KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	G. Vairetti, N. Kaplanis, E. De Sena, S. H. Jensen, S. Bech, M. Moonen, and T. van Waterschoot (2017), **The Subwoofer Room Impulse Response database (SUBRIR)** Journal of the Audio Engineering Society, Vol. 65, No. 05, May 2017
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	https://doi.org/10.17743/jaes.2017.0007
Journal homepage	http://www.aes.org/journal/
Author contact	giacomo.vairetti@esat.kuleuven.be *+ 32 (0)16 321817*
IR	url in Lirias https://lirias.kuleuven.be/handle/123456789/572970

(article begins on next page)

The Subwoofer Room Impulse Response database (SUBRIR)

GIACOMO VAIRETTI[*1], **NEOFYTOS KAPLANIS**[2,4], *AES Student Member*,
ENZO DE SENA[3], **SØREN HOLDT JENSEN**[4], **SØREN BECH**[2,4], *AES Fellow*,
MARC MOONEN[1], *AES Associate Member*, **and TOON VAN WATERSCHOOT**[1,5], *AES Associate Member*
({giacomo.vairetti, marc.moonen, toon.vanwaterschoot}@esat.kuleuven.be)
(e.desena@surrey.ac.uk) ({neo, sbe}@bang-olufsen.dk) (shj@es.aau.dk)

[1]*KU Leuven, Dept. of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, 3001 Leuven, Belgium.*
[2]*Bang & Olufsen A/S, Peter Bangs Vej 15, 7600 Struer, Denmark.*
[3]*Institute of Sound Recording, University of Surrey, Guilford, Surrey, GU2 7XH, UK.*
[4]*Dept. of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7B, 9220 Aalborg, Denmark.*
[5]*KU Leuven, Dept. of Electrical Engineering (ESAT), ETC, AdvISe Lab, Kleinhoefstraat 4, 2440 Geel, Belgium.*

This report introduces a new database of room impulse responses (RIRs) measured in an empty rectangular room using subwoofers as sound sources. The purpose of this database, publicly available for download, is to provide acoustic measurements within the frequency region of modal resonances. Performing acoustic measurements at low frequencies presents many difficulties, mainly related to ambient noise and to unavoidable nonlinearities of the subwoofer. In this report, it is shown that these issues can be addressed and partially solved by means of the exponential sine-sweep technique and a careful calibration of the measurement equipment. A procedure for estimating the reverberation time at very low frequencies is proposed, which uses a cosine-modulated filterbank and an approximation of the RIRs using parametric models in order to reduce problems related to low signal-to-noise ratio and to the length of typical band-pass filter responses.

0 INTRODUCTION

Room impulse response (RIR) measurements are essential to assess the performance of acoustic signal enhancement algorithms, e.g. for applications such as dereverberation [1], source separation [2], source localization [3], blind acoustic parameter estimation [4], convolutive reverb [5], and many others. Several available RIR databases [1–7] are intended for different audio signal processing tasks, each requiring a different choice of measurement technique and of the measuring equipment. For instance, the databases in [6] and [7] contain binaural and head-related RIRs, and are useful in hearing-aids applications. Other databases present specific configurations of the microphones, usually arranged into arrays. What is common to all these databases is that they use full-range loudspeakers, whose frequency response typically has a lower bound of 50-100 Hz. While these databases cover a frequency range sufficient for the development and evaluation of speech enhancement algorithms, information about a significant portion of the modal response of the room is missing.

Nowadays, home audio systems generally include a subwoofer, which is intended for the reproduction of low-frequency content typically in the region between 20 Hz and 150 Hz. In this frequency range, small-sized typical rooms operate within the modal frequency region [8]. In small-sized rooms, most of the acoustical problems are actually due to poor acoustics at very low frequencies (LFs). The modal resonances are usually well separated, energetic, and detectable by the human ear [9], thus degrading the perceived sound quality. A subwoofer with small enough lower cut-off frequency can even partially excite the so-called cavity mode (i.e. the modal resonance centered at 0 Hz). Therefore, algorithms for home audio system applications, such as room compensation algorithms, should be validated also on

[*]To whom correspondence should be addressed. Tel: +32-16-321817; e-mail: giacomo.vairetti@esat.kuleuven.be

RIRs measured within the frequency region of modal resonances. Moreover, such RIRs may provide new insights and be useful to validate physical models of room acoustics, although detailed information about the boundaries conditions are not available. To the authors' best knowledge, a RIR database measured at very LFs is not yet available.

The Subwoofer Room Impulse Response (SUBRIR) database introduced in this report is a collection of RIRs measured in a standard domestic listening room using a subwoofer as the sound source. Two subwoofers with different characteristics and two types of omnidirectional microphones were used to measure the RIR at different locations, for a total of 96 measurements[1]. Performing acoustic measurements at very LFs presents some difficulties, mainly related to LF ambient noise and to unavoidable nonlinear distortions of the subwoofer [11].

Nonlinear distortions can be divided into two categories: *regular* nonlinear distortions refer to systematic and reproducible distortions, such as harmonic spectral components, whose impact to the overall performance of the loudspeaker can be controlled in the design process [12]. *Irregular* nonlinear distortions are instead due to loudspeaker defects and are less easily reproducible and controllable [13]. The main irregular distortion artifact noticed in the measurements presented in this report was recognized as the so-called *rub & buzz* distortion [13–15]. This is a signal-dependent distortion caused by defects due to manufacturing errors, aging or overload. Possible causes of this type of distortion are buzzing parts (e.g. a loose glue joint), the voice coil rubbing or bottoming (i.e. hitting the backplate due to over-displacement), loose particles, air leakages, etc.

The family of methods for measuring RIRs known to have a high immunity against distortion artifacts is the one where a sweep is used as the excitation signal [16–18]. This report shows that the Exponential Sine-Sweep (ESS) technique [19] is particularly suitable for measuring good quality LF-RIR measurements regardless of all the difficulties mentioned above. The ESS is known to provide a better signal-to-noise ratio (SNR) and a better rejection of distortion artifacts than other RIR measurement techniques [20–23].

This report also outlines a procedure to estimate reverberation time (RT) at very LFs. Indeed, the standard specifications [24] are not applicable in this frequency region due to the low SNR [25] and to the influence of the response of the band-pass filters of the filterbank [26]. The proposed approach uses a cosine-modulated filterbank, which reduces the bias introduced by typical filterbanks at low frequencies, and a representation of the RIRs using orthonormal basis function (OBF) models [27], which allows to remove the effect of the noise floor.

The report is structured as follows. In Section 1, a brief summary of the ESS technique is given, together with comments on advantages and disadvantages of the technique. Section 2 describes the room in which the measurements were performed, together with details of the measurement equipment. In Section 3, an analysis of the measurements performed is given; the recorded signals and the retrieved RIRs are analyzed and guidelines on how to obtain good quality measurements are provided. In Section 4, values for the frequency-dependent RT at LFs are estimated with the proposed approach. Section 5 concludes the report and summarizes the recommendations for performing LF-RIR measurements.

1 THE EXPONENTIAL SINE-SWEEP (ESS) MEASUREMENT TECHNIQUE: A SUMMARY

This section reviews the key points of the ESS technique and discusses its applicability in measuring LF-RIRs. A detailed treatment of the ESS technique can be found in [19,20].

The excitation signal used by the ESS technique is a sweep signal with instantaneous frequency (IF) increasing exponentially with time. The IF at time t of the sweep signal of duration T is given by

$$f(t) = e^{(1-(t/T))\ln(f_a)+(t/T)\ln(f_b)} = f_a \left(\frac{f_b}{f_a}\right)^{(t/T)}, \quad (1)$$

where f_a and f_b are the starting frequency and stopping frequency, respectively. The instantaneous phase is obtained by integrating (1) between 0 and t, and used as the argument of a sinusoidal function, leading to the excitation signal,

$$s(t) = \sin\left(\frac{2\pi T}{\ln\left(\frac{f_b}{f_a}\right)}\left(f(t) - f_a\right)\right). \quad (2)$$

The excitation signal, $s(t)$, is fed to the loudspeaker and the response $y(t)$ is recorded with a microphone. The RIR $\hat{h}(t)$ is retrieved by linear convolution of the recorded signal $y(t)$ with the so-called inverse signal $v(t)$ ($\hat{h}(t) = y(t) \otimes v(t)$, with $\otimes$ indicating convolution). The inverse signal is built such that the linear convolution of the sweep signal with the inverse signal produces a shifted delta function $s(t) \otimes v(t) = \delta(t - T)$. The inverse signal can be obtained by time-reversing the sweep signal, plus an amplitude scaling to compensate for the different energy content at various frequencies, as

$$v(t) = C \cdot \left(\frac{f_b}{f_a}\right)^{-(t/T)} s(T - t). \quad (3)$$

[1] A subset of this database for one subwoofer and one microphone was already presented shortly in [10].

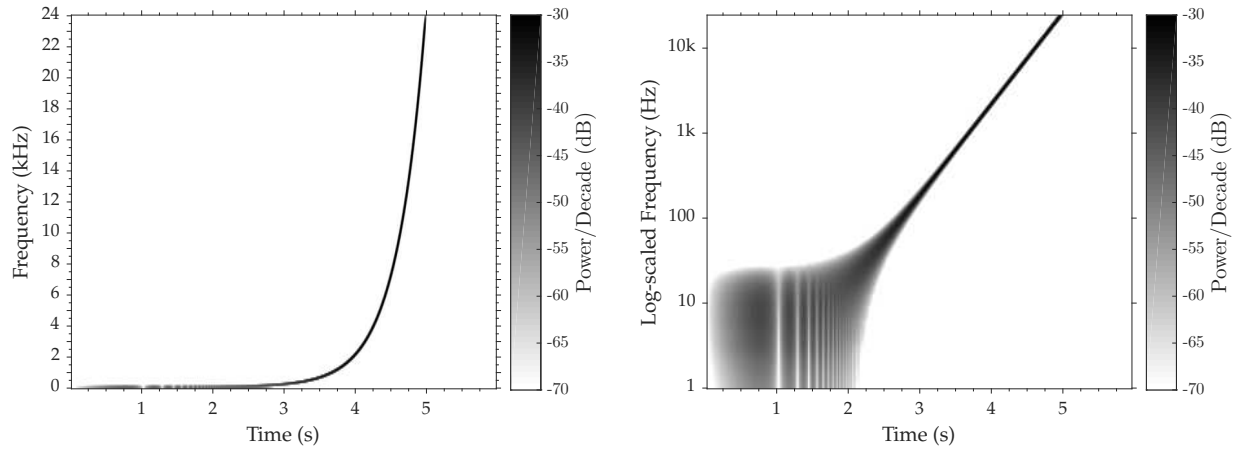


Fig. 1. The spectrogram of the sweep signal in a linear frequency scale (left) and in a logarithmic frequency scale (right). In both plots, the power resolution is linear.

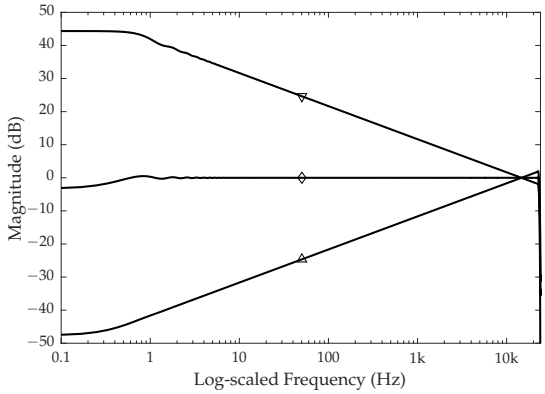


Fig. 2. The magnitude responses of the sweep signal ($\triangledown$), of the inverse signal ($\triangle$) and of the linear convolution between the two ($\diamond$).

Here, C is a normalization constant, modified from [23] to include start and stop frequencies different from 0 and the Nyquist frequency, respectively, as

$$C = \frac{2 f_b \ln(f_b/f_a)}{(f_b - f_a)\, T}. \tag{4}$$

The excitation signal used in the measurements presented in this report is the sweep signal defined in (2), with start frequency $f_a = 0.1\,\text{Hz}$ and stop frequency $f_b = f_s/2$, where $f_s = 48\,\text{kHz}$ is the sampling frequency. The duration of the sweep signal was set to $T = 5\,\text{s}$, followed by one second of silence, to ensure that the reverberant tail in the recorded signal has faded out.

The beginning and the end of the excitation signal is usually smoothed out using a tapering window in order to force the sweep to start and stop with zero phase, thus avoiding switching noise. In this way, ringing and ripples effects are reduced, at the expense of a slight deviation from the desired magnitude spectrum [20]. The tapering window used consisted of two ramp functions of length 1000 samples. The one at the beginning of the sweep signal was defined as a quarter of a cycle of a sinusoidal function (as suggested in [20]), while the one at the end of the sweep signal was a linear ramp function.

The spectrogram of the sweep signal is given in Figure 1 (using the `spgrambw` function included in the `voicebox` toolbox [28]), while the magnitude responses of the sweep signal, of the inverse signal and of the result of the convolution of the two is shown in Figure 2. From the latter, a slight deviation from the ideal uniformly flat magnitude response can be noticed. This effect is due to the tapering window and is only noticeable below 5 Hz, i.e. outside the frequency range of the subwoofers. The code for generating the sweep signal and its inverse was adapted from the code provided in [3].

The main sources of error in measuring RIRs are the presence of ambient noise, the nonlinear distortions caused by the loudspeaker, and the time-variance of the acoustic system due to changes in the room temperature or in the position of people. The ESS technique is known to be robust in tackling these issues [22, 29]. According to (1), the IF grows faster as time advances, with the result that the excitation signal has a magnitude spectrum with a pink characteristic (-3 dB/octave). High SNR can be achieved because also the ambient noise normally has a spectrum with a pink characteristic, rather than white.

The ESS technique is also quite robust against impulsive noise, provided that the impulsive event does not occur towards the end or just after the sweep signal [22]. Indeed, the time-frequency correspondence of the sweep signal guarantees that at time t all the spectral components with frequency above the IF of the sweep are shifted before the causal RIR after convolution [19, 20, 22, 29].

The same principle explains the ability of the ESS technique to partially reject regular nonlinear harmonic distortions caused by the loudspeaker when driven beyond its linear operating range [14]; each order of distortion creates a sweep with IF proportional to its order, e.g. the second-order distortion has IF increasing twice as fast as the IF of the sweep signal. It follows that the linear convolution with the inverse signal pulls back these distortions into the noncausal part of the RIR. However, this is not true for all harmonic distortion artifacts; each order of distortion also creates sweeps with IF proportional to submultiples of its order, which means that odd-order distortions produce artifacts with the same IF as the sweep signal, that overlap with the causal part of the retrieved RIR. The same arguments are valid for irregular distortions caused by defects [13], such as rub & buzz; the ESS technique is able to reject all the distortions with IF above the IF of the sweep.

A final consideration pertains to the sensitivity of the measurement technique to the time-variance of the acoustic system. This is important because a better measurement SNR can be achieved by synchronous averaging of multiple measurements recorded for the same source-receiver position pair [19–22]. It was shown in [22] that the ESS technique is more robust to time variations, compared to other techniques, and that an improvement of the SNR of 3 dB can be obtained by doubling the number of measurements (or alternatively the duration of the sweep signal). In addition, the time variance is more prominent at high frequencies, so that the SNR of the retrieved RIR at LFs can be increased by synchronous averaging of multiple measurements without introducing significant errors [22].

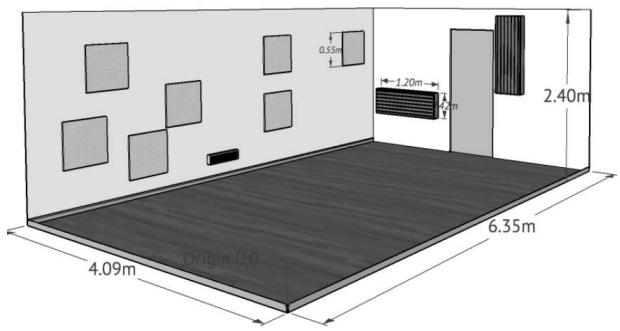


Fig. 3. A sketch of the room at B&O headquarters, Struer, Denmark.

f_n (Hz)	n_x	n_y	n_z	f_n (Hz)	n_x	n_y	n_z
0	0	0	0	83.91	2	0	0
27.02	0	1	0	87.19	1	1	1
41.95	1	0	0	88.16	2	2	1
49.91	1	1	0	89.63	0	2	1
54.05	0	2	0	91.28	1	3	0
68.42	1	2	0	98.96	1	2	1
71.50	0	0	1	99.81	2	2	0
76.44	0	1	1	108.09	0	4	0
81.07	0	3	0	108.10	0	3	1
82.90	1	0	1	110.24	2	0	1

Table 1. The theoretical value of the eigenfrequencies, with the corresponding mode index numbers [8].

q	x	y	z	p	x	y	z
1	1.12	1.56	1.50	1	3.84	3.84	0.53
2	0.77	4.04	1.80	2	2.90	0.80	0.53
3	2.04	2.47	0.90	3	3.63	5.83	0.53
4	1.62	5.32	0.60	4	2.35	4.55	1.13
5	3.05	3.06	1.50				
6	3.09	5.07	1.00				

Table 2. Source-receiver positions. The source position corresponds to the center of the subwoofer cone.

2 MEASUREMENT SETUP

2.1 Room description

The measurements were conducted in an empty small-sized room, aiming to model a typical domestic listening environment. The room dimensions were 4.09 m L $\times$ 6.35 m W $\times$ 2.40 m H, which satisfy the IEC 60268-13 specifications [30] and ensure a reasonably uniform distribution of low-frequency room modes. The theoretical values of the central frequencies of the first 20 room modes are given in Table 1 [8]. The structure is based on a brick construction comprising of lightly plastered painted walls, a wooden acoustic floating floor, and a wooden suspended false ceiling filled with absorptive material. The IEC 60268-13 standard requires the room to be filled with ordinary room furnishings, semi-covered floor and reflective roof to achieve a certain degree of diffusion and absorption and meet a 'typical' RT (e.g. $\mathrm{RT}_{200\mathrm{Hz}-4\mathrm{kHz}} = 0.3 - 0.6\,\mathrm{s}$). During the measurements described here, the room was empty but included a total of 16 high-frequency acoustic panels (8 panels on each side wall), measuring $0.5 \times 0.5 \times 0.025$ m each, and 2 Helmholtz absorbers ($1.20 \times 0.42 \times 0.13$ m) with resonance frequency 200 Hz and 300 Hz, attached on the rear wall. A sketch of the room is given in Figure 3. The air conditioning was kept off to limit possible low-frequency noise, but the room temperature was kept monitored at 21 °C ($\pm$1°C).

2.2 Measurement equipment

Two types of subwoofers were used as sound sources. The first, denoted here as Subwoofer A, was a purpose-made loudspeaker based on a closed-box design (Genelec 1094), comprising of an 18″ driver in a rigid wooden cabinet (V$\approx$168 ℓ) and capable of reproducing frequencies well below 20 Hz (-6 dB_{SPL} at 14 Hz, based on near-field measurements described below). The second, denoted here as Subwoofer B, was a Genelec 7050B comprising of an 8″ driver in a spiral bass reflex design and a metallic cylindrical cabinet, having a high-pass filter with cut-off frequency

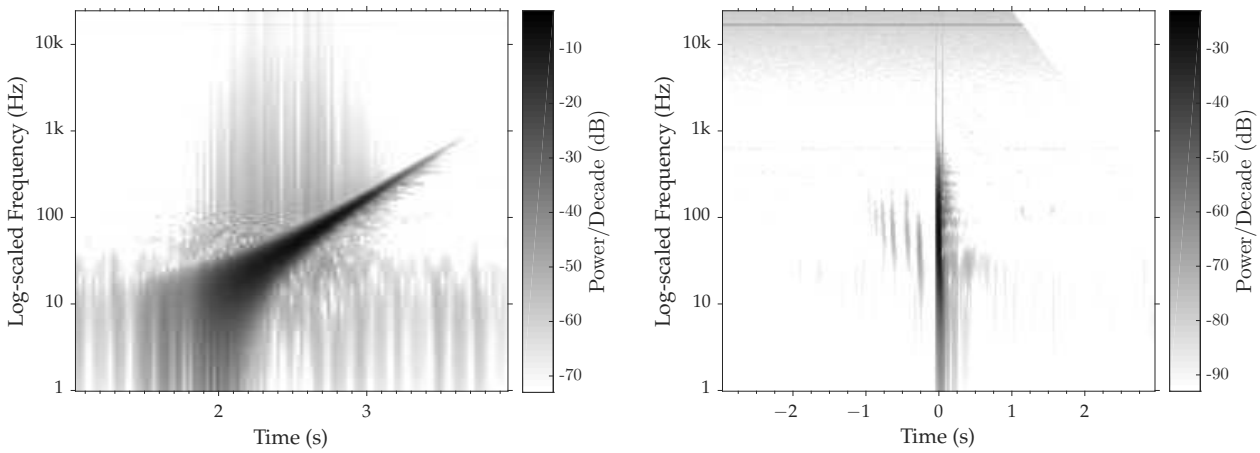


Fig. 4. The spectrogram of the near-field recording $S_4^A M_{\mathrm{NF}}^C R_1$ (left) and of the retrieved RIR (right). Notice the rub & buzz distortions above the sweep signal in the left plot, and the harmonic nonlinear distortions in the anti-causal part of the RIR in the right plot.

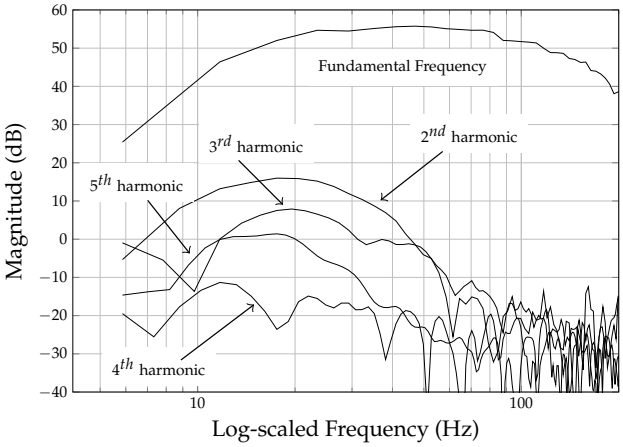


Fig. 5. The harmonic distortion magnitude response for subwoofer A up to the fifth order.

of 25 Hz and a low-pass filter with cut-off frequency set at 120 Hz [31].

The responses were recorded by two microphones connected to a B&K 2669 preamplifier and a B&K NEXUS 2690-A conditioner. The first microphone, denoted here as Microphone C, was a B&K 4939 (1/4″), with a 0° incidence frequency range from 4 Hz to 100 kHz (±2 dB), thermal noise level of 28 dBA and sensitivity of 4 mV/Pa. The second microphone, denoted here as Microphone D, was a B&K 4133 (1/2″), with a 0° incidence frequency range from 4 Hz to 40 kHz (±2 dB), thermal noise level of 20 dBA and sensitivity of 12.5 mV/Pa. Microphones and subwoofers were connected to an RME UCX audio interface. No signal processing was enabled within the signal chain.

A total of 96 RIRs were measured in the room using the two subwoofers and the two omnidirectional microphones. Each subwoofer was placed at four positions in the room and measured at six microphone positions, completing a set of 24 source-receiver combinations, in conformity with ISO 3382-2 [24] for precision measurements. The source-receiver positions are summarized in Table 2. The notation $S_p^s M_q^m R_r$ will be used to refer to a particular recorded signal, with $s = \{A, B\}$ indicating the two subwoofers and $m = \{C, D\}$ indicating the two microphones, $p = \{1, \ldots, 4\}$ and $q = \{1, \ldots, 6\}$ indicating the source and receiver positions, respectively (see Table 2), and $r = \{1, \ldots, 10\}$ indicating the number of a particular recording.

2.3 Near-field and calibration measurements

In general, measuring the free-field response of a LF source requires rooms with very large dimensions. Keele [32] suggested that such measurements could be realized within a non-anechoic environment, by placing the receiver at a point of maximum pressure i.e. at the apex of the driver. The near-field measurements presented here were performed for subwoofer A placed at position $p = 4$ (see Table 2) with the microphone capsule placed at a distance of 5 mm on axis from the driver's cone at maximal outward displacement, as recommended in [32]. For subwoofer B, information is provided by the manufacturer.

Figure 4 shows the spectrogram of the near-field recording and of the retrieved RIR. In the spectrogram on the recorder signal, impulsive noise can be seen above the sweep. This artifact, which is not visible in the retrieved RIR, is often referred as rub & buzz distortion and is likely generated by the voice coil periodically beating some internal parts of the speaker, such as connection wires, loose particles or other defects [13, 14]. These distortions have a low level compared to the recorded sweep signal, approximately -50 dB below the peak of the signal , and will be either

shifted in the non-causal part of the RIR or made not visible in the spectrogram of the retrieved RIR by the presence of the room resonances. It should be noticed that, being these types of distortion deterministic, averaging over multiple measurements will not decrease their level [13, 14].

Harmonic regular nonlinear distortions cannot be easily noticed in the spectrogram of the recorded signal, but become visible in the spectrogram of the retrieved RIR in the right plot of Figure 4; distortions at least up to the fifth order appear in the anti-causal part of the RIR. The level of the harmonic distortions is reported in Figure 5, where the magnitude response of the linear component and of the first four higher harmonics are depicted on a logarithmic frequency scale. Notice that the harmonic distortions are more prominent between 10 and 50 Hz, and tend to decay at higher frequencies. What is recorded above 90 Hz is practically ambient noise (the measured SNR was around 70 dB). A similar plot for Subwoofer B is provided in [31].

The microphones were calibrated with a B&K 4231. The output level of each subwoofer was then adjusted so that the sound level at 0.50 m was equal for the two subwoofers (56 dBC_{RMS} / peak 70 dB $SPL_{at\,53\,Hz}$)[2] when placed at the center of the room. Some of these calibration measurements are included in the database for reference.

3 MEASUREMENT ANALYSIS AND POST-PROCESSING

3.1 Recorded signals

For each source-receiver position pair, 10 recordings were performed sequentially. The analysis of the recorded signals is important to detect possible issues and assess the quality of the measurements.

Figure 6 shows the spectrograms of the first recordings for the position pair $(p,q) = (4,6)$ and for the four combinations of subwoofers and microphones. The following considerations apply in general for the other recordings and for the other source-receiver position pairs. The sweep signal is only partially reproduced, according to the frequency range of the subwoofer response (see Section 2). In comparison with the synthesized sweep signal in the right plot of Figure 1 or with the near-field measurement in Figure 4, it can be noticed how the recorded sweep is smeared out in time due to reverberation; in particular, from these plots we can expect a strong resonance between 20 and 30 Hz, corresponding to the first axial room mode (see Table 1). In these plots, all the difficulties inherent to LF-RIR measurements discussed earlier are visible. First, the LF ambient noise and the pink characteristic of its spectrum are evident. Second, irregular nonlinear distortion artifacts (or rub & buzz) for both subwoofers can be observed above the recorded sweep signal, as discussed for the near-field measurements (cfr. Section 2.3 and Figure 4). Finally, a steady component appearing in all measurements at 16 kHz can be observed in Figure 6. This disturbance, which is well above the frequency region of interest, was generated by a power adapter of one of the devices used for the measurements. From the comparison between different combinations of subwoofer and microphone, it can be seen how the 1/2″ microphone (M^D) (plots on the right in Figure 6) provides a lower noise level ($\approx$ 5 dB difference), which is in agreement with specifications (see Section 2).

Figure 7 shows the magnitude response of recordings for the source-receiver position pair $(p,q) = (4,6)$ with microphone D. It is clear that subwoofer A has a larger operational frequency range than subwoofer B. In particular, subwoofer A is able to partially excite the cavity mode (left plot); subwoofer B, on the other hand, has a frequency range between 25 Hz and 120 Hz (center plot). The same plot shows the presence of LF noise, which is not visible due to the cavity modal resonance in the left plot. Strong noise components are present at very LFs and have a harmonic structure, with fundamental frequency at 3.7 Hz (see right plot); as these components occur below the operating range of the subwoofer, they are unlikely related to the nonlinearities of the subwoofer, and are probably due to some external disturbance. Regarding the rub & buzz distortion artifacts noticed in Figure 6, their characteristic impulsive nature does not allow them to be seen in the magnitude response, since they mix up with the ambient noise. According to Klippel [13,14], these types of distortion would produce a harmonic spectrum if driven with a constant tone, which is not the case for a sweep with time-varying IF like the ESS sweep signal.

3.2 Retrieved RIRs

The linear convolution necessary to retrieve the RIR is performed in the frequency domain by multiplying the Discrete Fourier Transform (DFT) of the recorded signal and of the inverse signal, computed with a DFT size equal to twice the number of samples of the signals ($2(T+1)f_s$), and then performing an inverse DFT. Figure 8 shows the spectrograms of the RIRs retrieved from the signals recorded at source-receiver position pair $(p,q) = (4,6)$ using microphone D only (see right column of Figure 6). Compared to the spectrograms of the recorded signals, the LF noise in the retrieved RIRs is significantly reduced, as a consequence of the higher SNR achieved with the ESS technique at LFs. On the other hand, the ambient noise at high frequencies is amplified in the retrieved RIRs, as well as the 16 kHz steady component;

[2] C-weighted RMS value obtained by reproducing pink noise at equal output level as the sine-sweep. Peak SPL obtained by reproducing sine-sweeps.

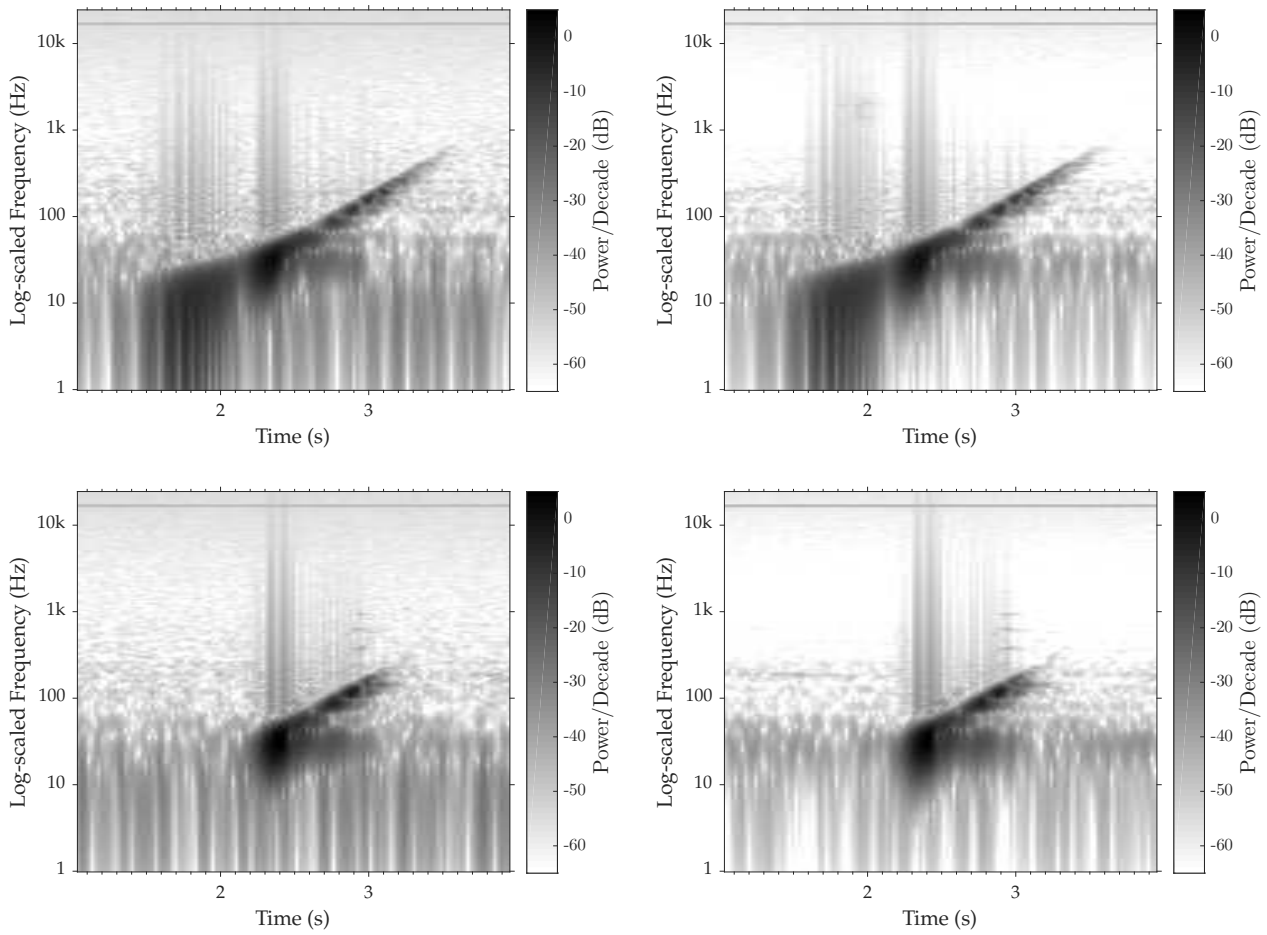


Fig. 6. The spectrogram of the recorded signals $S_4^A M_6^C R_1$ (top left), $S_4^A M_6^D R_1$ (top right), $S_4^B M_6^C R_1$ (bottom left), and $S_4^B M_6^D R_1$ (bottom right). Notice the differences in the frequency response of the two subwoofers (top vs. bottom) and in the level of the ambient noise (left vs. right), and the steady component at 16 kHz. Also notice the wide power range.

this is probably due to the fact that the ambient noise spectrum is not exactly pink.

Another effect is visible in these spectrograms; an impulsive event appears in both cases as a downward slanted line starting in the anti-causal part of the response, likely to be attributed to a strong occurrence of the rub & buzz distortion. It is not clear if the impulsive event affects the linear causal part as well, its level being close to the ambient noise level. The same can be said for regular harmonic nonlinear distortions, which are not clearly distinguishable from the background noise (except for a 2nd harmonic appearing in the bottom plot). Finally, well-separated room resonances with long decay are particularly noticeable as a smearing in time of the response in the causal part.

3.2.1 Post-processing

In order to limit the presence of nonlinear distortions, a relatively low sound level of the subwoofer has been set (see Section 2.3). As a consequence, the SNR of the RIRs retrieved from a single recording is not very high. In order to increase the SNR. the following post-processing operations are suggested. First, it is strongly recommended to perform a synchronous averaging over the RIRs retrieved from different recordings for a given source-receiver position pair and for a given subwoofer-microphone combination; as discussed already in Section 1, the robustness to time variations of the ESS technique, especially at LFs, allows to perform such an averaging over the different recordings, thus obtaining an SNR improvement of 3 dB per doubling of the number of realizations [20, 22]. Notice that synchronous averaging could also be performed on the recorded signals before retrieving the RIRs by linear convolution, and that an alternative would be to double the length of the sweep signal.

The ESS technique, however, has a poor noise rejection at high frequencies; a simple low-pass filtering can be applied to get rid of the high frequency noise (as well as the 16 kHz component). Finally, the non-causal part of the RIR can be discarded, if the interest is limited to the causal part only. A ready-to-use set

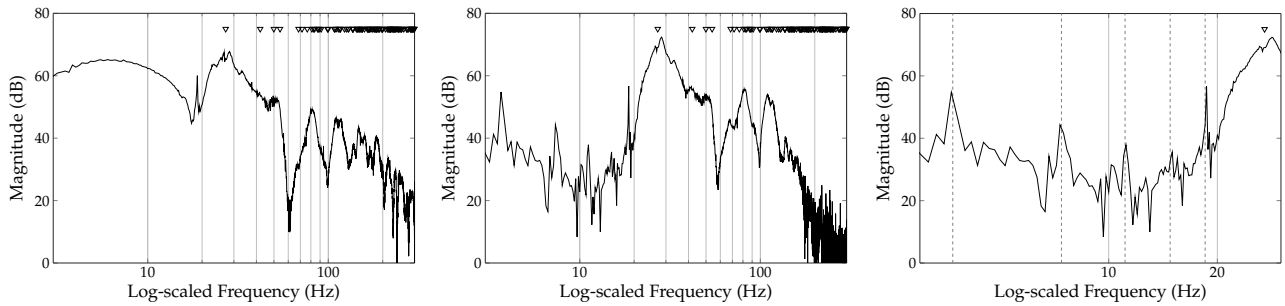


Fig. 7. The magnitude response of the recorded signals $S_4^A M_6^D R_1$ (left) and $S_4^B M_6^D R_1$ (center). The frequency range between 3 Hz and 30 Hz (right) of the latter, showing the harmonic noise component (dashed lines). In all plots, the theoretical values of the eigenfrequencies ($\triangledown$) are shown (see Table 1).

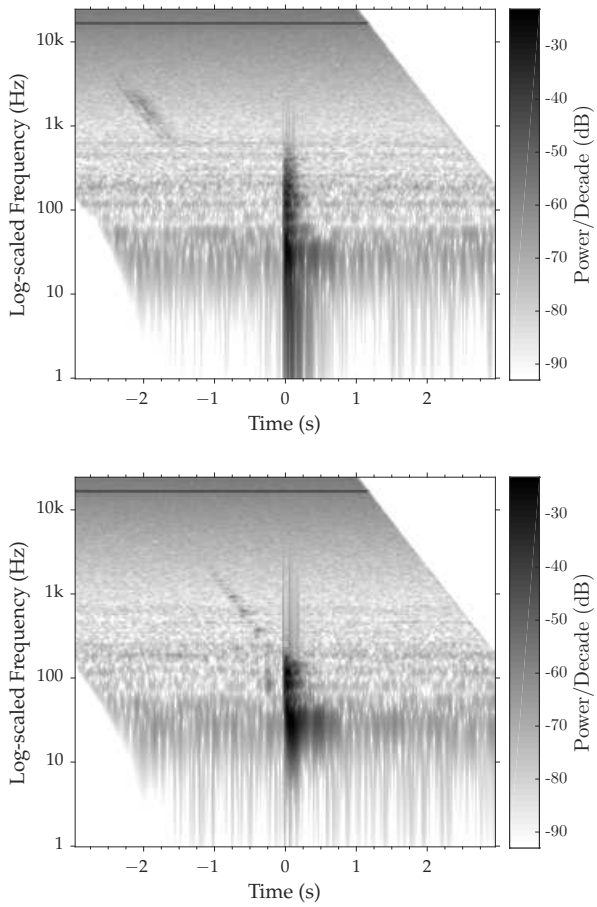


Fig. 8. The RIRs retrieved from the recorded signals $S_4^A M_6^D R_1$ (top), $S_4^B M_6^D R_1$ (bottom).

of post-processed RIRs, measured with subwoofer B and microphone D, for which a low-pass filter with cut-off frequency at 1 kHz and 100 Hz roll-off has been used, is available for download[3].

An example of the result of averaging is given in Figure 9, comparing the spectrogram and magnitude response of the RIR retrieved from a single recording (top) and after synchronous averaging over 10 recordings (bottom), with source-receiver position pair $(p,q) = (3,5)$, and with subwoofer B and microphone D. From the magnitude responses, computed over the causal part of the RIR, it can be seen how averaging is able to reduce the noise level by at least 10 dB, including the very LF disturbance already noticed in Figure 7. From the spectrograms, it can be observed how the reduction in the noise level makes the nonlinear distortions more visible; the fact that the impulsive occurrences of the rub & buzz effect are not reduced in level after averaging, is a confirmation of the deterministic nature of these events. As a consequence, great care has to be taken in the setup of the subwoofer sound level during calibration, so that nonlinear distortions are kept to a minimum. The effect of synchronous averaging can be also seen in Figure 10, showing the RIR measured at position pair $(p,q) = (3,5)$ for a single recording and after averaging over 10 recordings.

4 REVERBERATION TIME

The RT (or T_{60}) is defined as the time instant when the RIR energy decays by 60 dB from its peak value. This is usually calculated on the basis of the energy decay curve (EDC), i.e. the total amount of energy remaining in the impulse response at a given time [33]. The RT is taken as the time instant when the EDC drops below -60 dB. In most measurements, however, the noise floor level is above -60 dB and therefore this definition cannot be used in practice. In these cases, the RT is calculated using linear regression analysis and the least-squares fit procedure [24]. The decay curve is approximated by a line interpolating the EDC instead of using the EDC itself: the T_{10} is defined by interpolating the EDC between -5 and -15 dB, the T_{20} between -5 and -25 dB, and the T_{30} between -5 and -35 dB. The slope of the line interpolating the EDC within a given integration interval provides the decay rate d (in dB/s), from which an estimate of the RT is given as $-60/d$ [24]. The ISO 3382-2 standard [24] also requires the noise floor level to be at least 10 dB below the lower limit of integration, so that the the

[3]https://lirias.kuleuven.be/bitstream/123456789/572970/3/SUBRIR_SpB_MicD_RIRs.zip (password: subrir2016)

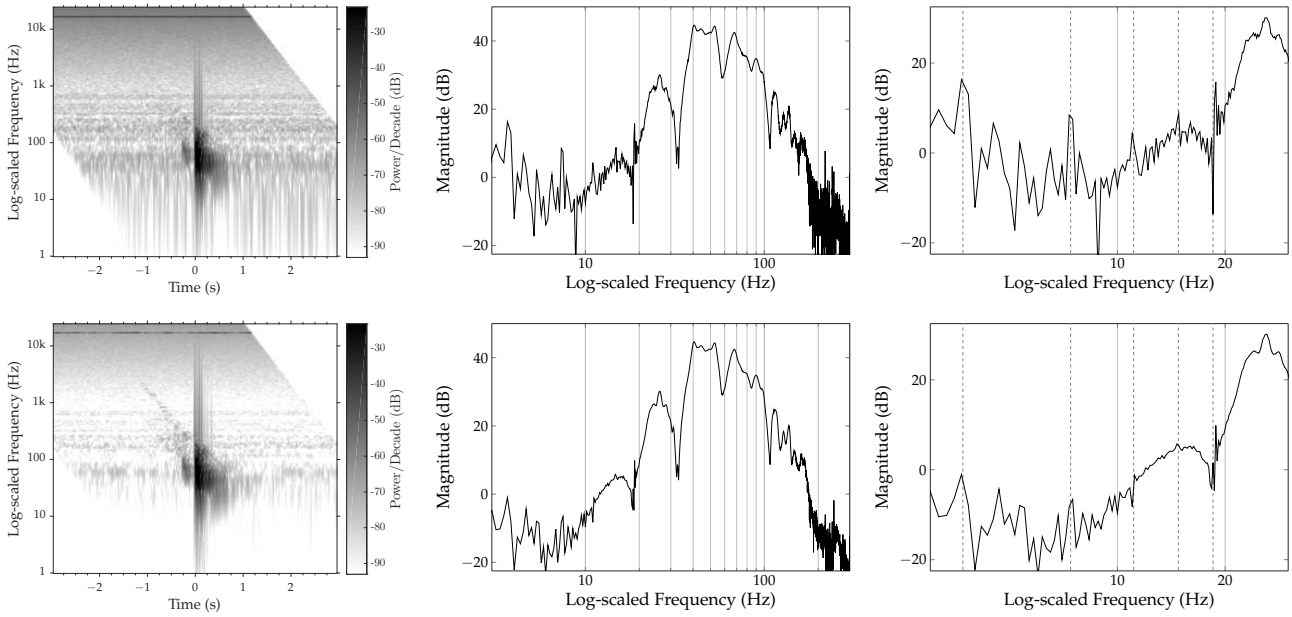


Fig. 9. Synchronous averaging. The spectrogram and the magnitude response of the RIR retrieved from a single recording $S_3^B M_5^D R_1$ (top row) and the corresponding responses after averaging over 10 recordings. The frequency range between 3 Hz and 30 Hz (right) showing the harmonic noise component (dashed lines).

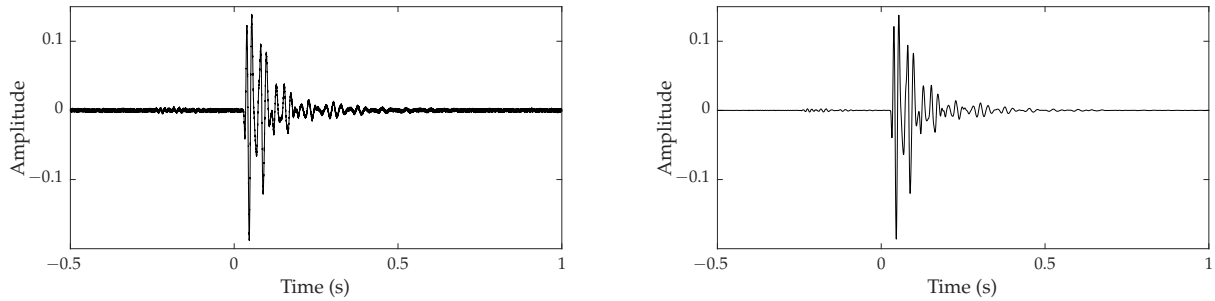


Fig. 10. The retrived RIR $S_3^B M_5^D$ before (left) and after post-processing (i.e. synchronous averaging over 10 recordings and low-pass filtering) (right).

T_{30} can be reliably estimated only for an SNR of at least 45 dB.

Frequency-dependent values of the RT are generally estimated using a bank of full-octave or one-third-octave band-pass filters [24]. Estimating the RT in subbands at very LFs is problematic. The main issues are related to low SNR, to complex modal decays (such as beating modes or double decays) [25], and to the influence of the bandpass filters of the filterbank [26]. Let us first focus on the latter. At very LFs, typical filterbanks have band-pass filters with a very narrow bandwidth, resulting in a long decay which may exceed the RT of the RIR. For instance, one-third-octave filterbanks yield a strong overestimation of the RT up to approximately 60 Hz.

In order to reduce the influence of the filters, a cosine-modulated filterbank with all filters having the same bandwidth can be used. The cosine-modulated filterbank used has 10 channels evenly distributed over the range 0 Hz to 200 Hz, and was generated with an FIR prototype filter designed using the approach in [34], with a stop-band attenuation of 60 dB. The so-obtained band-pass filters have a fixed bandwidth of 20 Hz and a decay rate of 135 ms, which is expected to be lower than the RT of the room.

Another issue is associated to the low SNR of the RIR measurements, which results in a dynamic range not sufficient for the estimation of the T_{30}. Figure 11 shows that the T_{30}^{RIR} estimate is strongly biased due to the presence of noise, while the T_{10} estimate remains largely unaffected. A conservative choice would then involve using T_{10} for all frequency bands. However, as explained later in this section, the T_{10} estimates sometimes fail to capture phenomena such as double decays and beating modes. An alternative is to visually inspect the EDCs in each frequency bands (or estimate their noise floor level) and choose the most appropriate definition of the RT in each case.

In order to overcome this issue, this paper uses an approach similar to [25]. Here, instead of calculating

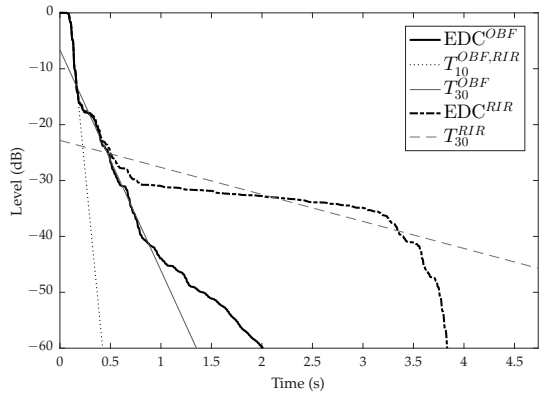


Fig. 11. The EDCs calculated from a RIR ($S_2^A M_3^D$) after post-processing and from its OBF approximation for the subband centered at 30 Hz. The interpolation lines for estimating the T_{30} and the T_{10} are also shown.

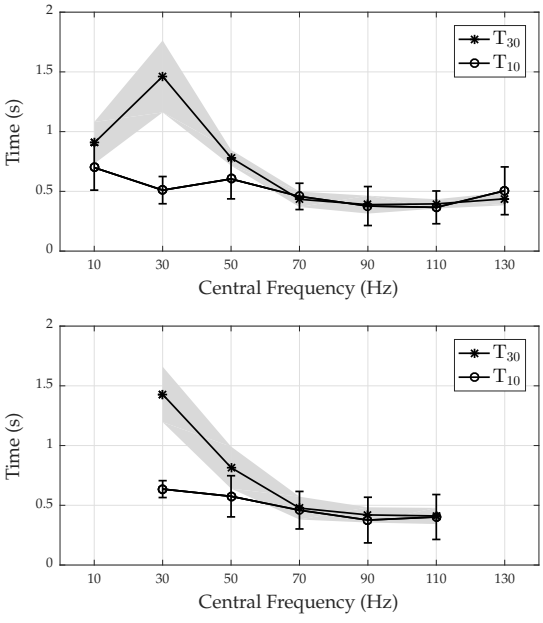


Fig. 12. The average T_{30} ($*$) and the average T_{10} ($\circ$) for subwoofer A (top) and B (bottom) estimated from the OBF approximations of the RIRs retrieved from the signals recorded using microphone D. The shaded area and the vertical lines show the standard deviation for the T_{30} and the T_{10}, respectively.

the RT of the noisy RIR directly, it is calculated based on a best-fitting noiseless parametric room model. More specifically, the RIRs of the database are first approximated by an orthonormal basis function (OBF) model [27], which provides a representation of a RIR as a linear combination of resonant responses. The model parameter values are estimated using the OBF-GMP algorithm described in [10], which is a scalable greedy algorithm with no limitations in the model order. The number of resonances used in the approximation was set to 70, which provided an accurate approximation (average normalized mean square error of -37 dB) without overfitting. This resulted in a nearly noiseless representation of the RIRs, as shown in Figure 11. The figure shows the EDCs of a post-processed RIR and of its OBF approximation for the subband centered at 30 Hz. Here, it is clear that the T_{30} value (which is calculated by interpolating the EDC between -5 and -35 dB) greatly overestimate the RT. On the other hand, the value obtained from the EDC of the OBF approximations is largely unaffected by noise. Notice also that the T_{10} is correctly estimated in both cases, as shown in Figure 11, with the two interpolating lines for the T_{10} overlapping.

Figure 12 shows the average RT values in each subband estimated from the OBF approximation of the RIRs retrieved from the signals recorded with microphone D (for microphone C, similar curves are obtained). Only the subbands centered within the limits of the frequency response of the subwoofers are considered. It can be seen that, while the T_{30} is around 400 ms above 75 Hz, it has much higher values at very LFs. This is probably due to the fact that the first axial mode, the one with theoretical frequency at 27 Hz, is very prominent. The influence of this mode can be clearly seen in both plots of Figure 12 in the T_{30} curve, where the highest values for the RT correspond to the band centered at 30 Hz. The T_{10} is also of interest in the modal region, where the low modal density gives rise to double decays and fluctuations due to beating modes [25]. A particularly large difference between the two decay rates is observed in Figure 11 for the frequency band around 30 Hz, and this is the reason why the T_{10} fails to capture the room resonant behavior in that region, as indicated in Figure 12.

5 CONCLUSION

A new RIR database measured with subwoofers as sound sources has been introduced, filling the gap of available acoustic measurements at LFs. Common difficulties in performing acoustical measurements at LFs have been addressed. The main issues proved to be a prominent LF ambient noise and the presence of impulsive irregular nonlinear distortions due to defects of the subwoofer (rub & buzz).

The ESS technique has been chosen to estimate the RIRs, due to its robustness to nonlinear distortions and its capability of providing a higher SNR at LFs. However, not all distortions can be isolated using the ESS technique, with impulsive distortions and odd-order harmonic distortions partially overlapping with the causal RIR. For this reason, near-field and calibration measurements become important to verify the nonlinear behavior of the subwoofer and to set the subwoofer level accordingly, so as to avoid distortion artifacts or at least to reduce them to an acceptable level.

Synchronous averaging of the recordings for the same source-receiver position pair is also recommended, since it allows to achieve an SNR increase of 3 dB for each doubling of the number of recordings. The

same increase can be achieved by doubling the length of the sweep signal, but with an increased risk of impulsive events occurring during the sweep.

Common difficulties in estimating the frequency-dependent RT at very LFs have been also addressed. The influence of the band-pass filters has been reduced by using a fixed-bandwidth cosine-modulated filterbank, while the problem of low SNR has been tackled by estimating the RT from a noiseless approximation of the RIRs obtained with OBF models.

The SUBRIR database is available for download[4] and it is expected to find application in the testing of acoustic signal enhancement algorithms intended for music reproduction and in the validation of physical models for room acoustics. The database has already been used in the validation of algorithms for multichannel room acoustic system identification with fixed-pole adaptive digital filters [10, 35–37].

6 ACKNOWLEDGMENT

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of (i) the FP7-PEOPLE Marie Curie Initial Training Network 'Dereverberation and Reverberation of Audio, Music, and Speech (DREAMS)', funded by the European Commission under Grant Agreement no. 316969, (ii) KU Leuven Research Council CoE PFV/10/002 (OPTEC), (iii) KU Leuven Impulsfonds IMP/14/037, and (iv) was supported by a Postdoctoral Fellowship (F+/14/045) of the KU Leuven Research Fund. The authors would like to thank Bang & Olufsen A/S for the use of their premises and equipment. The scientific responsibility is assumed by its authors.

7 REFERENCES

[1] J. Y. Wen, N. D. Gaubitch, E. A. Habets, T. Myatt, and P. A. Naylor, "Evaluation of speech dereverberation algorithms using the MARDY database," in *Proc. Int. Workshop Acoust. Signal Enhancement (IWAENC 2006)*, Paris, France, 2006.

[2] E. Hadad, F. Heese, P. Vary, and S. Gannot, "Multichannel audio database in various acoustic environments," in *Proc. Int. Workshop Acoust. Signal Enhancement (IWAENC 2014)*, Antibes-Juan Les Pins, 2014, pp. 313–317. https://doi.org/10.1109/iwaenc.2014.6954309

[3] J. K. Nielsen, J. R. Jensen, S. H. Jensen, and M. G. Christensen, "The single- and multichannel audio recordings database (SMARD)," in *Proc. Int. Workshop Acoust. Signal Enhancement (IWAENC 2014)*, Antibes-Juan Les Pins, France, 2014, pp. 40–44. https://doi.org/10.1109/iwaenc.2014.6953334

[4] J. Eaton, N. D. Gaubitch, A. H. Moore, and P. A. Naylor, "The ACE challenge - corpus description and performance evaluation," in *Proc. 2015 IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA 2015)*. IEEE, 2015, pp. 1–5. https://doi.org/10.1109/TASLP.2016.2577502

[5] R. Stewart and M. B. Sandler, "Database of omnidirectional and B-format room impulse responses," in *Proc. 2010 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP 2010)*, Dallas, USA, 2010, pp. 165–168. https://doi.org/10.1109/icassp.2010.5496083

[6] M. Jeub, M. Schäfer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *Proc. Int. 16th Conf. Digital Signal Process.*, Santorini, Greece, 2009, pp. 1–5. https://doi.org/10.1109/icdsp.2009.5201259

[7] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses," *EURASIP J. Adv. Signal Process.*, vol. 2009, p. 6, 2009. https://doi.org/10.1155/2009/298605

[8] H. Kuttruff, *Room acoustics*. Spon Press, 2009.

[9] S. E. Olive, P. L. Schuck, J. G. Ryan, S. L. Sally, and M. E. Bonneville, "The detection thresholds of resonances at low frequencies," *J. Audio Eng. Soc.*, vol. 45, no. 3, pp. 116–128, 1997. http://www.aes.org/e-lib/browse.cfm?elib=7868

[10] G. Vairetti, E. De Sena, T. van Waterschoot, M. Moonen, M. Catrysse, N. Kaplanis, and S. H. Jensen, "A physically motivated parametric model for compact representation of room impulse responses based on orthonormal basis functions," in *Proc. 10th Eur. Congr. Expo. Noise Control Eng. (EURONOISE 2015)*, Maastricht, The Netherlands, 2015.

[11] W. Klippel, "Tutorial: Loudspeaker nonlinearities - causes, parameters, symptoms," *J. Audio Eng. Soc.*, vol. 54, no. 10, pp. 907–939, 2006. http://www.aes.org/e-lib/browse.cfm?elib=13881

[12] W. Klippel and R. Werner, "Loudspeaker distortion - measurement and perception, part 1: Regular distortion defined by design," in 26th *Tonmeistertagung*, Leipzig, Germany, 2010.

[13] ——, "Loudspeaker distortion - measurement and perception, part 2: Irregular distortion caused by defects," in 26th *Tonmeistertagung*, Leipzig, Germany, 2010.

[14] W. Klippel and U. Seidel, "Measurement of impulsive distortion, rub and buzz and other disturbances," in *Preprints AES* 114th *Conv.*, Amsterdam, The Netherlands, 2003. http://www.aes.org/e-lib/browse.cfm?elib=12550

[15] W. Klippel, "Rub and buzz and other irregular loudspeaker distortion (tutorial)," in *Preprints AES* 134th *Conv.*, Rome, Italy, 2013.

[16] R. C. Heyser, "Acoustical measurements by time delay spectrometry," *J. Audio Eng. Soc.*, vol. 15, no. 4, pp. 370–382, 1967. https://doi.org/10.1121/1.2020540

[17] A. J. Berkhout, D. de Vries, and M. M. Boone, "A new method to acquire impulse responses in

[4] https://lirias.kuleuven.be/handle/123456789/572970 (password: subrir2016)

concert halls," *J. Acous. Soc. Am.*, vol. 68, no. 1, pp. 179–183, 1980. https://doi.org/10.1121/1.384618

[18] P. M. Clarkson, J. Mourjopoulos, and J. Hammond, "Spectral, phase, and transient equalization for audio systems," *J. Audio Eng. Soc.*, vol. 33, no. 3, pp. 127–132, 1985. http://www.aes.org/e-lib/browse.cfm?elib=4461

[19] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Preprints AES* 108th *Conv.*, Paris, France, 2000. http://www.aes.org/e-lib/browse.cfm?elib=10211

[20] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *J. Audio Eng. Soc.*, vol. 49, no. 6, pp. 443–471, 2001. http://www.aes.org/e-lib/browse.cfm?elib=10189

[21] G.-B. Stan, J.-J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," *J. Audio Eng. Soc.*, vol. 50, no. 4, pp. 249–262, 2002. http://www.aes.org/e-lib/browse.cfm?elib=11083

[22] A. Torras-Rosell and F. Jacobsen, "Measuring long impulse responses with pseudorandom sequences and sweep signals," in *Proc. 39th Int. Congr. Noise Control Eng. (INTER-NOISE 2010)*, Lisbon, Portugal, 2010.

[23] M. Holters, T. Corbach, and U. Zölzer, "Impulse response measurement techniques and their applicability in the real world," in *Proc. 12th Int. Conf. Digital Audio Effects (DAFx 2009)*, Como, Italy, 2009. http://dafx09.como.polimi.it/proceedings/

[24] ISO 3382-2:2008, "Acoustics - measurements of room acoustic parameters - part 2: Reverberation time in ordinary rooms," 2008. https://doi.org/10.3403/30081124

[25] M. Karjalainen, P. Ansalo, A. Mäkivirta, T. Peltonen, and V. Välimäki, "Estimation of modal decay parameters from noisy response measurements," *J. Audio Eng. Soc.*, vol. 50, no. 11, pp. 867–878, 2002. http://www.aes.org/e-lib/browse.cfm?elib=11059

[26] F. Jacobsen, "A note on acoustic decay measurements," *J. Sound Vibration*, vol. 115, no. 1, pp. 163–170, 1987. https://doi.org/10.1016/0022-460X(87)90497-4

[27] P. Heuberger, P. van den Hof, and B. Wahlberg, *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer, 2005. https://doi.org/10.1007/1-84628-178-4

[28] M. Brookes, "VOICEBOX: A speech processing toolbox for MATLAB." Imperial College, Software Library, 2011. http://www.ee.imperial.ac.uk/hp/staff/dmb/voicebox/voicebox.html

[29] A. Torras-Rosell and F. Jacobsen, "A new interpretation of distortion artifacts in sweep measurements," *J. Audio. Eng. Soc.*, vol. 59, no. 5, pp. 283–289, 2011. http://www.aes.org/e-lib/browse.cfm?elib=15929

[30] IEC 60268-13:1998, "Sound system equipment - part 13: Listening tests on loudspeakers," 1998. https://doi.org/10.3403/01396801

[31] *Genelec 7050B Active Subwoofer - Operating manual*, Genelec Oy, 2005, D0061R001. http://www.genelec.com/

[32] D. Keele Jr., "Low-frequency loudspeaker assessment by nearfield sound-pressure measurement," *J. Audio Eng. Soc.*, vol. 22, no. 3, pp. 154–162, 1974. http://www.aes.org/e-lib/browse.cfm?elib=2774

[33] M. R. Schroeder, "New method of measuring reverberation time," *J. Acoust. Soc. Am.*, vol. 37, no. 3, pp. 409–412, 1965. http://doi.org/10.1121/1.1909343

[34] Y.-P. Lin and P. Vaidyanathan, "A Kaiser window approach for the design of prototype filters of cosine modulated filterbanks," *IEEE Signal Proc. Letters*, vol. 5, no. 6, pp. 132–134, 1998. http://doi.org/10.1109/97.681427

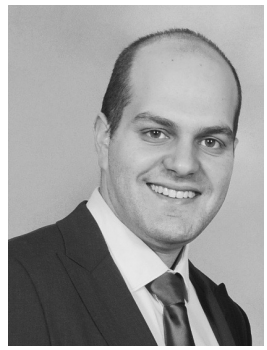
[35] G. Vairetti, E. De Sena, M. Catrysse, S. H. Jensen, M. Moonen, and T. van Waterschoot, "Room acoustic system identification using orthonormal basis function models," in *Proc. 60th Int. Conf. Audio Eng. Soc.* Leuven, Belgium: AES, 2016. http://www.aes.org/e-lib/browse.cfm?elib=18086

[36] ——, "Multichannel identification of room acoustic systems with adaptive IIR filters based on orthonormal basis functions," in *Proc. 2016 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP 2016)*, Shanghai, China, 2016. https://doi.org/10.1109/icassp.2016.7471628

[37] ——, "A scalable algorithm for physically motivated and sparse approximation of room impulse responses with orthonormal basis functions," KU Leuven, Tech. Rep., 2016. ftp://ftp.esat.kuleuven.be/pub/stadius/gvairett/

THE AUTHORS


Giacomo Vairetti


Neofytos Kaplanis


Enzo De Sena


Søren Holdt Jensen


Søren Bech


Marc Moonen


Toon van Waterschoot

Giacomo Vairetti received the B.Sc. in 2010 and the M.Sc. (cum laude) in 2012, both in Computer Engineering at Politecnico di Milano (Italy). He was a visiting student at the Signal Processing and Acoustics Dept. of Aalto University (Finland) in 2012 and at the Electronic Systems Dept. of Aalborg University (Denmark) in 2014. He is currently pursuing a Ph.D. in Electrical Engineering at KU Leuven (Belgium), where he was a Marie Curie Fellow. His research interests are in signal processing and system identification, applied to room acoustic modeling, sound synthesis, and audio reproduction.

●

Neofytos Kaplanis holds a B.Mus. Tonmeister (Surrey, UK) and a M.Sc. in auditory neuroscience (London, UK). He has been an R&D acoustic engineer at HARMAN Automotive and a visiting researcher at University of London (2013), KU Leuven (2014), and Aalto University (2015). In 2013 he joined Bang & Olufsen as an acoustic research fellow, pursuing a Ph.D. (EE) from Aalborg University. His research interest centers upon the auditory human perception, aiming to merge physical phenomena with their underlying cognitive and perceptual properties.

●

Enzo De Sena received the B.Sc. in 2007 and M.Sc. (cum laude) in 2009, both from the Universita degli Studi di Napoli "Federico II" (Italy) in Telecommunication Engineering. In 2013, he received the Ph.D. degree in Electronic Engineering from Kings College London (UK), where he was also a Teaching Fellow from 2012 to 2013. Between 2013 and 2016 he was a Postdoctoral Research Fellow at the Katholieke Universiteit Leuven (Belgium). Since September 2016 he is a Lecturer in Audio at the Institute of Sound Recording at the University of Surrey (UK). He held visiting positions at Stanford University (USA), Aalborg University (Denmark) and Imperial College London (UK). He is a former Marie Curie Fellow. His current research interests include room acoustics modelling, surround sound, microphone beam forming and binaural modelling. For more information, see www.desena.org.

●

Søren Holdt Jensen received the M.Sc. degree in electrical engineering from Aalborg University (AAU), Denmark, in 1988, and the Ph.D. degree (in signal processing) from the Technical University of Denmark (DTU) in 1995. He is Full Professor in Signal Processing at Aalborg University. Before joining the Electronic Systems Dept. (AAU), he was with the Telecommunications Laboratory of Telecom Denmark, Ltd, Copenhagen; the Electronics Institute of Technical University of Denmark; the Scientific Computing Group of Danish Computing Center for Research and Education (UNI•C), Lyngby; the Electrical Engineering Dept. of KU Leuven, Belgium; and the Center for PersonKommunikation (CPK) of AAU. His current research interest are in statistical signal processing, numerical algorithms, optimization engineering, machine learning, and digital processing

of acoustic, audio, communication, image, multimedia, speech, and video signals. He is co-author of the textbook *Software-Defined GPS and Galileo Receiver—A Single-Frequency Approach*, Birkhäuser, Boston, USA, also translated to Chinese: National Defence Industry Press, China. Prof. Jensen has been Associate Editor for the IEEE Transactions on Signal Processing, IEEE/ACM Transactions on Audio, Speech and Language Processing, Elsevier Signal Processing, and EURASIP Journal on Advances in Signal Processing. He is a recipient of an individual European Community Marie Curie (HCM: Human Capital and Mobility) Fellowship, former Chairman of the IEEE Denmark Section and the IEEE Denmark Sections Signal Processing Chapter (founder and first chaiman). He is member of the Danish Academy of Technical Sciences (ATV) and has been member of the Danish Council for Independent Research (2011–2016) appointed by Danish Ministers of Science.

●

Søren Bech received a M.Sc. and a Ph.D. from the Acoustic Technology (AT) Dept. of the Technical University of Denmark. From 1982–92 he was research Fellow at AT studying perception and evaluation of reproduced sound in small rooms. In 1992 he joined Bang & Olufsen where he is Director of Research. In 2011 he was appointed Professor in Audio Perception at Aalborg University and he is Adjunct Professor at Surrey University (UK), and McGill University (Canada). He is a Fellow of the Acoustical Society of America and Audio Engineering Society (AES). He is past Governor and Vice-President of the AES and now serves as associate technical editor of the AES Journal. He has been vice-chair of the International Telecommunication Union working group 10/3. In 2006 he and Dr. Zacharov published the book Perceptual Audio Evaluation - Theory, Method and Application (Wiley and Sons). His research interest includes psychoacoustics and in particular human perception of reproduced sound in small and medium sized rooms. Other interests include experimental procedures and statistical analysis of data from sensory analysis of audio and video quality.

●

Marc Moonen is a Full Professor at the Electrical Engineering Dept. of KU Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL and audio signal processing. He received the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), the 2004 Alcatel Bell (Belgium) Award (with Raphael Cendrillon), and was a 1997 Laureate of the Belgium Royal Academy of Science. He received journal best paper awards from the IEEE Transactions on Signal Processing (with Geert Leus and with Daniele Giacobello) and from Elsevier Signal Processing (with Simon Doclo). He was chairman of the IEEE Benelux Signal Processing Chapter (1998-2002), a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications, and President of EURASIP (European Association for Signal Processing, 2007-2008 and 2011-2012). He has served as Editor-in-Chief for the EURASIP Journal on Applied Signal Processing (2003-2005), Area Editor for Feature Articles in IEEE Signal Processing Magazine (2012-2014), and has been a member of the editorial board of IEEE Transactions on Circuits and Systems II, IEEE Signal Processing Magazine, Integration-the VLSI Journal, EURASIP Journal on Wireless Communications and Networking, and Signal Processing. He is currently a member of the editorial board of EURASIP Journal on Advances in Signal Processing.

●

Toon van Waterschoot received the MSc (2001) and PhD (2009) degrees in Electrical Engineering, both from KU Leuven, Belgium, where he is currently a tenure-track Assistant Professor. He has previously held teaching and research positions with the Antwerp Maritime Academy, the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT), and the Research Foundation - Flanders (FWO) in Belgium, with Delft University of Technology in The Netherlands, and with the University of Lugano in Switzerland. His research interests are in signal processing, machine learning, and numerical optimization, applied to acoustic signal enhancement, acoustic modeling, audio analysis, and audio reproduction. He has been serving as an Associate Editor for the Journal of the Audio Engineering Society (AES) and for the EURASIP Journal on Audio, Music, and Speech Processing, and as a Guest Editor for Elsevier Signal Processing. He is a Member of the Board of Directors of the European Association for Signal Processing (EURASIP) and a Member of the IEEE Audio and Acoustic Signal Processing Technical Committee (AASP-TC). He was the General Chair of the 60th AES International Conference in Leuven, Belgium (2016), and has been serving on the Organizing Committee of the European Conference on Computational Optimization (EUCCO 2016) and the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2017). He is a member of EURASIP, IEEE, ASA, and AES.