second validation section with an emulated scenario has been realized in order to create different distortion levels.

### A. Real scenario

Fig. 1 represents the scheme of the HRTFs measurement. The acquisition is performed by a Bruel&Kjaer Head and Torso Simulator Type 4128, with right and left Ear Simulators (Type 4158 and 4159) connected to a Sound Card Focusrite 2i2 through the Bruel&Kjaer microphone preamplifier Type 2829. For the reproduction system, the car sound system has been considered connecting them to the same sound card using the Auxiliary car audio port. The reproduction and the acquisition synchronization is managed by a PC through the NU-Tech software [20] exploiting ASIO drivers. All the measurements have been performed within the semi-anechoic chamber of the A3lab group (Dept. of Information Engineering, UNIVPM) to ensure a small environmental noises as shown in Fig. 2.


Fig. 1. Overall scheme of the acquisition procedure

Fig. 2. Car used for the experiments with the B&K mannequin. The experiments have been performed inside the semi-anechoic chamber of the A3Lab group at Università Politecnica delle Marche.

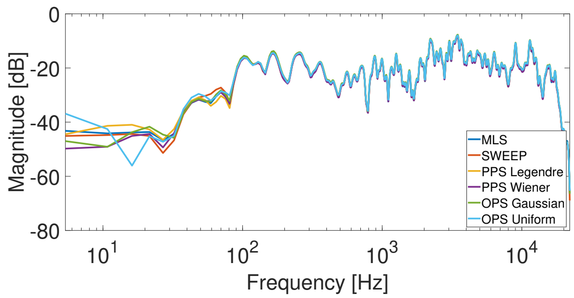
Several experiments have been performed considering the driver and passenger positions. The measurements performed with OPS (with input samples having Gaussian and uniform distribution) and PPSs for LN and WN filters are compared with measurements based MLSs, and exponential sweeps.

The OPSs have memory length $N = 2048$ samples, order $K = 3$, diagonal number $D = 3$, period $L = 262,144$. The OPS input samples are also quantized in the set $[-512 : +512]/512$. The PPSs for WN and LN filters have $N = 2048$, $K = 3$, $D = 3$, and $L = 262,120$ (to have a period comparable with the OPS) and the samples are represented 24 bits. The MLSs have period $2^{18} - 1$. The exponential sweeps have length $262,144$ and sweep from around 20 Hz till $22,050$ Hz. The sampling frequency is $44,100$ Hz. The same power has been considered for all input signals.

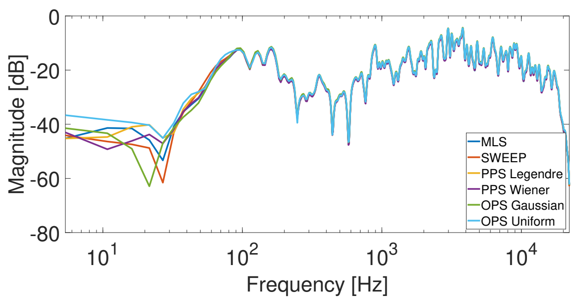
For the sake of brevity just the results for the driver position have been reported in Fig. 3. It is clear that all the methodologies present similar results since the car audio system shows a quasi-linear behaviour, following the results we have obtained in [14] for the specific application of room response identification. Other results will be presented in the next section considering a more nonlinear system.

### B. Emulated scenario

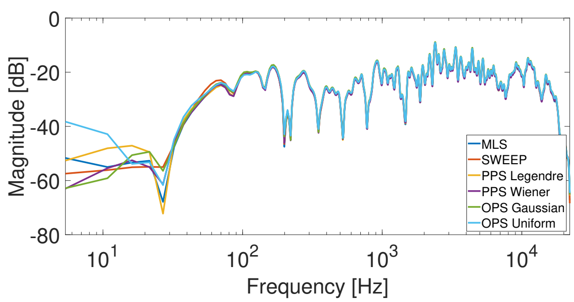
To test the proposed approach with different levels nonlinearity and to study the effect of noise, an emulated system has been considered. The input signals of the previous experiment were applied to a Behringer MIC 100 vacuum tube preamplifier and the corresponding output, recorded with a Focusrite Scarlett 2i2 audio interface, was convolved with four HRTFs of 8192 samples previously measured inside the real car environment. Specifically, the HRTFs were those measured with the exponential sweep in the previous experiment. The preamplifier has a potentiometer that allows to set different levels of nonlinear distortion. Twenty-one different settings have been considered and Fig. 4 shows the second, third and total harmonic distortion on a tone at 1 kHz having the same power of the applied signals at the different distortion settings.



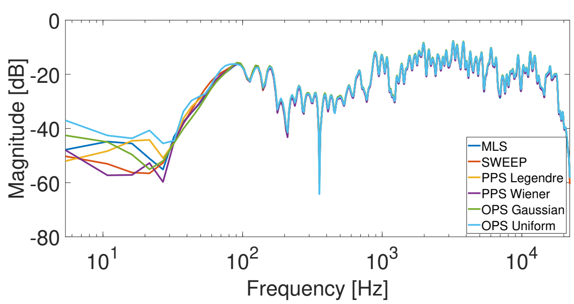

(a)




(b)




(c)




(d)

Fig. 3. HRTFs measurement at driver position, where 3(a) is the response of the left ear from the left speakers, 3(b) is the response of the left ear from the right speakers,3(c) is the response of the right ear from the left speakers and 3(d) is the response of the right ear from the right speakers. A smoothing of 1/12 was applied to the responses.

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. Many of the harmonic distortions of Fig. 4 are much greater than those normally found in an impulse response measurement system.

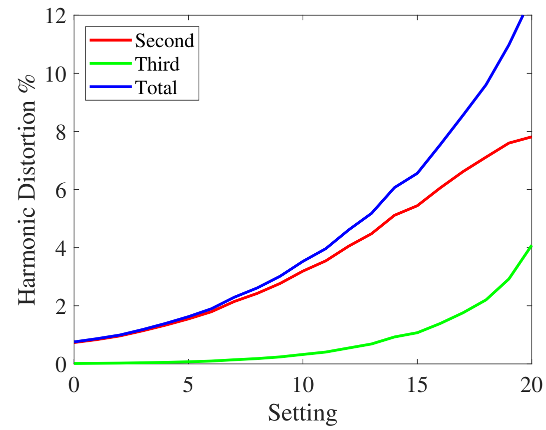



Fig. 4. Second, third, and total harmonic distortion of the MIC-100 preamplifier at the different settings.

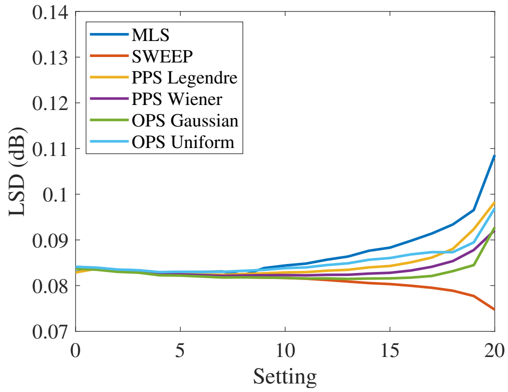
The objective of this experiment is to stress the robustness towards the nonlinearities of the different methods. In what follows the different methods will be compared in terms of log-spectral distance (LSD) between the measured impulse response and the convolved HRTF impulse response. Note that the resulting LSD is affected by the impulse response of the DAC, power amplifier, and ADC, which are not compensated. The LSD is defined in a 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:

$$\mathrm{LSD} = \sqrt{\frac{1}{k_2 - k_1 + 1} \sum_{k=k_1}^{k_2} \left[ 10 \log_{10} \frac{|H(k)|^2}{|\hat{H}(k)|^2} \right]^2}, \quad (5)$$

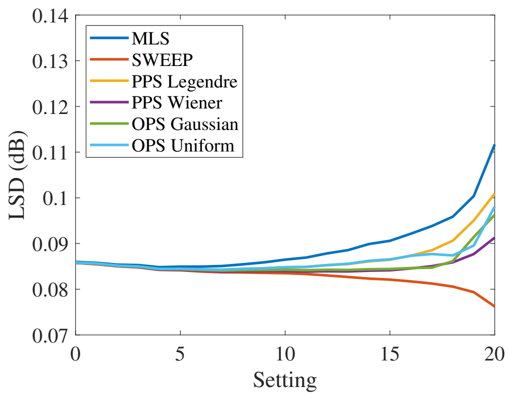
where $H(k)$ is the reference HRFT and $\hat{H}(k)$ is the measured HRTF. In the experiment, the LSD was measured in the band $[100, 18000]$ strictly inside the pass-band.

Fig. 5 shows the LSD measured at the different settings for the four HRTFs when no artificial noise is added to the output. The only noise in the system is that generated by the power pre-amplifier and the SNR is larger than 60 dB. In these conditions, all the curves coincide at the lowest settings, i.e., for low distortions, but they separate at larger distortions. The more flat is the curve, the more immune is the method towards the nonlinearities. The PPSs and the OPSs, especially the PPS for WN filters and the OPS for Gaussian input, provide the best results. For the exponential sweep, the LSD tends to reduce with larger distortions, which means that the alterations introduced by the nonlinearities on the measured response tend to compensate the effect of the power amplifier response in this case, but the measure is anyway altered with respect to the lower distortion conditions.

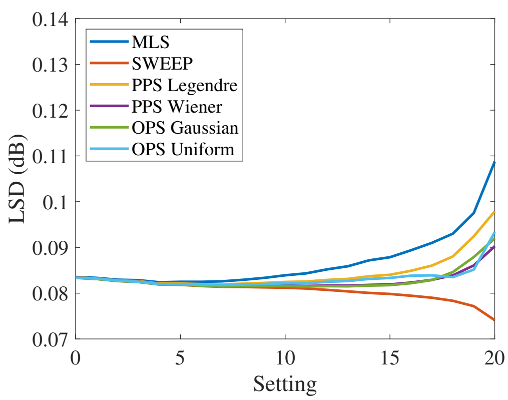
To study the effect of noise, a white Gaussian noise has been added on the output signals used in the measurement to reach a 40 dB SNR. Each measurement has been repeated 100 times with a different output noise and the resulting LSDs values have been averaged. Fig. 6 shows the results obtained in these conditions. It is immediately evident the rise of the curves



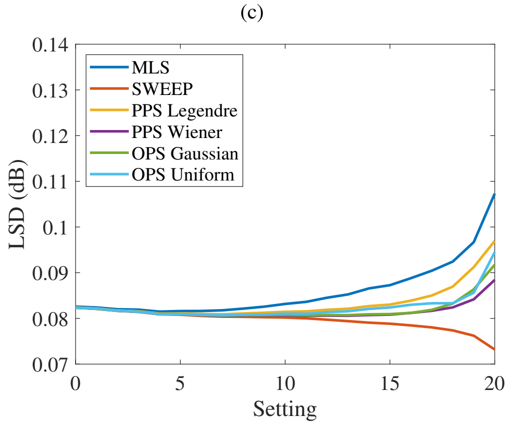

(a)




(b)

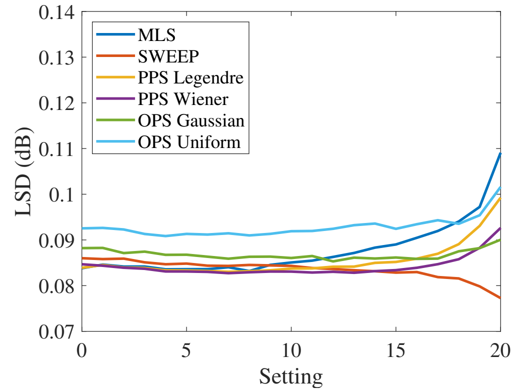



(c)

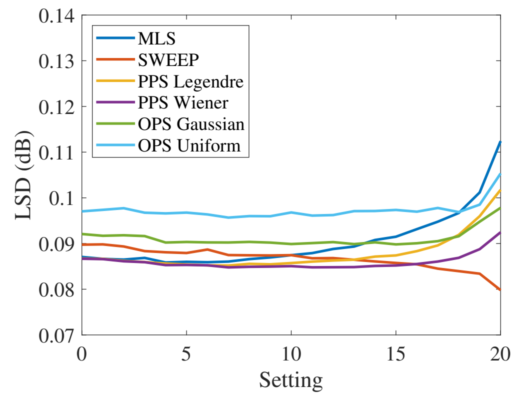



(d)

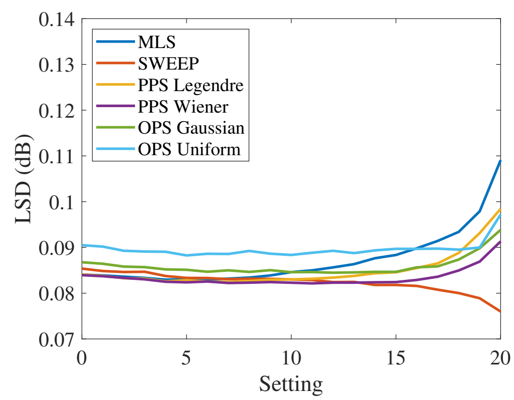
Fig. 5. Log-spectral distance in the band [100, 18000] Hz at the different settings for 5(a) the response of the left ear from the left speakers, 5(b) the response of the left ear from the right speakers,5(c) the response of the right ear from the left speakers and 5(d) the response of the right ear from the right speakers and no artificially added noise, i.e. with SNR larger than 60 dB.



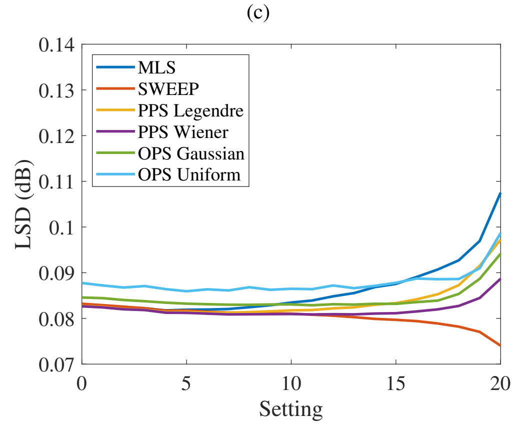

(a)




(b)




(c)




(d)

Fig. 6. Log-spectral distance in the band [100, 18000] Hz at the different settings for 6(a) the response of the left ear from the left speakers, 6(b) the response of the left ear from the right speakers,6(c) the response of the right ear from the left speakers and 6(d) the response of the right ear from the right speakers with a 40 dB output noise.

obtained with the OPSs, due to a noise gain larger than 1. The noise gain on the OPS with Gaussian input is 8.8 dB, and on the OPS with uniform input is 12.5 dB. These large values are due to the short length of the OPS input sequences. The effect of noise on short length OPSs, i.e., with length similar to those used for HRTF measurements, is much larger than that experienced on room impulse response (RIR) measurements [19]. In the experiment, the best results are obtained by the PPSs, which have a noise gain 1, in particular the PPSs for WN filters are the less affected by nonlinearities and are good candidate for HRTF measurements.

## IV. Conclusions

The paper has presented a novel HRTFs measurement method robust towards the nonlinearities that may affect the car sound system (i.e., power amplifier or the loudspeakers). Starting from previous methodologies presented by the same authors, the proposed approach is applied for the first time to the measurement of the HRTFs. Using a periodic input sequence, the HRTFs are obtained by computing the cross-correlation between the output signals and an appropriate perfect or orthogonal periodic sequence over a period. The robustness towards measurements is achieved taking into account the nonlinearities in the development of the PPS and OPS sequences. In contrast to other cross-correlation methods discussed in the literature, the approach based on OPSs 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. The first experiment on a real scenario with small nonlinearities has demonstrated the validity of the approach in comparison with other methods of the state of the art. To stress the robustness towards the nonlinearities of the different methods, an emulated scenario has been analyzed considering the HRTFs previously measured in the car environment. Also in this case we have tested the potentiality of the proposed approach and, in particular, the PPS for WN filters has shown a good immunity to the system nonlinearities and noise confirming to be a good candidate for HRTFs measurements.

## Acknowledgment

This work is supported by Marche Region in implementation of the financial programme POR MARCHE FESR 2014-2020, project "Miracle" (Marche Innovation and Research fAcilities for Connected and sustainable Living Environments), CUP B28I19000330007.

## References

[1] S. Li and J. Peissig, "Measurement of head-related transfer functions: A review," *Applied Sciences*, vol. 10, no. 14, p. 5014, 2020.

[2] T. Dupré, S. Denjean, M. Aramaki, and R. Kronland-Martinet, "Spatial sound design in a car cockpit: Challenges and perspectives," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2021, pp. 1–5.

[3] D. Pinardi, A. Farina, and J.-S. Park, "Low frequency simulations for ambisonics auralization of a car sound system," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2021, pp. 1–10.

[4] C. Lin, Y. Chen, and Y. Wang, "Partial update adaptive filtering based on head-related model for in-vehicle audio enhancement," in *2012 International Conference on Connected Vehicles and Expo (ICCVE)*, 2012, pp. 226–230.

[5] F. Piazza, S. Squartini, R. Toppi, M. Navarri, M. Pontillo, F. Bettarelli, and A. Lattanzi, "Industry-oriented software-based system for quality evaluation of vehicle audio environments," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 3, pp. 855–866, 2006.

[6] W. Gardner and K. D. Martin, "Hrtf measurements of a kemar," *The Journal of the Acoustical Society of America*, vol. 97, no. 6, pp. 3907–3908, 1995.

[7] Q. Ye, Q. Dong, Y. Zhang, and X. Li, "Fast head-related transfer function measurement in complex environments," in *Proceedings of the 20th International Congress on Acoustics, Sydney, Australia*, 2010, pp. 23–27.

[8] S. Muller and P. Massarani, "Transfer-Function Measurement with Sweeps," *J. Audio Eng. Soc.*, vol. 49, no. 6, pp. 443–471, Jun. 2001.

[9] M. Rothbucher, K. Veprek, P. Paukner, T. Habigt, and K. Diepold, "Comparison of head-related impulse response measurement approaches," *The Journal of the Acoustical Society of America*, vol. 134, no. 2, pp. EL223–EL229, 2013.

[10] B. Xie, *Head-related transfer function and virtual auditory display*. J. Ross Publishing, 2013.

[11] A. Carini, S. Cecchi, L. Romoli, and G. L. Sicuranza, "Perfect Periodic Sequences for Legendre Nonlinear Filters," in *Proc. 22nd European Signal Processing Conference*, Lisbon, Portugal, Sep. 2014, pp. 2400–2404.

[12] A. Carini, S. Cecchi, and L. Romoli, "Room Impulse Response Estimation using Perfect sequences for Legendre Nonlinear filters," in *Proc. 23nd European Signal Processing Conference*, Nice, France, Aug. 2015.

[13] A. Carini, L. Romoli, S. Cecchi, and S. Orcioni, "Perfect Periodic Sequences for Nonlinear Wiener Filters," in *Proc. 24th European Signal Processing Conference*, Budapest, Hungary, Aug. 2016.

[14] A. Carini, S. Cecchi, A. Terenzi, and S. Orcioni, "A room impulse response measurement method robust towards nonlinearities based on orthogonal periodic sequences," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 3104–3117, 2021.

[15] 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.

[16] A. Carini, S. Cecchi, and S. Orcioni, "Orthogonal LIP nonlinear filters," in *Adaptive learning methods for nonlinear system modeling*. Elsevier, 2018, pp. 15–46.

[17] A. Carini, S. Orcioni, A. Terenzi, and S. Cecchi, "Nonlinear system identification using Wiener basis functions and multiple-variance perfect sequences," *Signal Processing*, vol. 160, pp. 137–149, 2019.

[18] A. Carini, S. Cecchi, L. Romoli, and G. L. Sicuranza, "Legendre nonlinear filters," *Signal Processing*, vol. 109, pp. 84–94, 2015.

[19] 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.

[20] A. Lattanzi, F. Bettarelli, and S. Cecchi, "NU-Tech: the entry tool of the hArtes toolchain for algorithms design," in *Proc. 124th Audio Engineering Society Convention*, Amsterdam, The Netherlands, May 2008, pp. 1–8.