

COVER PAGE

Document downloaded by @DAEL

Fri Nov 20 03:27:06 2026
For personal use

When automatic English translation is provided, only the original document is authentic.
The EAA cannot be held responsible of any translation error

Bibliographical reference

De-Noising of a Room Impulse Response by Applying Wavelets, Đorđe M. Damnjanovič, Dejan G. Čirič and Bratislav B. Predič, Acta Acustica **vol. 104** (Number 3), 2018, pp. 452-463

DOI
https://doi.org/10.3813/AAA.919187

Acta Acustica united with Acustica
Vol. 104 (2018) 452 – 463

DOI 10.3813/AAA.919187

De-Noising of a Room Impulse Response by Applying Wavelets

Đorđe M. Damnjanović[1)], Dejan G. Ćirić[2)], Bratislav B. Predić[2)]

[1)] Faculty of Technical Sciences Čačak, University of Kragujevac, Svetog Save 65, 32000 Čačak, Serbia.
djordje.damnjanovic@ftn.kg.ac.rs

[2)] Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia

Summary

One of the techniques most commonly used for de-noising of audio signals, images or speech is wavelet de-noising. However, with regard to de-noising of room impulse responses (RIRs), the usage of wavelets can be hardly found in literature. This paper explores the potentials of applying wavelets for the reduction of negative effects of noise in these responses, since noise represents a serious problem in RIRs. The analysis is based on implementing various wavelets on both synthesized and measured RIRs containing different noises. The influence of changing the parameters of wavelets is observed in order to find an optimal set of parameters that will provide the best de-noising results. The dynamic range improvement of the backward integrated RIR as well as the noise floor reduction are used as control measures to assess the wavelet de-noising quality. The results show that the noise floor can be reduced and dynamic range can be improved significantly (e.g., by more than 10 dB) by applying the optimal set of wavelet parameters. The wavelet de-noising is compared with some other noise compensation methods using the estimated reverberation time as a control measure. In addition to de-noising, wavelets seem to be promising for detection of the truncation time of a RIR.

PACS no. 43.55.Br, 43.60.Hj

1. Introduction

Signal corruption by noise represents one of the important problems present in many areas including acoustics. To overcome the noise disturbance, the usage of de-noising techniques has become common.

However, it is hard to define an optimal technique for noise removal. In an ideal scenario, a de-noising method would completely remove noise without disturbing the signal. However, this is hardly realizable in practice.

Some well-known de-noising techniques are infinite impulse response and finite impulse response filtering, adaptive and optimal filtering, Wiener filtering, wavelets, etc. [1, 2, 3, 4]. A technique widely applied in de-noising, especially of speech signals is the wavelet technique [5]. This technique has certain advantages coming from the nature of wavelets [6, 7, 8]. Different wavelet packages and different sets of parameters have been applied in de-noising. In that regard, it is very important to have some prior information about the signal to be de-noised and the wavelets that will be used. To achieve a better performance, it is necessary to adjust the setup parameters.

A similar situation regarding noise disturbance exists in acoustic measurements, or more specifically in the measurements of room impulse responses (RIRs). Noise is an inevitable component of any measured RIR and it represents one of the major weaknesses of RIRs [9, 10, 11]. Noise can cause serious problems in estimating the acoustic parameters from the measured RIRs [9, 12, 13]. As a consequence of noise, the decay curve obtained by the backward integration of a RIR – also known as the Schroeder's decay curve [14] or energy decay curve (EDC) [9] has a reduced slope from a certain point. This point is located in the vicinity of the knee where the main decay of a RIR intersects the noise floor.

Different methods have been developed to compensate the effects of noise in RIRs. Some of the ways how to handle the noise, such as truncation of a RIR at a certain (truncation) time and correction for truncation are described in the ISO 3382 standard [15]. Also, a well-known method to reduce the noise problem is the averaging of responses that gives an improved signal-to-noise ratio (SNR). It is worth pointing out that averaging is more sensitive to time-variance than the usage of a longer excitation signal (e.g., swept sine) [16, 17]. In addition, there are methods oriented towards the reduction of noise influence directly in the Schroeder's decay curve. They include subtraction of the mean square value of noise from the squared RIR before the backward integration [18].

Possibilities to implement the wavelet technique in de-noising of RIRs are investigated in this paper. To the best of the authors' knowledge, wavelets have not been used for that purpose so far. There are many types of wavelets

Received 31 May 2017,
accepted 30 March 2018.

© S. Hirzel Verlag · EAA

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

that can be applied, and several parameters that can affect the results. This is why a detailed analysis of the effects of changing wavelets and their parameters is carried out here. The main objective is to find an optimal set of wavelet parameters that would give the best results – the largest increase of the dynamic range of both the processed RIR and corresponding EDC and acceptably small disturbance of the useful signal that might be caused by de-noising. The wavelet de-noising is compared with two other compensation methods based on the RIR truncation and noise subtraction, truncation and correction for truncation (here denoted as subtraction-truncation-correction). These two methods compensate the noise effects in the EDCs, whereas the wavelet method de-noises RIRs directly. Therefore, only the wavelet de-noising could provide a potential benefit in the applications such as auralization.

One more possible usage of wavelets in the RIR processing is the estimation of truncation time located at the knee (intersection of the main RIR's decay and noise floor) and truncation of a RIR at that time. This possibility is also explored here. The truncation times of RIRs obtained by the wavelets are compared to the times estimated by the nonlinear decay model procedure recently presented in [19].

2. Background on wavelets

In the simplest form, a wavelet is a waveform of effectively limited duration that has an average value of zero [6, 20, 21]. Basis wavelet function, or so-called "mother" wavelet, can define a wavelet basis by scaling and translation process [6, 21],

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right), \quad a > 0, \tag{1}$$

where a is the scaling parameter and b is the translation parameter. If the values of the parameters are chosen correctly, the segments of the examined function may be represented with a higher or lower resolution, depending on the shape and complexity of the function [6, 21].

Decomposition of a signal is an inherent process of wavelet de-noising, where the signal is divided into subsignals according to different frequency contents. Two most common wavelets used for de-noising are Haar and Daubechies. Some other wavelets such as Coiflets, Symlets, biorthogonal and Meyer may potentially provide good results in de-noising.

In order to remove noise from a signal, wavelet coefficients including both the approximation and detail coefficients (low and high frequency components) are first obtained by the multilevel signal decomposition [6, 8]. The next step is thresholding, which refers to the comparison of each coefficient with the threshold. The purpose of this comparison is to decide whether the coefficients constitute a desirable part of the original signal or not. The wavelet thresholding can be realized in two different ways – hard and soft thresholding [6, 8].

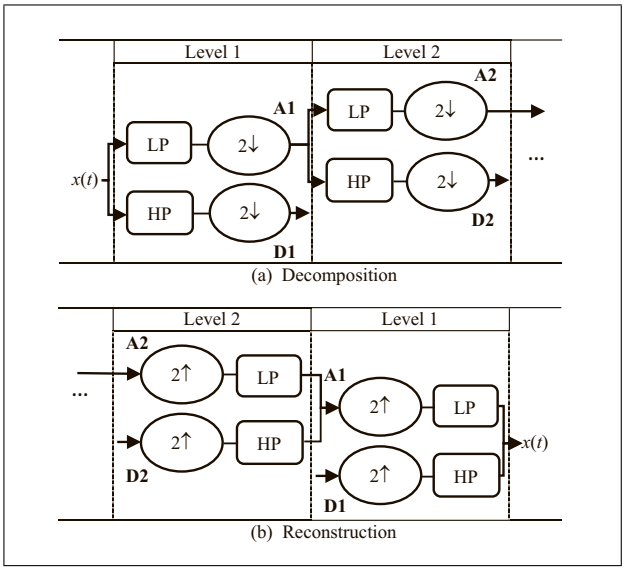


Figure 1. Block diagram of wavelet transform: (a) decomposition and (b) reconstruction.

In addition to the method (way) of thresholding, there are other parameters related to thresholding that have to be considered, such as the threshold value. Different threshold selection algorithms can be applied: minimax, universal, rigorous, sure, etc. [6, 8, 22].

The final step in wavelet de-noising is signal reconstruction based on multilevel wavelet decomposition structure. Signal components generated after decomposition have to be assembled back into the reconstructed original signal without loss of information [6]. The described procedure is presented as a block diagram in Figure 1, where LP stands for a low-pass filter, HP for a high-pass filter, A for approximation coefficients, D for detailed coefficients, $2\uparrow$ for up-sampling and $2\downarrow$ for down-sampling [5, 6].

3. Methods of analysis

The research presented here is done using various synthetized and measured RIRs. Two types of synthetized RIRs are generated. The first one, generated by the image source model (ISM) [23], contains certain irregularities, while the second one, generated by the procedure described in [24], has a truly exponential decay as in diffuse sound fields. In the RIR synthesis by the ISM, different synthesis parameters are applied such as wall reflection coefficients, reverberation time, source/receiver locations and room dimensions. One of the parameters is also the transition point from early reflections (computed according to the ISM [23]) to diffuse field (the late reflections generated as decaying random noise, see [25]). This transition is located at the point where the RIR's energy has decayed by 10 dB to 20 dB. The reverberation times calculated from the synthesized RIRs have values from 0.52 s to 1.18 s. The RIRs with exponential decays have the same reverberation times as those of the RIRs generated by the ISM. The synthesized RIRs are contaminated with white or pink noises of different levels ranging from −30 dB to

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

−70 dB in order to have the RIRs with different SNRs, that is, dynamic ranges. The noise levels are determined with reference to the maximal level of the RIR in the initial region.

Besides the synthetized RIRs, measured RIRs are also used here. These RIRs were measured in a hallway and laboratory similar to a classroom using the swept sine technique with the excitation signals of different lengths [19]. The reverberation times calculated from the broadband RIRs of the hallway and laboratory are about 1.8 s and 0.8 s, respectively. The measured RIRs have natural ambient noise. In this type of RIRs, the reference (noiseless) RIR is missing. Consequently, the RIR measured with much longer excitation signal and having a larger dynamic range is used as a reference one. Five RIRs of each of the mentioned synthesized types with various noises and noise levels (yielding more than 60 RIR-noise combinations overall) together with ten measured RIRs are used in this study.

In the simulation part of the presented research, the whole procedure of RIR measurement is simulated in the following way. The synthesized noiseless RIR is first convolved with the excitation signal (exponential swept sine) and synthetic noise is added to the result. In this way, the room response with noise is obtained. The RIR is extracted by deconvolving the room response with the inverse filter. Wavelet de-noising is applied to both the room response and the extracted RIR. Wavelets are also applied to the synthesized RIRs with the added noise directly (without convolution and deconvolution). In addition, both the measured room responses to the excitation swept sines and the extracted measured RIRs are de-noised by wavelets.

In the wavelet de-noising, the effects of changing the wavelet (function) and parameters are analyzed. In that regard, different wavelets and parameters are used. Thus, the wavelets included in the analysis belong to the following families: Haar, Daubechies, Coiflets, Symlets, biorthogonal, reverse biorthogonal and Meyer. The wavelet parameters that are changed throughout the analysis are: threshold selection rule, thresholding, threshold rescaling and decomposition level.

The de-noising results are evaluated in the following way. De-noised synthesized RIRs are compared with the corresponding noiseless and noisy RIRs. The comparison is done in both the time and frequency domain. In addition, the EDCs of de-noised synthesized RIRs are compared with the EDCs of the corresponding noiseless and noisy RIRs. In the latter comparison, special attention is paid to the dynamic range of the EDC, that is, to the improvement of this range. The dynamic range of a noisy EDC is determined here as a difference between the initial maximal level of the EDC (0 dB) and the level along the EDC where the absolute value of deviation of that particular EDC from the noiseless EDC becomes greater than the pre-defined threshold, as it is shown in Figure 2. The difference between the dynamic ranges of EDCs of a noisy RIR after and before the wavelet de-noising represents a dynamic range improvement.

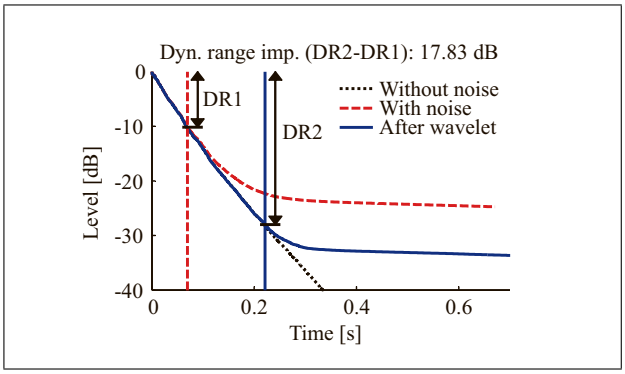


Figure 2. (Colour online) Obtaining the dynamic range improvement as a difference between the dynamic ranges of the EDCs of a noisy RIR after (DR2) and before the wavelet de-noising (DR1).

An initial study was carried out in order to select a suitable threshold value. Different thresholds were used ranging from 0.05 dB to 2 dB. It was found that the choice of threshold value affects the resulting dynamic range. Some threshold values lead to a dynamic range change that corresponds well to a visual inspection of the change of EDC dynamic range. As a result, the values of 0.2 dB and 0.3 dB are set as optimal thresholds for the synthesized and measured RIRs, respectively.

The wavelet de-noising algorithm is applied in the Matlab software package using the *wden* function. This function contains various wavelet parameters. Options for the threshold selection rule are *rigrsure*, *minimaxi*, *sqtwolog* and *heursure*. The thresholding can be *hard* or *soft* defining what happens with the coefficients above the threshold value. These coefficients are shrunk or scaled in *soft* thresholding, which is not the case in *hard* thresholding. Options for the threshold rescaling are *one*, *sln* and *mln*. Wavelet decomposition level is also a parameter of the *wden* function. Wavelet parameters are changed to take different options through the analysis.

The wavelet de-noising is applied to third-octave and octave bandpass filtered RIRs as well as broadband RIRs. The filtering is carried out using Butterworth filters that are in accordance with the IEC 61260 [26]. The reverberation time, also used as a control measure, is estimated by performing a linear regression in a range of the generated EDCs starting at the −5 dB point. The end point of the estimation range is located at the knee (decay level of a particular EDC at the knee) for the EDCs obtained using the subtraction-truncation-correction method, and at the same time point for the noiseless EDCs. The end point is positioned 10 dB above the knee (10 dB above the decay level of a particular EDC at its knee) for the EDCs of noisy RIRs and RIRs after the wavelet de-noising. The EDCs obtained using the truncation method have the end point at the decay level that corresponds to the time shifted for 50 ms from the knee towards the RIRs start in order to avoid the sudden drop of the EDC at the knee. When comparing the wavelet de-noising with the compensation methods from the literature, the truncation time used in both the trun-

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

cation and subtraction-truncation-correction method is determined by the nonlinear decay model [19]. The correction term for the subtraction-truncation-correction method is also calculated by the nonlinear decay model as done in [27].

4. Results of analysis

4.1. Effects of wavelet parameters

One of the RIRs synthesized by the ISM corrupted by synthetic white noise at a level of −55 dB, filtered in the third-octave band at 1 kHz, is used within this section to present the de-noising results. The Haar wavelet and different decomposition levels (from level 1 to level 10) are first implemented.

The illustration of the effects of using different thresholding (*soft* and *hard*) and different threshold selection rules (*sqtwolog*, *minimaxi*, *rigrsure* and *heursure*) is presented in Figure 3. According to some authors [5, 6, 8], one of the most commonly used threshold selection rules is minimax (*minimaxi* when applied in Matlab). This is why special attention is paid to this selection rule. The results for decomposition level of 5 are given in Figure 3 since it leads to the best results in this particular case. As a consequence of the nature of the algorithms [6], *rigrsure* and *heursure* threshold selection rules provide similar results.

The wavelet de-noising with the *soft* thresholding and decomposition level lower than 5 provides only a slight improvement in the EDC dynamic range (e.g., up to 1 dB or 2 dB) or no improvement at all independently of the threshold selection rule. When the decomposition level is increased to 5, the EDCs obtained after the application of such wavelets might go below the reference EDC generated from the noiseless RIR in one time range, as shown in Figure 3a. This means that the decay rate is changed especially in the late part of the decay prior to the knee (before the knee). With a higher decomposition level, the noise floor in the EDC is reduced more, but the degradation (loss) of the reverberation decay is also increased, causing a larger change – an increase of the decay rate. Consequently, the dynamic range is usually smaller than the one obtained without the wavelet de-noising. If the time domain is observed, a relatively large loss of the original signal appears, especially close to the end of the reverberation decay, that is, close to the knee. A change of wavelet function does not lead to a significant change in the dynamic range improvement.

If the *soft* and *hard* thresholdings are compared, the transition region from almost no impact of wavelets (in the reverberation decay part of the RIR) to the strongest impact of wavelets (in the noise part of the RIR) is wider in the *soft* thresholding. The degradation of the reverberation decay is larger in this region, especially for higher decomposition levels – in the case presented here, higher than 5. This is why better results are achieved using the *hard* thresholding.

For the applied decomposition level of 5, the change of the selection rule has a certain impact on the decay curve.

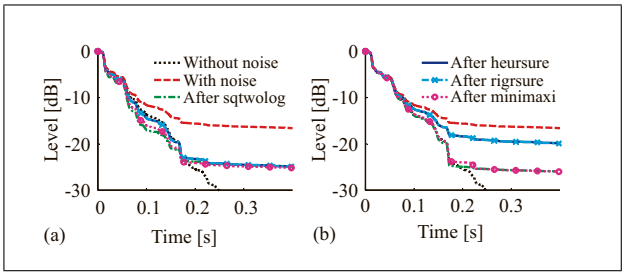


Figure 3. (Colour online) Effects of changing threshold selection rule (*sqtwolog*, *minimaxi*, *rigrsure* and *heursure*) on the wavelet de-noising of the RIR synthesized by the ISM filtered in the third-octave band at 1 kHz using the Haar wavelet and decomposition level of 5 for (a) *soft* thresholding and (b) *hard* thresholding; the legend of the presented curves is split up in 2 parts given in (a) and (b).

In the case of *hard* thresholding, the *rigrsure* and *heursure* rules lead to both the smallest noise floor reduction and the smallest dynamic range improvement, see Figure 3b. The *sqtwolog* threshold rule gives a somewhat larger dynamic range improvement, while the *minimaxi* rule provides the best results. In the case presented in Figure 3b, the dynamic range improvement achieved using the *minimaxi* rule is about 11 dB.

One more parameter that needs to be considered is the threshold rescaling. When the *one* rescaling method meaning no rescaling is applied, the results regarding the dynamic range improvement are very poor, so this method is not used in further research. The RIRs and corresponding EDCs after de-noising done by using the *sln* rescaling are very similar to those before the de-noising. So, the effects of such de-noising are minimal. The best results regarding both the dynamic range improvement and noise floor reduction are achieved using the *mln* rescaling.

Many authors who investigated wavelet de-noising did not consider the level of decomposition parameter. However, in the case of RIR de-noising it is very important to observe the impact of this parameter and find an optimal level. The effects of changing decomposition level by using the Haar wavelet function, *minimaxi* threshold selection rule, *hard* thresholding and *mln* rescaling method on the RIR in the time domain are presented in Figure 4. Noise reduction is almost negligible up to a certain decomposition level (the level of 3 in this particular case). For the decomposition levels higher than this level, noise is reduced more by a higher decomposition level.

Differences between the noisy RIRs and noiseless RIR from Figure 4 are shown in Figure 5. The wavelet transform does not reduce noise in the RIR part up to a certain time located somewhat before the knee, regardless of the decomposition level. Instead, the noise is slightly increased in that part, compared to the noise before the wavelet application, see Figure 5. Noise reduction by wavelets starts somewhat before the knee. There is a region with gradually decreasing noise starting before and finishing after the knee. After that transition region, the pattern of noise becomes steady. The level of this steady

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

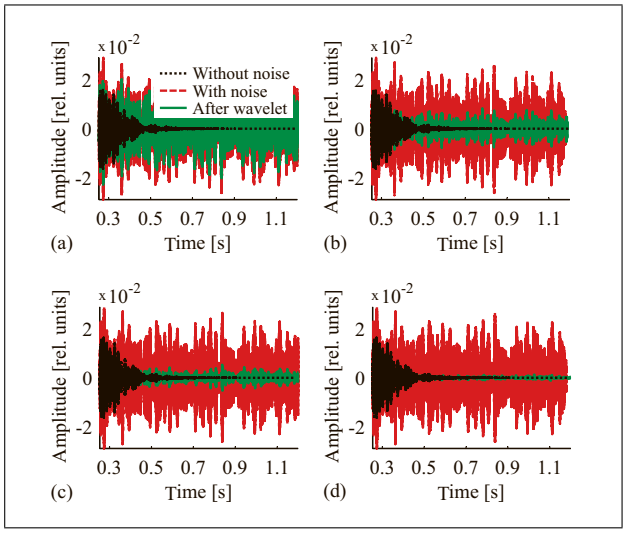


Figure 4. (Colour online) Effects of changing decomposition level by applying the Haar wavelet, *minimaxi* threshold selection rule, *hard* thresholding and *mln* rescaling method on the RIR synthesized by the ISM filtered in the third-octave band at 1 kHz (the RIRs after the knee are shown): decomposition level of (a) 4, (b) 5, (c) 6 and (d) 7 is used.

noise is lower when a higher decomposition level is applied.

Observing the EDCs, it can be noticed that the increase of decomposition level up to a certain level (the level of 3 in this case) has almost no effect on the EDCs. These EDCs are almost identical to the noisy EDC before the wavelet de-noising. Further increase of decomposition level results in the decrease of the noise floor causing the second part of the EDC after the knee to go down, see Figure 6. At the same time, the EDC dynamic range is increased by decomposition level increase up to a certain level (in this case up to the level of 5, here called saturation level). In that regard, the level of decomposition of 5 yields a significant improvement of the EDC dynamic range, and the decay curve has a shape common for the EDCs. The absolute difference between this EDC and the reference (noiseless) EDC is smaller than the threshold up to a particular point (decay level). This point is located about 11 dB below the corresponding point in the EDC obtained without the wavelet de-noising.

In the EDCs obtained by using decomposition levels higher than the saturation level, the dynamic range improvement is smaller than for the saturation level, and deviations from the reference EDC are larger, as shown in Figure 6. With such decomposition levels reverberation decay rates are higher than the corresponding one of the reference EDC as a consequence of signal loss or degradation in the vicinity of the knee.

If the wavelet parameters – threshold selection rule, thresholding, rescaling method and decomposition level – are set to *minimaxi*, *hard*, *mln* and 5, respectively, and the wavelet function is changed, there are significant effects on both the dynamic range improvement and noise floor reduction. The differences of these control measures can

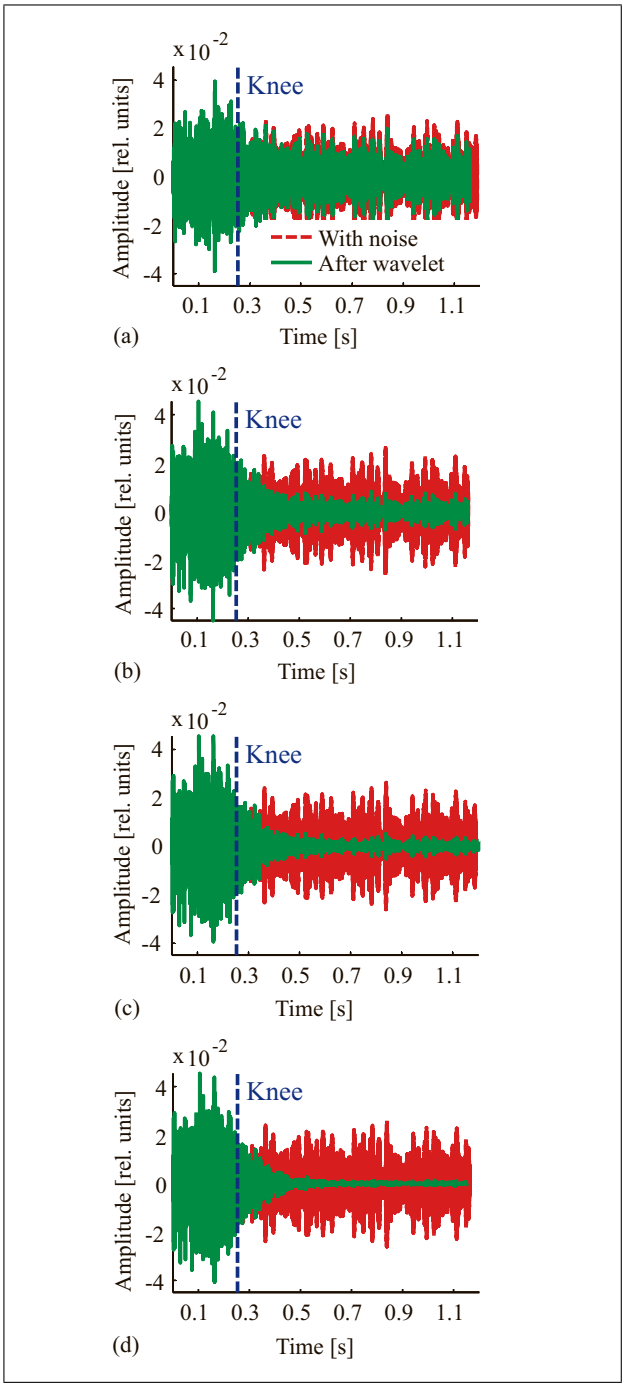


Figure 5. (Colour online) Effects of changing decomposition level: differences between the noisy RIRs (before and after the wavelet application) and the noiseless RIR from Figure 4.

even be larger than 10 dB depending on the wavelet family and number of vanishing moments applied. For each one of the used wavelet families, there is an optimal number of vanishing moments yielding the best de-noising results for that particular family. Some wavelet functions lead to similar results. An example is the Daubechies wavelet with two vanishing moments (Daubechies 2 or *db2* when applied in Matlab), the Coiflet wavelet with one vanishing moment (Coiflet 1 or *coif1* in Matlab) and the Symlet wavelet with two vanishing moments (Symlet 2 or *sym2*

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

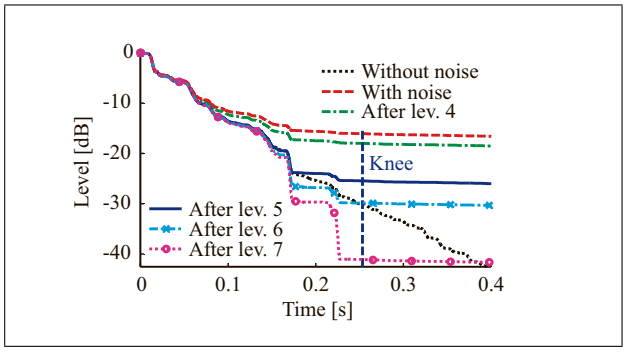


Figure 6. (Colour online) Effects of changing decomposition level: EDCs of the noisy RIRs (before and after the wavelet application) and the noiseless RIR from Figure 4.

in Matlab). In the case of the analyzed RIR, these wavelet functions, but also the Haar wavelet (*haar* when applied in Matlab), provide the best results regarding the dynamic range improvement.

This first phase of the research shows that there is mutual dependence of the wavelet parameters. Based on the presented results, it can be considered that the most promising set of the parameters (leading to the largest dynamic range improvement so far) consists of *minimaxi* threshold selection rule, *hard* thresholding and *mln* rescaling method.

4.2. Effects of changing noise and RIR

When the level of synthetic white noise added to the RIR is changed, the situation remains almost the same. The best results related to the dynamic range improvement are still obtained using the most promising set of the parameters and decomposition level of 5 if the RIR filtered in the third-octave band at 1 kHz is considered. The largest dynamic range improvement, about 12.5 dB, is achieved for the noise levels in the range between −50 dB and −60 dB. The dynamic range improvements are reduced by up to a few dB for noise levels higher than −50 dB and lower than −60 dB. Figure 7 illustrates the dynamic range improvements. The noise floor reduction also depends on the noise level. This reduction increases with the noise level up to a certain level (typically about −40 dB or −30 dB) where the saturation is reached.

The effects of changing wavelet functions are significant here, too. Only some of the wavelet functions lead to the mentioned dynamic improvement greater than 10 dB, including Daubechies 2, Haar and Coiflet 1 wavelets. Besides, it is noticed that the most stable results independently of the noise level are obtained with the Haar and Symlet 1 wavelets.

Random realizations of noise with the same statistical properties typically cause a change of the dynamic range improvement of up to a few dB. The effects of random noise realizations are also presented in Figure 7, where the EDCs for ten different noise realizations generated from trial to trial are shown. The results for noise floor reduction are more consistent, and they are typically within 0.5 dB.

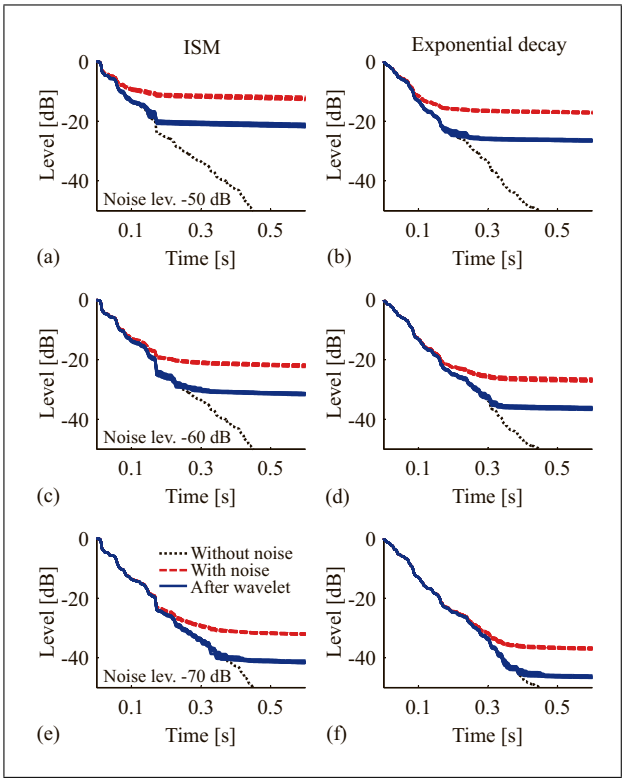


Figure 7. (Colour online) Dynamic range improvement in the RIR synthesized by the ISM and RIR with exponential decay both having different noise levels and random noise realizations (10 random noises generated in 10 trials are used) filtered in the third-octave band at 1 kHz achieved using the *minimaxi* threshold selection rule, *hard* thresholding, *mln* rescaling method, *haar* wavelet and decomposition level of 5.

Another type of noise considered here is pink noise added to the synthesized noiseless RIRs. The wavelet de-noising behaves in a similar way for pink and white noise when the RIRs filtered in the third-octave bands are considered. In the band at 1 kHz, the most promising set of the parameters, decomposition level of 5 and *haar* wavelet provide the dynamic range improvement of about 8.5 dB and noise floor reduction of about 11 dB for both white and pink noise on average (taking into account the RIRs with noise levels from −70 dB to −40 dB in steps of 10 dB).

Changing the RIR in the analysis, but still using only the RIRs synthesized by the ISM, leads to a certain change of the dynamic range improvement, which is typically within a few dB. When the most promising set of wavelet parameters is applied to the RIRs with exponential decays, the dynamic range improvement is similar to the cases of the RIRs synthesized by the ISM, see Figures 7a and 7b. Since the former RIRs are more regular, the shape of the EDCs after wavelet de-noising is more regular, and the dynamic range improvement may be larger in some cases by up to a few dB.

4.3. Optimal decomposition levels for filtered RIRs

The performed analysis has shown that for both white and pink noise, the largest dynamic range improvement and

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

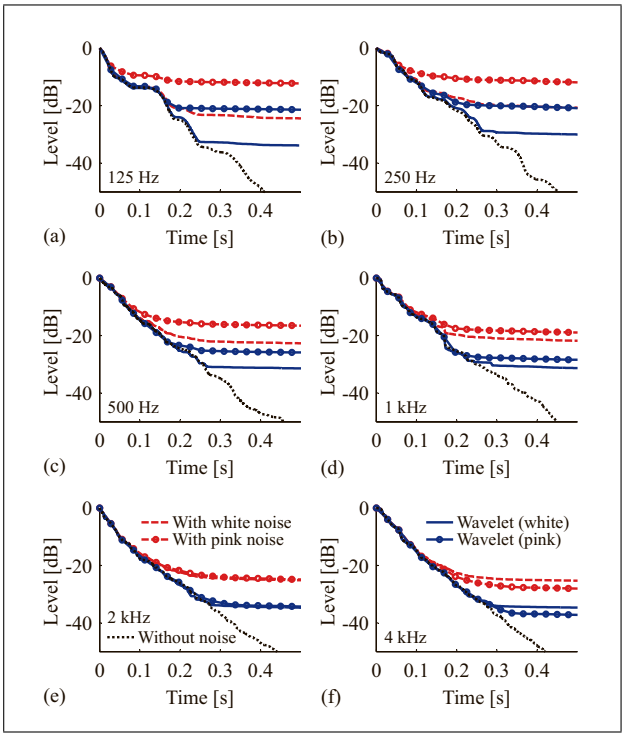


Figure 8. (Colour online) Effects of wavelet de-noising of the RIR synthesized by the ISM (contaminated by either white or pink noise) on EDCs in different frequency bands using the most promising (optimal) set of wavelet parameters, optimal decomposition levels and *haar* wavelet; the legend of the presented curves is split up in 2 parts given in (e) and (f).

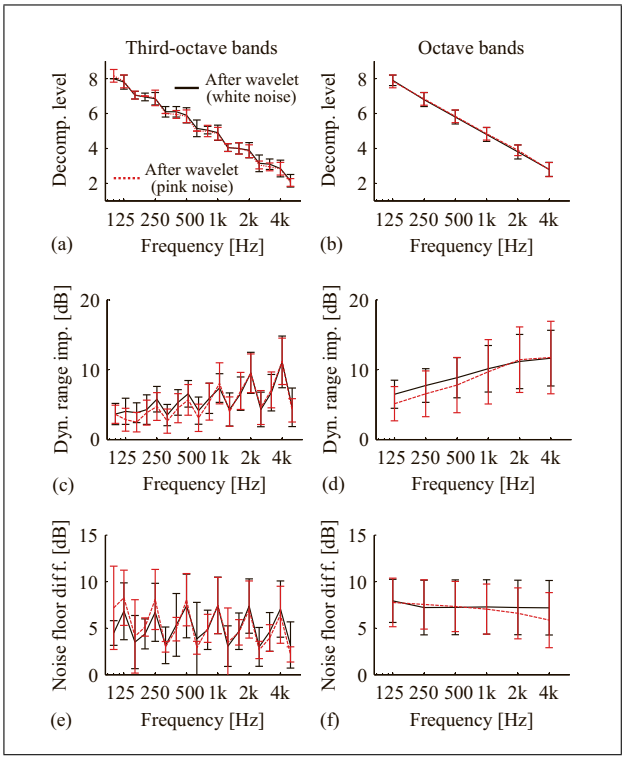


Figure 9. (Colour online) Results of the detailed analysis using 10 random realizations of noise, 10 synthesized RIRs and 2 noise levels (−60 dB and −50 dB): (a) and (b) optimal decomposition levels, (c) and (d) dynamic range improvement as well as (e) and (f) noise floor reduction given as mean – standard deviation in the third-octave and octave frequency bands.

noise floor reduction are obtained using the same set of parameters (*minimaxi* threshold selection rule, *hard* thresholding and *mln* rescaling method). Other combinations of wavelet parameters can typically provide only minor dynamic range improvement and noise floor reduction, e.g., a few dB.

The optimal decomposition level that leads to the best results depends on the frequency range, that is, central frequency of the third-octave or octave band. The lower the central frequency, the higher the decomposition level. The wavelet function affects the results, too. The most consistent (stable) results are obtained using the *haar* wavelet. The effects of wavelet de-noising in different frequency bands are illustrated in Figure 8.

A detailed analysis using ten random realizations of noise, ten different synthesized RIRs and two noise levels (−60 dB and −50 dB) shows that the optimal decomposition level ranges from 8 for the lowest third-octave frequency band at 100 Hz to 2 for the highest frequency band at 5 kHz, see Figure 9a. A similar situation exists for the octave bands, see Figure 9b. The mean values – one standard deviation of the dynamic range improvement and noise floor reduction taken over all analyzed cases are shown in Figures 9c, 9d, 9e and 9f. The overall mean dynamic range improvement is about 7.5 dB and the overall standard deviation of this improvement is about 3 dB. The overall mean value and standard deviation of the noise floor reduction is about 6 dB and 2.5 dB, respectively.

When noise in a RIR has some specific spectral shape, e.g., components at low and high frequencies are greater than those at mid frequencies, the dynamic range improvement and noise floor reduction are also within the limits presented in Figure 9. This is a consequence of stability of the results that generally do not depend on the noise type and level. The noise floor reduction is directly related to the SNR gain.

4.4. De-noising in broadband synthesized RIRs

During the analysis of wavelet de-noising of broadband RIRs, some effects already seen in the de-noising of RIRs filtered in third-octave and octave bands are observed, but also some unique effects are distinguished. The latter ones are described in this section. The optimal set of wavelet parameters for de-noising of broadband synthesized RIRs contaminated with white noise consists of *minimaxi* threshold selection rule, *hard* thresholding and *mln* rescaling method. Using this optimal set of the parameters and changing the wavelet function, there is only a slight effect on both the dynamic range improvement and noise floor reduction.

The effects of changing the decomposition level in the RIR with white noise of level of −50 dB are shown in Figure 10a. The EDC for the optimal decomposition level (the level of 3 in this case) has a shape common for the EDCs, and it does not go below the reference EDC, which is not

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

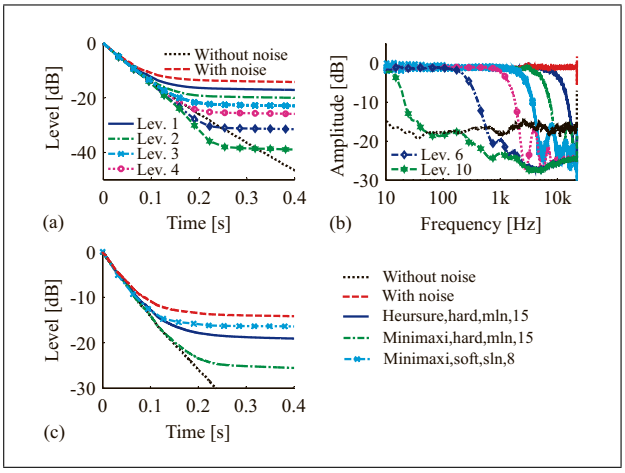


Figure 10. (Colour online) Effects of de-noising of broadband RIRs synthesized by the ISM contaminated by (a) and (b) white noise and (c) pink noise: (a) EDCs and (b) spectra of the noisy RIRs (before and after applying the Daubechies 2 wavelet, *minimaxi* threshold selection rule, *hard* thresholding and *mln* rescaling method) and noiseless RIR, where the spectra are obtained taking the RIRs from the knee to the RIR end; (c) EDCs showing the effects of using 3 optimal sets of wavelet parameters; 10 EDCs and 10 spectra obtained using 10 random noise realizations are shown for the decomposition level of 3 in (a) and (b), respectively.

the case with the EDCs obtained using the decomposition levels higher than the optimal one. The optimal decomposition level depends on the noise level. For the noise level of −50 dB and lower, the optimal decomposition level is 3, and for the noise level of −40 dB and higher, the optimal level is 4. In that regard, the largest dynamic range improvement of about 18 dB is achieved for the noise level of −60 dB. The improvement is about 15 dB for the noise level of −50 dB, and about 12 dB for the noise level of −40 dB.

The spectra of the part of the RIRs from the knee to the end of the RIRs where noise is dominant are presented in Figure 10b. Here, an increase of the decomposition level leads to an extension of the frequency range towards lower frequencies where noise is reduced by the wavelets. It can be considered that the frequency range of noise is reduced by the increase of decomposition level.

When the wavelet de-noising is applied to the room response to the excitation signal, the obtained dynamic range improvement is similar to the case when de-noising is applied to the extracted RIR. If a RIR is measured using the exponential swept sine technique, the RIR extraction by the inverse filter changes the spectrum of noise present in the room response to the excitation. This is why it might be necessary to change the decomposition level (reduce it from 3 to 2 here) in order to reach the largest dynamic range improvement.

In the wavelet de-noising of broadband synthesized RIRs contaminated with pink noise, several combinations of wavelet parameters may be considered to be optimal sets leading to the best results, as shown in Figure 10c. The first set leading to the most consistent results is related to the *heursure* or *rigrsure* threshold selection rule, *hard* thresholding, *mln* rescaling method, decomposition level of 15 and *haar* wavelet function. These parameters provide the dynamic range improvement of about 3 dB and noise floor reduction of about 5 dB. In the second set of wavelet parameters, only the *heursure* or *rigrsure* threshold selection rule from the first set is replaced by the *minimaxi* rule. In this case, the dynamic range improvement can be greater than 10 dB or 15 dB, while the noise floor reduction can be even greater than 20 dB. The third set of the parameters consists of the *minimaxi* selection rule, *soft* thresholding, *sln* rescaling method, decomposition level of 8 and *haar* wavelet function. The dynamic range improvement is about 7 dB and the noise floor reduction is typically from 2.5 dB to 3 dB.

4.5. Wavelet de-noising in measured RIRs

The results for the measured RIRs filtered in third-octave or octave bands are similar to those for the filtered synthesized RIRs. The best de-noising results are again obtained by using the most promising set of parameters and the *haar* wavelet, see Figures 11a to 11d. The optimal decomposition level leading to the largest dynamic range improvement decreases with the increase of the central frequency of the third-octave or octave band of interest. It is important to note that the optimal decomposition levels given in Figures 9a and 9b are also valid for the measured RIRs. The effects of changing the decomposition level around the optimal level are rather large. This means that a level higher than the optimal one causes significant degradation of the reverberation decay. A decomposition level lower than the optimal level results in a smaller dynamic range improvement, but this level might also be acceptable since the degradation of the EDC is typically negligible.

In the measured broadband RIRs, the most promising set of wavelet parameters (the *minimaxi* threshold selection rule, *hard* thresholding and *mln* rescaling method) also gives the best results. The differences in the de-noising are rather large when other wavelet parameters are applied. Considering the wavelet function, the most consistent results are obtained using the *haar* wavelet. Regarding the decomposition level, the best results are achieved using the levels of 3 and 4 for some measured broadband RIRs, while for some other RIRs the best results are obtained using the levels of 5 and 6. The higher level is more appropriate for the RIRs where the noise floor is higher. An illustration of the effects of the wavelet de-noising in the measured broadband RIRs is given in Figures 11e and 11f.

4.6. Wavelet de-noising versus two other methods

A comparison of the wavelet de-noising with two other noise compensation methods (truncation and subtraction-truncation-correction) for reverberation time estimation is described in this section. For that purpose, the optimal set of wavelet parameters is applied to the synthesized RIRs of both mentioned types filtered in the third-octave bands. When the range for reverberation time estimation (linear

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

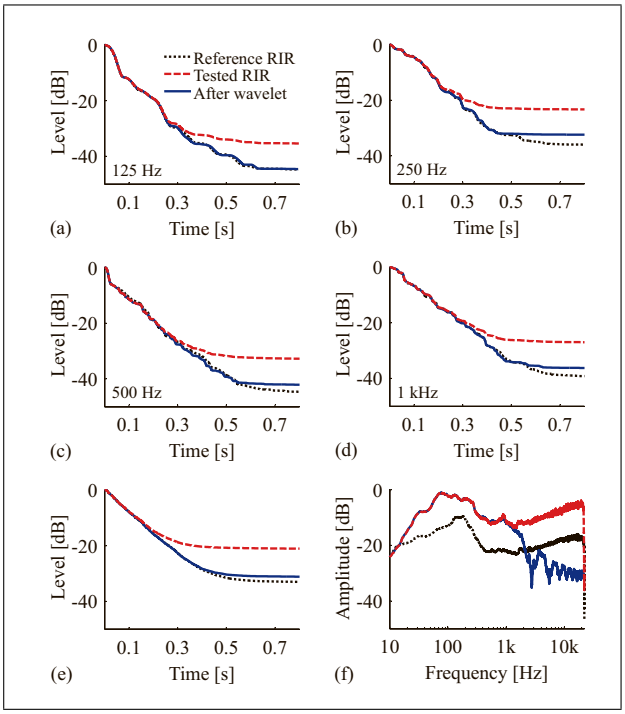


Figure 11. (Colour online) Effects of wavelet de-noising of: (a) to (d) the measured RIRs filtered in third-octave bands, and (e) to (f) the measured broadband RIR (tested RIR) using the optimal set of wavelet parameters, optimal decomposition levels and *haar* wavelet: (a) to (e) EDCs, and (f) spectra of the RIRs before and after the wavelet application obtained taking the RIRs from the knee to the RIR end; the reference RIR is measured using 64 times longer excitation signal (swept sine) than in the case of the tested RIR.

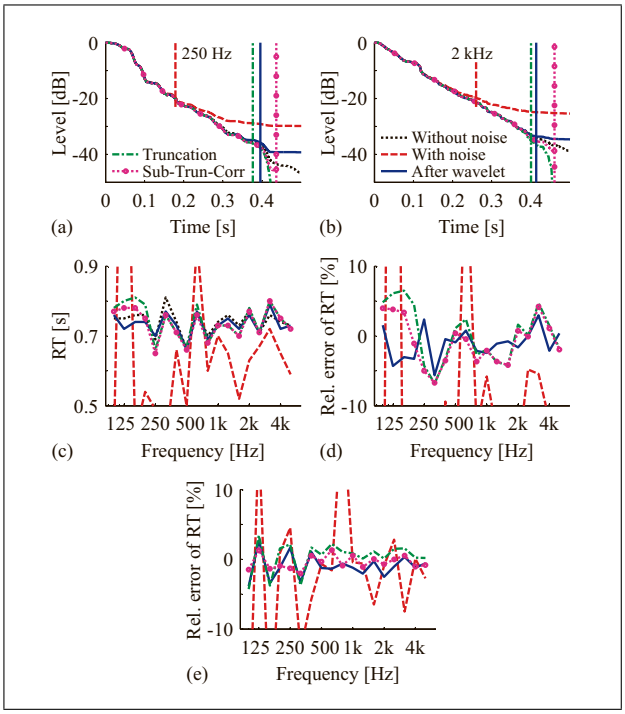


Figure 12. (Colour online) Effects of noise compensation on reverberation time estimation: (a) and (b) EDCs obtained from the synthesized RIR without noise and with noise of level of - 60 dB in the third-octave bands (the end points of estimation ranges are denoted by the vertical lines of the same line styles as the corresponding EDCs); (c) the reverberation time (RT) and (d) relative error of this time calculated using the same RIR as in (a) and (b); (e) mean values of the relative error of reverberation time calculated from 14 cases of the synthesized RIRs with various noises; the curve legend is split up in 2 parts given in (a) and (b).

regression range) is determined as described in Section 3, the noise compensation methods, including the wavelet de-noising, enable a considerable extension of the estimation range, see Figure 12a and 12b. Consequently, the error of the estimated reverberation time is reduced. Considering the matching of the compensated EDCs to the noiseless EDC, generally similar matching is achieved by the applied noise compensation methods up to a certain point, e.g., up to the point indicated by the dash-dotted vertical line associated with the truncation method in Figures 12a and 12b.

The estimated reverberation times, and associated relative errors (given as signed errors), in third-octave bands, for a single RIR, are shown in Figures 12c and 12d. The mean results based on 14 cases of the synthesized RIRs with various noises having three noise levels (−70 dB, −60 dB and −50 dB) are given in Figure 12e. Considerably larger errors of the reverberation time are obtained when none of the compensation methods is applied. Comparing the noise compensation methods, slightly better results (more accurate reverberation times) are produced using the subtraction-truncation-correction method, while the wavelet de-noising is typically comparable to the truncation method. It is worthwhile to note that in most cases where the noise compensation methods are applied, the relative errors of the reverberation time are within the limits of just noticeable difference (±JND), which is 5 % for the reverberation time [13].

As the wavelet de-noising is sensitive to the wavelet parameters, RIR and noise, in some cases, the EDCs obtained after the wavelet de-noising match the noiseless EDCs in a better way than the EDCs obtained using two other compensation methods. Similarly, in some cases, the wavelet de-noising leads to more accurate estimated reverberation times than the compared compensation methods, especially the truncation method. This emphasizes the potentials of the wavelet de-noising, but also its sensitivity to the parameters and RIR properties.

4.7. Truncation time estimation by wavelet de-noising

By applying the most promising set of wavelet parameters with a particular decomposition level and *haar* wavelet function to the RIR filtered in a third- octave or octave band, specific shapes of the de-noised RIR and corresponding EDC are obtained, see Figure 13. The decomposition level should be high enough, e.g. 15, regardless of the frequency band of interest. Typically, a de-noised filtered RIR follows the noisy filtered RIR approximately up to the knee. After that point, the RIR amplitude drops

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

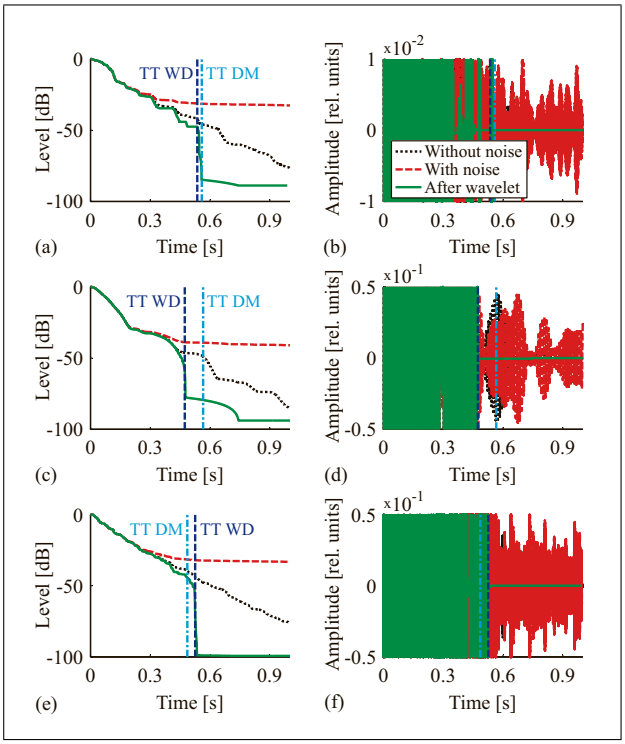


Figure 13. (Colour online) Three typical cases appearing in estimation of the truncation time by the wavelet de-noising: sudden drop of (a) EDC and (b) RIR located at the knee, sudden drop of (c) EDC and (d) RIR before the knee and existence of the mid-region, and sudden drop of (e) EDC and (f) RIR after the knee; truncation times (TT) estimated by the nonlinear decay model (DM) and wavelets de-noising (WD) are presented by the vertical lines.

suddenly to values close to 0, see Figure 13b. This sudden drop also causes a similar drop in the EDC, as can be seen in Figure 13a. Such a behavior can be used for estimation of the truncation time at the knee, that is, at the sudden drop of the de-noised RIR. In the analysis here, the intersection of the RIR's main decay and the noise floor estimated by the nonlinear decay model [19] is taken as the knee (reference truncation time).

Potential problems in the truncation time estimation are related to the dislocation of the RIR sudden drop from the knee. This dislocation happens for some of the tested synthesized RIRs typically at low and mid frequencies. In the majority of such cases at frequencies up to about 400 Hz, there is a mid-region between the sudden drop of the de-noised RIR and the region of RIR amplitude values close to 0, see Figures 13c and 13d. Here, the smallest deviation of the truncation time from the knee is commonly obtained if the truncation time is set at the end of the mid-region. Besides, the mid-region can have a more complex shape (multi-decay shape) in some rare cases. In order to have the truncation time closer to the knee here, it is preferable to set the truncation time in the middle of the mid-region.

At mid and high frequencies, the mid-region usually does not exist. Thus, when the sudden drop is dislocated from the knee, as can happen in the range from 1 kHz to 2.5 kHz, it is difficult to make any corrections of the estimated truncation time, see Figures 13e and 13f. Fortunately, the number of such cases is not that large, and the frequency range of this behavior is not that wide.

The estimation of the truncation time (knee) by the wavelet de-noising is compared with the procedure from the literature based on the nonlinear decay model [19]. The obtained times are presented as vertical lines in Figure 13. When the truncation time is estimated to be at the sudden drop of the de-noised RIR, the difference between this time and the time estimated by the nonlinear decay model is below 20 ms or even 10 ms in the majority of cases. This means that the de-noised RIR sudden drop is close to the knee. In a relatively small number of cases, the sudden drop is further dislocated from the knee (usually towards the RIR start). The difference in the estimated truncation times is then larger - it is typically between 20 ms to 60 ms, although it can be as large as 100 ms. When it is possible to apply the correction using the mid region, the difference between the estimated truncation times is again up to 20 ms or 30 ms.

The truncation time estimation by the wavelet de-noising can be realized in an automated way, especially if the truncation time is detected at the de-noised RIR sudden drop. The correction of the truncation time estimation using the mid-region is more complex for the automated application, although it is not impossible to realize the correction as an automated procedure.

5. Discussion

Generally speaking, the choice of wavelet parameters including the threshold selection rule, thresholding, rescaling, decomposition level as well as the wavelet function can have a significant impact on the RIR de-noising. Thus, some combinations of wavelet parameters lead to a larger noise floor reduction. Unfortunately, it typically comes with the larger degradation of the reverberation decay, especially degradation of the late decay that can result in an incorrectly estimated decay rate. This is why it is desirable to have an optimal set of wavelet parameters yielding as large as possible noise floor reduction with no or minimal degradation of the reverberation decay.

When the wavelet de-noising is applied to synthesized RIRs filtered in third-octave or octave bands, independently of the noise type, noise level and RIR characteristics, there is a unique optimal set of wavelet parameters. This set includes the *minimaxi* selection rule, *hard* thresholding, *mln* rescaling and *haar* wavelet. The optimal decomposition level depends on the central frequency of the frequency band, and it ranges from 8 to 2. The mean dynamic range improvement is from about 4 dB to above 10 dB, while the mean noise floor reduction is about 5 dB and 7 dB in the third-octave bands and octave bands, respectively. The results for de-noising of measured RIRs filtered in third-octave or octave bands are similar to those for the filtered synthesized RIRs. Comparing the wavelet de-noising of broadband and filtered RIRs, more stable and consistent results are obtained in the latter case.

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

The dynamic range improvement is a measure prone to variations of several dB caused by differences in random realizations of noise as well as differences in the RIRs. The differences in the noise floor reduction caused by the differences in the applied RIRs and random realizations of noise might also be several dB, e.g., when the *minimaxi* selection rule is applied at a higher decomposition level such as 15. However, in many cases, especially when using a lower decomposition level such as 3 or 5, these differences are smaller than 0.5 dB.

Although the wavelet de-noising seems to be a complicated method, following the instructions given here it can easily be applied even as an automated method. In Matlab, it can be realized using the existing function (wden). The truncation method and especially subtraction-truncation-correction method may require sophisticated procedures for truncation time estimation (and correction term calculation). The accuracy of the reverberation time estimation presented in Section 4.6 for the truncation method and subtraction-truncation-correction method depends on the applied procedure. For some procedures, the accuracy is worse than the one presented here, see, e.g. [19] and [27]. Taking into account the mentioned potentials of the wavelet de-noising to be easily applied in an automated way and to give more accurate results than the compared methods in some cases, the wavelet de-noising can be considered to be one of alternative compensation methods. An additional benefit of the wavelet de-noising is related to potential usage of this approach for estimation of the truncation time.

6. Conclusions

The research presented here shows that RIRs can be de-noised by applying wavelets. However, due to the signal (reverberation decay) degradation, special care should be given to such an application. It is observed that the RIR de-noising strongly depends on the wavelet parameters. Some combinations of parameters lead to a rather small noise floor reduction, while some other parameter combinations lead to a large noise floor reduction, but also a large degradation of the reverberation decay. Such de-noising cannot be considered acceptable. However, there are sets of wavelet parameters that enable significant noise floor reduction with minimal or almost no degradation of the reverberation decay.

The optimal set of wavelet parameters for de-noising of a RIR filtered in third-octave or octave bands comprises the *minimaxi* threshold selection rule, *hard* thresholding, *mln* rescaling method and *haar* wavelet. The optimal decomposition level takes a value of 8 at low frequencies, and it decreases with frequency increase reaching a value of 2 at about 5 kHz. This optimal set provides the best results independently of the noise and RIR characteristics. In most of the cases of wavelet de-noising of a broadband RIR, the optimal parameter set also consists of the *minimaxi* rule, *hard* thresholding, *mln* method and *haar* wavelet, while the decomposition level takes a value from a wide range from 3 to 15. Such a parameter set enables the EDC dynamic range to be improved and the noise floor to be reduced typically from 9 dB to even 20 dB.

The wavelet de-noising enables a significant reduction of the error caused by noise in the estimation of the reverberation time. In addition to RIR de-noising, the wavelets might be applied for the estimation of RIR truncation time at the knee or in its vicinity.

Acknowledgement

The authors thank the anonymous reviewers for constructive reviews and many helpful comments. They also thank Professor of English Lena Tica for proofreading the manuscript.

References

[1] A. Kashyap, M. Prasad: Audio noise cancellation using Wiener filter based LMS algorithm using LabVIEW. Int J Emerging Technology and Advanced Engineering **3** (2010) 599–601.

[2] P. Górski, L. Morzynski: Active noise reduction algorithm based on NOTCH filter and genetic algorithm. Arch Acoust **38** (2013) 185–190.

[3] B. J. Saikia, U. Baruah: A GA-NN based wavelet method of speech signal denoising. IEEE Int Conf Communication and Informatics, Coimbatore, India, Jan. 2015.

[4] Đ. Damnjanović, M. Gojković, D. Ćirić: Audio signal denoising using wavelets algorithm in Matlab and LabVIEW. Proc Int Conf IcETRAN 2014, Vrnjačka banja, Serbia, Jun 2014, AKI1.2.1–6.

[5] G. Luo, D. Zhang: Wavelet denoising. Chapter 4 in: Advances in wavelet theory and their applications in engineering. Physics and Technology, InTech (2012) 59–81.

[6] B. Ergen: Signal and image denoising using wavelet transform. Chapter 21 in: Advances in wavelet theory and their applications in engineering. Physics and Technology, InTech (2012) 495–515.

[7] M. Y. Gokhale, D. K. Khanduja: Time domain signal analysis using wavelet packet decomposition approach. Int. J. Communications, Network and System Sciences **3** (2010) 321–329.

[8] R. Cohen: Signal denoising using wavelets. Project Report, Technion, Israel Institute of Technology, 2012.

[9] M. Guski, M. Vorländer: Comparison of noise compensation methods for room acoustic impulse response evaluations. Acta Acust united Ac **100** (2014) 320–327.

[10] D. G. Ćirić, M. A. Milošević: Transient noise influence in MLS measurement of room impulse response. Acta Acust united Ac **91** (2005) 110–120.

[11] A. Lundeby, T. E. Vigran, H. Bietz, M. Vorländer: Uncertainties of measurements in room acoustics. Acustica **81** (1995) 344–355.

[12] C. L. Christensen, G. Koutsouris, J. H. Rindel: The ISO 3382 parameters: Can we simulate them? Can we measure them? Proc Int. Symposium on Room Acoustics, Toronto, Canada, Jun 2013, paper P010.

[13] C. C. J. M. Hak, R. H. C. Wenmaekers, L. C. J. van Luxemburg: Measuring room impulse responses: impact of the decay range on derived room acoustic parameters. Acta Acust united Ac **98** (2012) 907–915.

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only

[14] M. R. Schroeder: Integrated-impulse method measuring sound decay without using impulses. J. Acoust. Soc. Am. **66** (1979) 497–500.

[15] ISO 3382-1: Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces. International Organization for Standardization, Geneva, 2009.

[16] M. Vorländer, M. Kob: Practical aspects of MLS measurements in building acoustics. Appl. Acoust. **52** (1997) 239–258.

[17] U. P. Svensson, J. L. Nielsen: Errors in MLS measurements caused by time-variance in acoustic systems. J. Audio Eng. Soc. **47** (1999) 907–927.

[18] W. T. Chu: Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method. J. Acoust. Soc. Am. **63** (1978) 1444–1450.

[19] M. Janković, D. G. Ćirić, A. Pantić: Automated estimation of the truncation of room impulse response by applying a nonlinear decay model. J. Acoust. Soc. Am. **139** (2016) 1047–1057.

[20] M. Sifuzzaman, M. R. Islam, M. Z. Ali: Application of wavelet transform and its advantages compared to Fourier Transform. J Physical Sciences **13** (2009) 121–134.

[21] S. R. Messer, J. Agzarian, D. Abbott: Optimal wavelet denoising for phonocardiograms. Microelectronics J. **32** (2001) 931–941.

[22] D. L. Donoho, J. M. Johnstone: Ideal spatial adaptation by wavelet shrinkage. Biometrika **81** (1994) 425–455.

[23] E. A. Lehmann, A. M. Johansson: Prediction of energy decay in room impulse responses simulated with an image-source model. J. Acoust. Soc. Am. **124** (2008) 269–277.

[24] D. G. Ćirić, M. A. Milošević: Optimal determination of truncation point of room impulse responses. Building Acoustics **12** (2005) 15–30.

[25] E. A. Lehmann, A. M. Johansson: Diffuse reverberation model for efficient image source simulation of room impulse responses. IEEE Trans. Audio, Speech, Language Process. **18** (2010) 1429–1439.

[26] IEC 61260: Electroacoustics - Octave-band and fractional-octave-band filters. International Electrotechnical Commission, Geneva, 1995.

[27] D. G. Ćirić, M. Janković: Correction of room impulse response truncation based on a nonlinear decay model. Appl. Acoust. **132** (2018) 210–222.

Document downloaded by @DAEL, 2026-11-20 03:27:09 - personal use only