Extraction of the envelope from impulse responses using pre-processed energy detection for early decay estimation

Simona De Cesaris, Dario D'Orazio, Federica Morandi, and Massimo Garai[a)]
Department of Industrial Engineering, University of Bologna, Bologna, Italy

(Received 27 January 2015; revised 5 September 2015; accepted 11 September 2015; published online 29 October 2015)

The Schroeder's backward integration method and its applications have been widely studied in the literature; some papers analyze the performance of the method, some others suggest various enhancement techniques. In spite of these findings, there exist several cases where the energy decay curve extracted using the classical backward integration method and the parameters computed from it seem not always representative of the phenomenon under study. Among them, the cases where the early decay is dependent on strong, distinct reflections occurring just after the direct wave, as in most Italian opera houses. Other cases are measured impulse responses with a very low signal-to-noise ratio or missing the direct wave. In the literature, alternatives to the Schroeder's method have been proposed, ranging from Hilbert transform to non-linear processing techniques. In this work a method for the extraction of the envelope based on pre-processed energy detection for early decay estimation is proposed. It is shown that it returns an envelope well matching the first part of the decay even in non-linear cases, returning detailed information on the first part of the decay. The performance of the proposed method is presented and discussed for some exemplary impulse responses measured in historical opera houses. A preliminary study on the perceptive relevance of the method is finally presented.
[http://dx.doi.org/10.1121/1.4931904]

[MV] Pages: 2513–2523

I. INTRODUCTION

It is well known that in a music hall the early sound reflections are strongly related to the perceived sense of reverberation, which in turn is more related to the early decay time (EDT) than to classical reverberation time (RT), e.g., estimated through T_{20} and T_{30}.[1–6] Unfortunately, the evaluation of EDT from measured impulse responses is more critical than that of RT, due to the most pronounced fluctuations of the decay curve calculated from the impulse response in the first 5 dB of decay. If the sound level decay is linear, the EDT and RT values are almost identical, but when the decay is non-linear the EDT and RT may have values differing more than the just noticeable difference (JND).[4,5] Given an impulse response, the standard method for the calculation of the reverberation time leads to EDT and RT values closely related to each other as they are derived from the energy decay curve, which accumulates energy from the tail to the beginning of the impulse response (IR).[4] Non-linear decays are typical of rooms with strong direct sound and/or early reflections,[4,7] e.g., Italian opera houses.[8–10] In these situations, the standard methods for calculating the reverberation time from a decay curve extracted from the impulse response is critical because the late part of the decay affects the evaluation of the slope on the very first part, which is the only one taken into account in the definition of EDT.[3] In order to highlight this peculiarity a different method may be required; it should be robust and independent from statistical assumptions on the first part of the impulse response.

Therefore, it is worth briefly revising the standard methods for calculating the reverberation time in order to clearly identify their points of strength and weakness.

In the following, the authors will refer to *energy decay curve* (EDC) to represent the Schroeder integral built on the IR; *envelope* (ENV) to describe a curve tangent to the local maxima of a squared IR. *Decay curve* indicates a generic curve built on the squared impulse response which represents the decay versus time (ENV and EDC are both included in this definition).

The original method to extract the reverberation time is the interrupted noise method;[5] the sound level decay curve is recorded after switching off an artificial sound source emitting a stationary noise. In order to attenuate the random fluctuations of the single decays, it is necessary to average a large number of decays. Schroeder demonstrated the equivalence between the sound level decay of an impulse response and the average of a large number of decays of an interrupted stationary noise when his backward integration method is used to extract the decay curve from the impulse response.[11,12] Other authors proposed some improvements of the Schroeder's method to minimize the detection error of the decay curve.[13–16] A recent review of these methods has been done by Guski and Vorländer.[17]

Other ways to extract the decay curve are based on neural networks,[18–20] segmentation procedures[21] or blind approaches.[22–24] Satoh *et al.*[25] proposed a method to extract the reverberation time based on the envelope instead of the decay curve. Recently a comparison between envelope and

[a)]Electronic mail: massimo.garai@unibo.it

backward integration has been discussed by Luizard and Katz.[26] Meissner[27] used the Hilbert transform to extract the envelope from the single frequency excited response.

The accuracy of the decay curves extracted from measurements has been discussed comparing different theoretical approaches and measurements.[16,28–31]

Starting from the energy decay curve, the reverberation time can be derived either from a linear fitting or using multi-order based methods.[15,32–34]

The aim of this work is to propose a decay curve that provides a well-suited description of the first part of the IR, independently from statistical assumptions. In fact the Schroeder's method assumes a Gaussian distribution of the power spectra,[11] but this is not always true in the very first part of the impulse response, where distinct early reflections may exist. The statistical properties of the impulse response have been generalized by different authors studying their Gaussian properties in the time domain[35–37] and in the time-frequency domain.[38] On the other hand, an alternative proposed in the literature is the construction of the envelope, which is not constrained by any assumption.[25–27] Its main drawback is that the envelope is not equivalent to the ensemble average of a large number of decay curves as the Schroeder's integral is.

For these reasons the envelope based method proposed here constructs the decay curve from the IR borrowing the idea of energy detection, already used in the 50's as a demodulator. This decay curve could then be studied using non-linear Bayesian techniques,[15,32–34] but in the present work it is fit with a linear regression in order to permit an effective comparison between the proposed method and the Schroeder's backward integration method.

The extraction of the decay rates from the envelope in general will return different values than those extracted from the energy decay curve, but on the other hand it is more sensitive to the actual distribution of the sound energy in the initial stage of the decay. Therefore, the envelope-based method might provide additional useful information on the first part of the decay.

The paper is organized as follows: in Sec. II an insight into the extraction techniques of the reverberation time is presented; in Sec. III the pre-processed energy detection method is described (the pseudocode of the algorithm is given in Appendix A); in Sec. IV the proposed method is tested against the most accredited literature method. The perceptive relevance of the proposed method is preliminary studied and the results are presented in Appendix B. Finally, the motivations, the results and the open questions are discussed in the conclusions.

II. EXTRACTION OF THE REVERBERATION TIME

Using a time-based representation dating back to Sabine,[39] the envelope of the late part of the decay is usually modeled using an exponentially decreasing function with a single time constant related to the reverberation time. This is acceptable only for the part of the decay after the mixing time[37] and only if the sound absorption coefficient is uniform in space.[40] Before the mixing time, the decay is multi-modal due to the contribution of single reflections. If the sound absorption is not uniform in the different directions of propagation, the decay after the first 5 dB may also have different time constants.[40]

Therefore, the evaluation of the reverberation time may be generalized as a fitting problem of a multi-modal decay.[41] Then the reverberation time is obtained from the decay rate of the decay curve of the squared impulse response. The problem, formulated as a least squares fitting, may be written as follows:

$$\min_{k_1\ldots k_n} \int_{t_1}^{t_2} [y(t) - x(t, k_1, \ldots, k_n)]^2\, dt, \tag{1}$$

where $y(t)$ is the dB-scaled decay curve of the squared impulse response and $k_1,\ldots, k_n$ are the n parameters of the optimal fitting function x. Actually, this formulation includes two different problems:

(1) The determination of the decay curve of the squared impulse response, $y(t)$;
(2) The determination of the best fit to this curve, $x(t)$.

The decay curve $y(t)$ may be extracted in different ways. The most used method is the Schroeder's backward integration:[11]

$$y_{\mathrm{EDC,sch}}(t) = \int_t^{t_u} h^2(\xi)\, d\xi, \tag{2}$$

where t_u is the upper limit of integration and $h(t)$ is the impulse response.

Another method to extract the energy decay curve is the moving integration, used for example, by Lundeby *et al.*:[16]

$$y_{\mathrm{EDC,lun}}(t) = \int_{t-T/2}^{t+T/2} h^2(\xi)\, d\xi, \tag{3}$$

where T is the width of the moving integration window. The choice of the interval of integration in Eq. (3) has been discussed in the literature in order to properly take into account the noise floor,[14,16,41] to exclude early reflections and more generally to optimize the envelope extraction technique.

All the methods mentioned above rely on the Schroeder's integral, Eq. (2). To highlight its peculiarities, the method, in its basic form, can also be written as[42]

$$y_{\mathrm{EDC,sch}}(t) = \int_0^{t_u} h^2(\xi)\, d\xi - \int_0^{t} h^2(\xi)\, d\xi. \tag{4}$$

Toward the end of the decay curve ($t \to t_u$), the second integral has a value comparable to the first one. Near the beginning of the decay curve ($t \to 0$) the second integral has a negligible value compared to the first one, so that the decay function $y(t)$ is determined mainly by the first integral. In other words, the slope of the very first dB of the decay is determined by the energy accumulated in the subsequent part of the decay by the first integral. The effect of the fluctuations in the very first part of the decay is taken into account only through the second integral, which may have a small value when $t \rightarrow 0$. Thus the Schroeder's method is less sensitive in the very first part of the decay when strong early reflections exist. In addition to this, as seen in Sec. I, the Schroeder's hypothesis regarding the Gaussian distribution of the power spectra[43] is not verified for the first part of the IR.

The decay curve could alternatively be calculated as an envelope, for example, using the Hilbert transform:[41,44] the impulse response $h(t)$ is converted into an analytic signal so that $h(t)$ is its real part and its imaginary part is the Hilbert transform of $h(t)$. In the literature the feasibility of the Hilbert transform method has been shown only for minimum-phase signals, but the acoustic impulse response of a room in general is not minimum-phase,[41,44] as a consequence the Hilbert transform may not return a correct envelope in real-life cases. Kumaresan and Rao[45] showed that in case of a single frequency excitation the minimum-phase condition is verified; the infrequent applications of Hilbert transform-based methods for calculating the reverberation time may be related to this need of a single frequency excitation of the sound field.[27]

Once the decay curve $y(t)$ is determined, its fitting function $x(t)$ in Eq. (1) must be calculated. It may depend on two or more parameters k_i. In the first order approach, the function x may be written as

$$x(t, k_1, k_2) = k_1 + k_2 t, \tag{5}$$

where k_2 is the slope of the linear curve and k_1 is related to the arrival time of the direct field.

Second order-based approaches have been suggested by Xiang,[15] who proposed an alternative fitting method using a non-linear iterative regression for evaluating the reverberation time from the energy decay curve. In 2002 Karjalainen *et al.*[41] studied non-linear optimization techniques by generalizing the method proposed by Xiang. Later, the decay function model established in Ref. 15 has been extended to multi-rate decay functions by Xiang and Goggans[32,33] and Xiang *et al.*:[34] they demonstrated that a model-based analysis using Bayesian probability theory is well-suited for the determination of the decay times from multi-decay rates measurements. These multi-order approaches are generally applied to coupled spaces, where the multiple slopes in the decay curve occur well after the mixing time. In fact, in the Bayesian analysis, all the data are taken from the normalized Schroeder's decay function starting from –5 dB.[33]

The literature cited so far does not deal specifically with the first 5 dB of the decay, both for what concerns the extraction of the decay curve $y(t)$ and the evaluation of the decay rate of $x(t)$. In fact none of the above mentioned methods is specifically tailored to accurately represent the first part of the decay, say from 0 to -5 dB, which is essential for the determination of the EDT.

III. THE PRE-PROCESSED ENERGY DETECTION METHOD (PPED)

In the light of the preceding discussion, the aim of the present work is to investigate an alternative method to calculate the envelope function $y_{\mathrm{ENV}}(t)$ [Eq. (1)] that can be fully representative of the first part of the decay of the squared IR.

The energy detection was already in use in the 50's as a demodulator. The original energy detection is based on a rectifier followed by a post detection low-pass filter.[46] In this work a modified energy detection algorithm is used. A suitable filter select only the lower components of the squared spectrum. A tuning of the filter permits an adequate extraction of the envelope. From this point of view, the energy detection method may be associated to the moving integration used in Eq. (3), where the center frequency of the filter is related to the width of the integration window. Both methods present some critical issues: in the method developed by Lundeby *et al.*[16] the optimal window width cannot be determined in advance and has to be found by iteration; in the original energy detection method, the cutoff frequency of the post detection filter cannot be determined in advance.

In the present work, a iterative pre-processing has been developed to improve the accuracy of the energy detection method (see Fig. 1).[47,48] The improved method is composed of three main steps (see the pseudocode in the Appendix A for details):

(1) The squared impulse response is pre-processed;
(2) The data from step 1 are filtered with a high-slope low-pass filter;
(3) The maxima of the data from step 2 are selected;
(4) The data from step 3 are fitted with a monotone piecewise cubic interpolation.

A. Pre-processing

The pre-processing (SHARPENING in the pseudocode) is an iterative algorithm that compares two by two

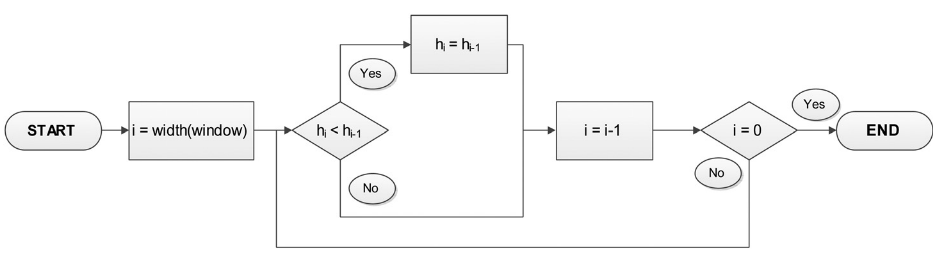


FIG. 1. Flow chart of one iteration of the pre-processing step.

the samples of the squared IR from the end to the beginning. At the first iteration the processed data correspond to the original squared IR. For the subsequent iterations, the processed data correspond to the output of the previous iteration. At each step the amplitude of each sample is compared with the one of the nearest sample and then the sample with the smaller amplitude value takes the value of the greater one (see Fig. 1). Considering only this step the number of iterations could not be fixed because it would be strictly dependent on the nature of the signal under test. However considering the whole PPED method it is possible to fix the number of iterations thanks to an accurate tuning of the post-detection filter (see step 2) that, combined with this pre-processing, eliminates any potential ambiguity. Testing the entire method on many room impulse responses (RIRs) measured with Fs = 48 000 Hz, it was found that a number of exactly ten iterations can be fixed *a priori* and never changed.

B. Post-detection filter

It is implemented, like the energy detection filter, as a high-slope low-pass filter. A relevant feature is the cut-off frequency of the filter. It was found that the correct cut-off frequency depends on the center of the frequency band of interest (Fc); thus for each octave band or one-third octave band the cut-off frequency of the post detection filter is different. As well as the number of iterations, also the shape of the filter seems to be dependent on the statistical characteristics of the signal.[46] The shape of this filter is a modified Blackmann Harris window: it is designed as a flat window (rectangular window) from 0 to Fc Hertz followed by a single side Blackmann Harris window centered at Fc and of width Fc (DESIGNFILTER in the pseudocode).

C. Selection of maxima

The maxima are selected imposing the monotonicity: from the end to the beginning of the data vector, the maxima are selected only if their amplitude is greater than the amplitude of the previous one (SELMAX in the pseudocode). To avoid inaccuracy in particular cases (two consecutive and very different local slopes of the envelope could lead to an incorrect estimate of a global reverberation time), the slope between two consecutive maxima related to the two closest couples of maxima (to left and to right) is also verified. In this way for each comparison three consecutive maxima are considered. If the two slopes are very different, the center point of the three is eliminated from the previous set of selected points.

D. Cubic interpolation

To obtain a good evaluation of the reverberation time through a linear fitting performed over the decay curve, the latter should be a decreasing function. In 1980 Fritsch and Carlson[49] proposed an algorithm performing a monotone piecewise cubic interpolation, which in the present work is applied to the entire sound level decay. In this step (CUBICPIECEHERMITE in the pseudocode) the maxima selected in step 3 become the *knots* of a piecewise cubic Hermite interpolation. In this way it is possible to impose the decreasing monotonicity not only at the *knots* (see Sec. III C) but also between two consecutive maxima. Because of the fact that the interpolation domain is the whole processed squared IR, this step returns an interpolation curve that describes the behavior of the squared IR globally as a function of time. If the samples of the measured IR are spaced T (= 1/Fs) seconds, also the points of the interpolation curve are spaced T seconds. In this way the method returns a monotonically decreasing curve with the same temporal definition of the measured IR (as the Schroeder integral does).

In Fig. 2 the steps are applied consecutively to a test impulse response and the intermediate results are graphically shown.

Two relevant features of the PPED method emerge from the description above. The first is that this method does not need a preliminary estimation of the background noise. The second is that it returns a very detailed decay curve, sensitive to each variation of the squared IR decay. This is a consequence of the fact that the method returns a "locally" defined decay curve.

IV. EXPERIMENTAL RESULTS

Different benchmarks have been proposed to test different methods for the extraction of the reverberation time.[50]

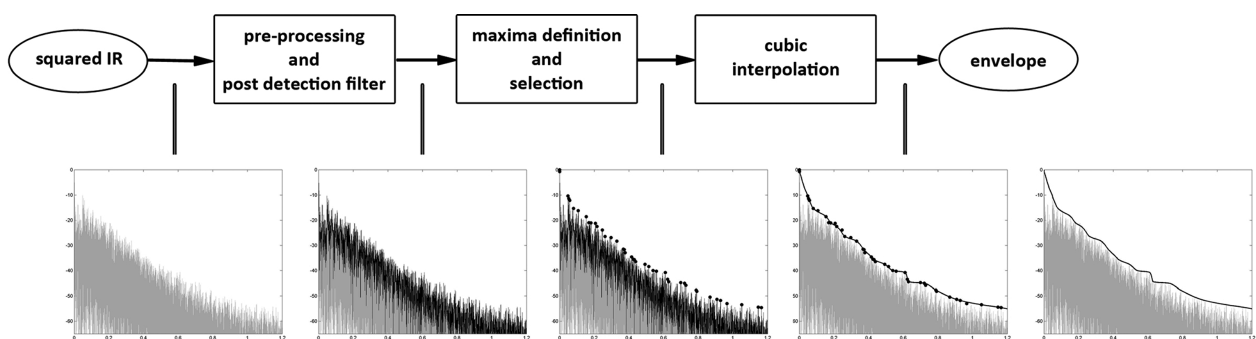


FIG. 2. Flow chart of the pre-processed energy detection method (PPED) for the evaluation of the envelope of the impulse response.

Some of them are composed by test signals synthesized using an exponentially decreasing shaped white noise, others by impulse responses measured in real rooms. It is well known that a real impulse response may be split into two regions: the early reflections, in which the acoustic field may be assumed deterministic and multi-modal, and the late reverberation, in which the acoustic field may be assumed Gaussian and mono-modal. The two regions have a temporal split point, the *mixing time*, that in real IRs corresponds to a point in the decay curve located above the -5 dB point. In case of synthesized IRs[51] this point is located at the 0 dB point of the decay (the mixing time corresponds to the flight time) and therefore the synthesized IRs are not useful to appreciate the differences between EDT and RT. Thus, in order to evaluate the performance of the PPED method, only measured IRs displaying typically occurring decays were used.

From the octave band filtered IR, the decay curve has been extracted with two different methods:

(1) The compensated Schroeder's method in the version presented by Guski and Vorländer[17] (method E: truncation, correction and subtraction), using the MATLAB ITA-Toolbox[52]);
(2) The pre-processed energy detection method (PPED).

A linear fitting was performed on the decay curves extracted with the two method in order to calculate three acoustic criteria according to ISO 3382-1:[5] EDT, T_{15}, and T_{30}. Though these criteria were originally defined over the energy decay curve, in the following they are calculated also over the envelope. This is somehow an extrapolation of the current definition of EDT, but it is the most effective way to permit a direct comparison between the two decay curves.

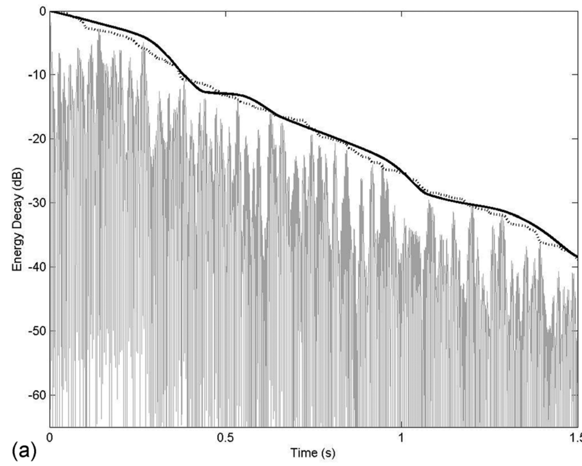
A. Test on a linear decay

Considering a IR with a theoretical exponential energy decay (linear when dB-scaled), it is expected that the two methods return similar results. In this work, a measured IR was chosen that displays an almost linear dB-scaled decay in order to get closer to this ideal case. It has been measured inside the Bayreuther Festspielhaus. Figure 3 shows the two decay curves at 125 and 1000 Hz: for this IR, displaying a linear decay, the two methods give similar decay curves, even if the PPED method returns a more detailed curve. In Table I the EDT, T_{15} and T_{30} calculated over the IR shown in Fig. 3 with the Schroeder's compensated method[17] are compared with the ones extracted with the PPED method.

The ratio between the difference of the values calculated with the two methods and the mean of the same values—for the sake of brevity called DR [Eq. (6)]—is also shown.

$$\mathrm{DR} = \frac{T_{\mathrm{PPED}} - T_{\mathrm{SCH}}}{\frac{T_{\mathrm{PPED}} + T_{\mathrm{SCH}}}{2}}. \quad (6)$$

As shown in Table I the DR values for T_{30} are smaller than the JND.



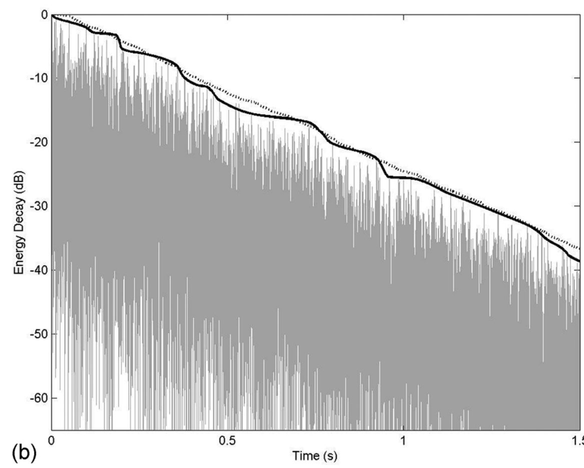


FIG. 3. Comparison of the two methods for the extraction of the decay curve. Gray curve: filtered test impulse response, measured inside the Bayreuther Festspielhaus. Pre-processed energy detection method (black solid line). Compensated Schroeder's integration method (dotted line). (a) IR filtered in the 125 Hz octave band. (b) IR filtered in the 1 kHz octave band.

The EDT concerns the first part of the decay curve, corresponding for about half of the fitted range to the early reflections. The T_{15} and T_{30} describe the part of the decay curve in which the acoustic field is diffuse. With the Schroeder's method, the different parts of the decay curve depend on each other, being all of them derived from a cumulative energy decay curve computed from the tail to the beginning of the impulse response [Eq. (4) and related comments in Sec. II]. In this way, the EDT values are related to the late part of the decay.[4] The local nature of the PPED method allows an independent definition of the different parts of the decay curve. In particular, the ability of the proposed method to identify the envelope of the early reflections permits to get a value of EDT representative of these reflections only, avoiding the influence of the late part of the decay. For the T_{15} and the T_{30}, mainly depending on a diffuse sound field condition, the decays extracted with the PPED method and the Schroeder's one may be assumed to be similar. Figure 4 reports (a) an IR measured inside the *Masini* theatre in

TABLE I. Reverberation time values calculated from a linearly decaying impulse response.

Octave band	Reverberation	Comp. Schroeder (Ref. 17)	PPED	DR
63 Hz	EDT (s)	2.81	3.55	0.23
	T_{15} (s)	1.87	2.35	0.20
	T_{30} (s)	2.36	2.51	0.06
125 Hz	EDT (s)	2.40	2.94	0.20
	T_{15} (s)	2.39	2.39	0.00
	T_{30} (s)	2.36	2.41	0.02
250 Hz	EDT (s)	2.68	2.03	−0.27
	T_{15} (s)	2.79	2.87	0.03
	T_{30} (s)	2.65	2.62	−0.01
500 Hz	EDT (s)	2.38	1.87	−0.24
	T_{15} (s)	2.61	2.33	−0.11
	T_{30} (s)	2.60	2.67	0.03
1 kHz	EDT (s)	2.34	2.60	0.10
	T_{15} (s)	2.43	2.44	0.00
	T_{30} (s)	2.44	2.47	0.01
2 kHz	EDT (s)	2.19	1.25	−0.55
	T_{15} (s)	2.13	2.27	0.06
	T_{30} (s)	2.18	2.13	−0.02
4 kHz	EDT (s)	1.87	1.43	−0.26
	T_{15} (s)	1.85	1.84	−0.01
	T_{30} (s)	1.84	1.78	−0.03

Faenza (Italy), and (b) the same IR where the direct component has been suppressed. The initial slope of the energy decay curve, built using the backward integration, is very similar in both cases. In fact, the EDT values calculated over the Schroeder's integral are almost identical: 1.09 s [Fig. 4(a)] and 1.10 s [Fig. 4(b)], respectively, even though the second IR lacks the direct sound field. The PPED method detects the two very different slopes and returns the values of 0.12 s [Fig. 4(a)] and 0.91 s [Fig. 4(b)].

If the useful decay range is not critical, it is expected that, decreasing the decay range of evaluation, the differences between the two methods increase. This is due to the fact that the PPED method returns an envelope which effectively describes the first part of the decay, independently from the latter part of it. Thus, reducing the evaluation range, the first part of the decay assumes a greater relative weight on the decay rate and the two methods return RT values that differ increasingly, as confirmed by the results shown in Table I and in Fig. 5.

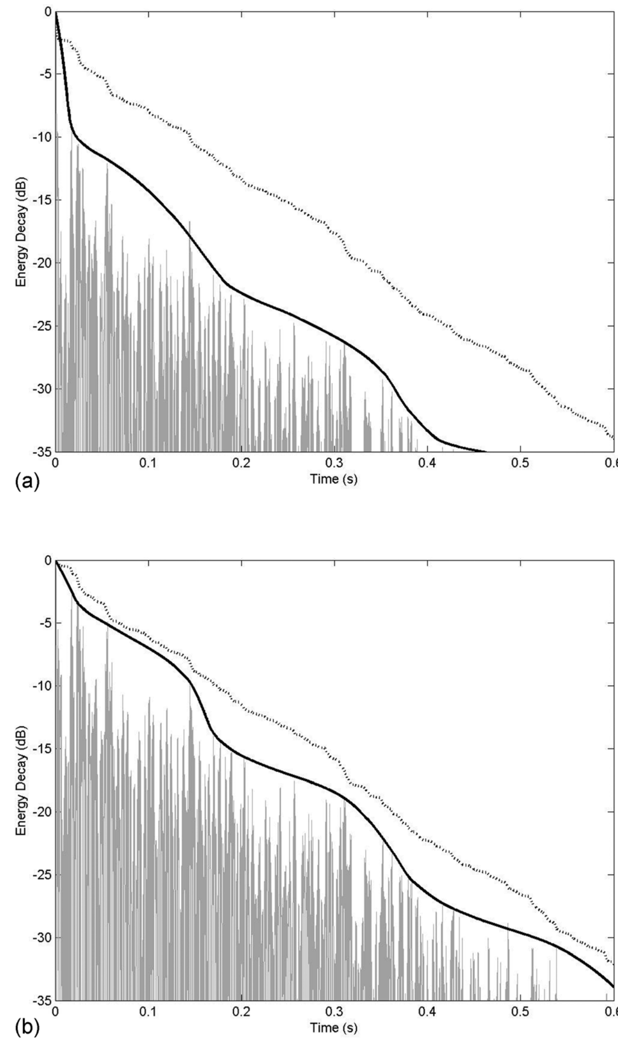


FIG. 4. Comparison of the energy decay curve and the envelope calculated on the same IR, measured in the *Masini* theatre in Faenza (Italy), unmodified and with the direct component suppressed. Filtered impulse response (gray curve). Pre-processed energy detection method (black solid line). Compensated Schroeder's integration method (dotted line). (a) Measured IR filtered at 500 Hz. (b) Measured IR filtered at 500 Hz with the direct component suppressed.

B. Test on other common decay types

The comparison of the compensated Schroeder's method[17,52] and the PPED method was further extended to different types of impulse responses.

In particular three IRs have been selected where the effect of early reflections returns different types of decay[4] in specific octave bands: the *cliff-type* decay, the *plateau-type* decay and the *sagging* decay. The same decays were identified by Hak *et al.*[7] in a different context. The first is a typical IR measured in a Italian opera house, the *Alighieri* theatre in Ravenna (Italy), having strong and distinct early reflections. For the *plateau-type*, an IR measured in the Bayreuther Festspielhaus with the source positioned in the orchestra pit was selected: it is a typical case where the early reflections are very weak or absent. Also the third IR (*sagging* decay) was measured in a Italian opera house, the *Bonci* theatre in Cesena (Italy); although this kind of decay is not typical of these theatres, it can occur for particular source-receiver combinations.

Figure 6 shows the three selected IRs and the decay curves extracted with the compensated Schroeder's method (dotted line) and with the PPED method (solid line). Table II reports the respective values of EDT.

Very different EDT values are obtained. As expected, the PPED method is very sensitive to early reflections and gives EDT values closely matching the very first part of the IR, as shown in Fig. 6. The fact that the envelope is "locally" defined implies that, relatively to the Schroeder's method, the new method returns greater EDT values in the case of

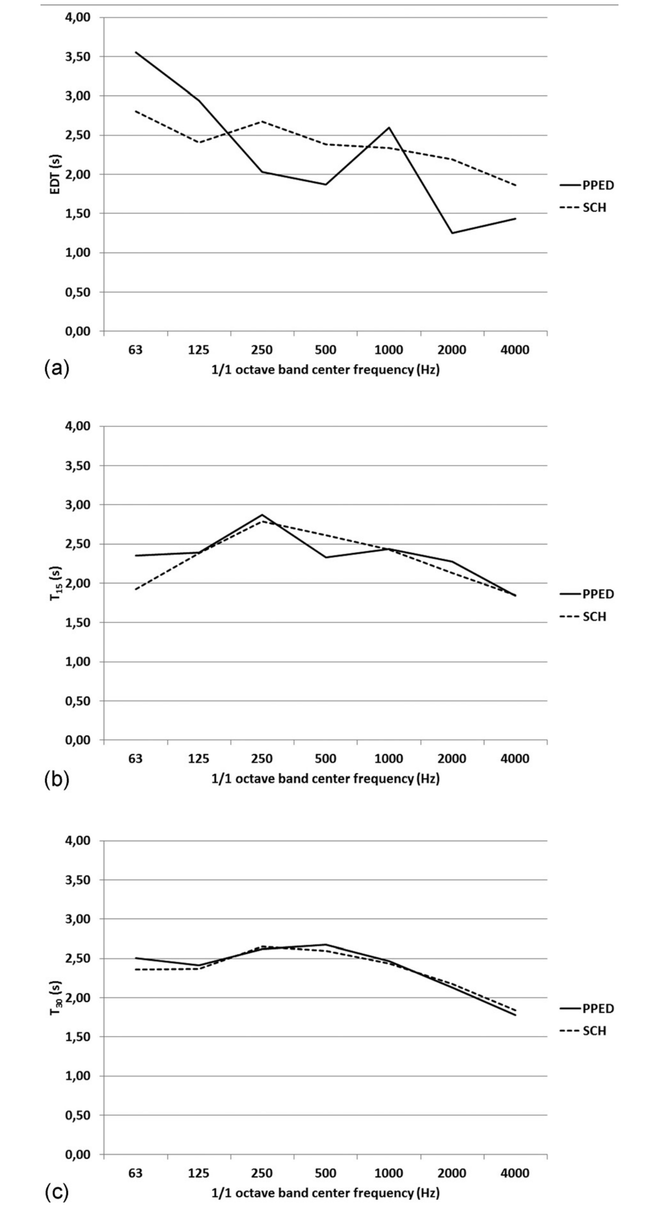


FIG. 5. Reverberation time values for the selected IR with linear energy decay (see Fig. 3). Pre-processed energy detection method (solid line). Compensated Schroeder's integration method (dotted line). (a) EDT. (b) T_{15}. (c) T_{30}.

plateau-type decay and smaller in the other cases. The differences between the two aforementioned methods are more evident in case of the *cliff-type* decay, where a very abrupt change in slope occurs in the first part of the decay.

V. LISTENING TESTS

In order to verify the subjective relevance of the PPED method, i.e., its ability to return EDT values closely correlated with the perceived reverberance, a listening test was performed in the Acoustic Laboratory of the University of Bologna (see Appendix B). The format chosen to spot the difference between the EDT calculated with the Schroeder integral and with the PPED method is a one-sided pairwise comparison test according to ISO 5495.[53]

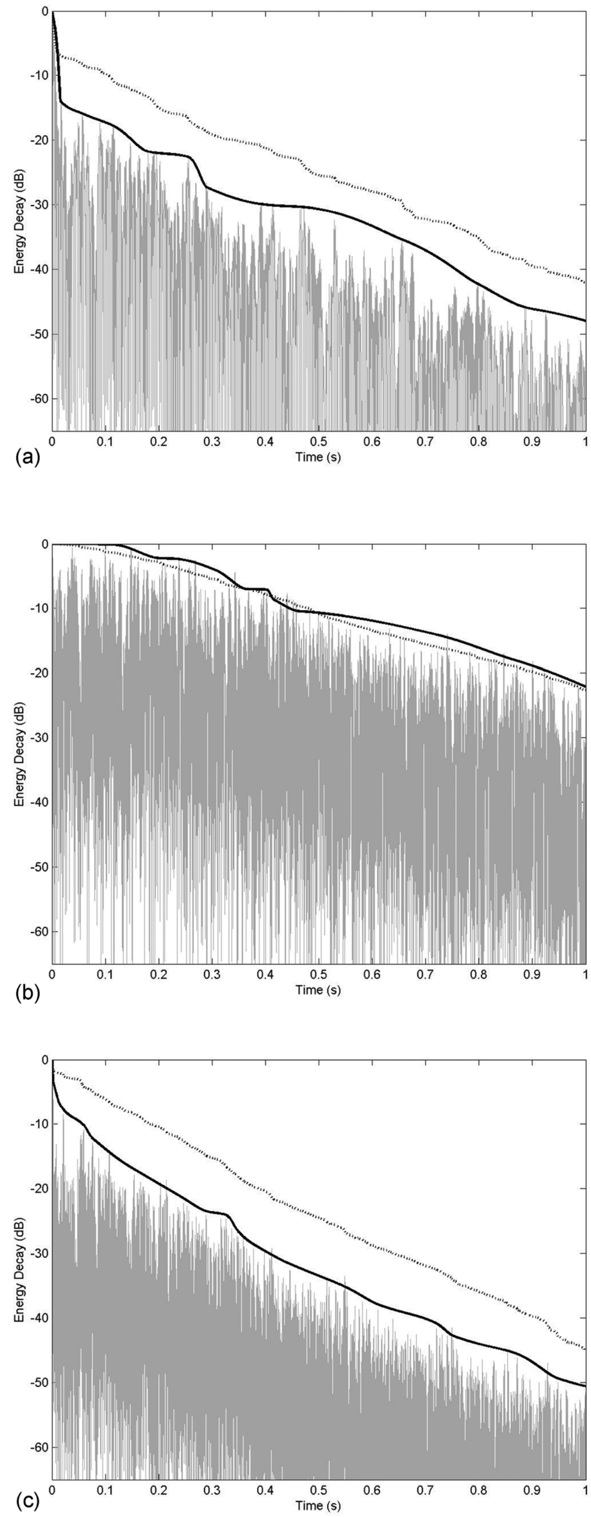


FIG. 6. Selected IRs with different type of decay due to the effect of early reflections (gray curve). Pre-processed energy detection method (solid line). Compensated Schroeder's integration method (dotted line). (a) *Cliff-type* decay. IR measured in the *Alighieri* theatre in Ravenna (Italy), filtered in the 250 Hz octave band. (b) *Plateau-type* decay. IR measured in the Bayreuther Festspielhaus, filtered in the 500 Hz octave band. (c) *Sagging* decay. IR measured in the *Bonci* theatre in Cesena (Italy), filtered in the 1 kHz octave band.

Six couples of stimuli were presented to 74 assessors in listening room compliant with ITU-R BS 1116-1.[56] Five out of six stimuli gave a correct response, thus these preliminary

TABLE II. EDT values calculated for the three types of IRs identified by Barron (Ref. 4).

Decay	Comp. Schroeder	PPED	Difference
Cliff-type (s)	1.27	0.09	−1.19
Plateau-type (s)	2.75	3.21	0.45
Sagging (s)	1.27	0.3	−0.97

results show that with a significance level of 5% it is possible to declare that the difference between the EDT values detected by the PPED method exists.

See Appendix B for details.

VI. CONCLUSIONS

Starting from the analysis of the previous literature concerning the methods for the evaluation of the energy decay curve or the envelope of an impulse response, a new extraction method, the pre-processed energy detection (PPED), has been proposed and implemented in a MATLAB toolbox. Its performance has been systematically tested versus the most accredited literature method: the Schroeder's backward integration method with noise subtraction, truncation and correction.[17]

The PPED method includes four steps. It extracts an envelope "locally matched" to the given impulse response which is different from the Schroeder's energy decay curve (compensated or not). While the energy decay curve represents a sort of energy average of the impulse response, the envelope is a curve tangent to the local maxima of the squared impulse response. Thus the decay rates (EDT, RT) extracted from the envelope are representative only of the decay range being evaluated. In the statistical part of the decay (the late part), the envelope and the energy decay curve seem very close to each other. In the deterministic part of the decay curve (the early part), the differences are more relevant. This can be related to the fact that the crest factor of the first part of the impulse response is not constant.

In case of linear decays, both methods provide a well-fitted description of the IR decay. But in other cases, differences between the two resulting decay curves can be found. In particular, when the early reflections are clearly distinguishable, like in Italian opera houses, or when they are nearly absent, like in the Bayreuther Festspielhaus when the sound source is inside the orchestra pit, there are remarkable differences between the EDT values calculated with the two methods. The proposed version of the energy detection method becomes relevant for the analysis of impulse responses with marked early reflections, where the detection of the different initial slope of the decay curve permits an evaluation of the early decay closely matching the local behavior of the IR. Listening tests have been performed in order to evaluate the perceptive relevance of the EDT calculated with the PPED method. The preliminary results showed that with a significance level of 5% the difference between the EDT extracted with the PPED method and with the Schroeder integral was correctly detected for five stimuli out of six.

APPENDIX A: PSEUDOCODE

Algorithm 1: PPED algorithm.

```
FcVect is the vector of the center frequencies of the octave band filters
RangeVect is the vector of the ranges to calculate the reverberation times
numIter is the number of iteration of the pre-processing block
DesignOctaveFilter is the function calculating the octave band filter
coefficients according to IEC 1260

procedure PPEDMAIN
  inIR ← READ(IRfile)
  for each Fc in FcVect do
    Hd ← DESIGNOCTAVEFILTER(Fc)
    inFilt ← APPLYFILTER(inIR, Hd)
    SQinFilt ← SQUARE(inFilt)
    OutTmp ← ENERGYDETECTION(SQinFilt, Fc, numIter)
    Env ← PEAKHERMITE(OutTmp, numIter)
    for each r in RangeVect do
      RT ← LINEARFIT(Env, RangeVect)
function ENERGYDETECTION(SQinFilt, Fc, numIter)
  InSharp ← SHARPENING(SQinFilt, numIter)
  LowPassF ← DESIGNFILTER(BlackmanHarris, Fc)
  Out ← APPLYFILTER(InSharp, LowPassF)
  OutTmp ← Out
  return OutTmp
function PEAKHERMITE(OutTmp, numIter)
  InSharp ← SHARPENING(OutTmp, numIter)
  MaxVal ← FINDMAX(InSharp)
  MaxValOK ← SELMAX(MaxVal)
  Env ← CUBICPIECEHERMITE(MaxValOK) % see (Frisch and Carlson[49])
  return Env
function SHARPENING(OutTmp, numIter)
  while n < numIter do
    initSample ← length(OutTmp)
    for Sample in OutTmp do
      if Sample < Sample − 1 then
        Sample ← Sample − 1
  return OutTmp
function SELMAX(MaxVal)
  initVal ← length(MaxVal)
  for Val in MaxVal do
    if Val > Val − 1 then
      MaxValOK ← Val
  return MaxValOK
```

APPENDIX B: PRELIMINARY RESULTS OF LISTENING TESTS

In order to verify the effectiveness of the PPED method, i.e., its ability to return EDT values closely correlated with the perceived reverberance, a listening test was performed in the Acoustic Laboratory of the University of Bologna. The format chosen to spot the difference between the EDT calculated with the Schroeder integral and with the PPED method is a one-sided pairwise comparison test.[53]

A. Setting

The stimuli presented to the participants consisted of measured impulse responses convolved with two short anechoic recordings:

TABLE III. Reverberation time values calculated from the selected impulse responses.

		IR_a		IR_b	
Couple	Reverberation	Comp. Schroeder (Ref. 17)	PPED	Comp. Schroeder (Ref. 17)	PPED
@ 500 Hz					
1–2	EDT (s)	1.55	1.51	1.55	**1.16**
	T_{30} (s)	1.65	1.60	1.60	1.59
3–4	EDT (s)	1.22	1.25	1.20	**0.79**
	T_{30} (s)	1.23	1.19	1.20	1.19
5–6	EDT (s)	1.82	1.75	1.81	**0.79**
	T_{30} (s)	1.68	1.68	1.66	1.66
@ 1000 Hz					
7–8	EDT (s)	1.13	1.14	1.11	**0.73**
	T_{30} (s)	1.11	1.07	1.11	1.06
9–10	EDT (s)	1.32	1.34	1.25	**0.81**
	T_{30} (s)	1.11	1.08	1.10	1.09
11–12	EDT (s)	1.32	1.34	1.30	**0.17**
	T_{30} (s)	1.11	1.08	1.11	1.11

(1) Clarinet solo from Mozart's opera Don Giovanni, 15 s duration (MOcla);
(2) Clarinet solo from Bruckner's Symphony no.8, II movement, 12 s duration (BRcla).

These two anechoic motifs were chosen in the set recorded at Aalto University[54] considering their spectral characteristics and their τ_e value. In particular, the constraints that were considered are

(1) Short τ_e value.[55]
(2) Energy mainly concentrated in one octave band, in order to correlate directly the calculated values of EDT to the perceived reverberation in the correspondent octave band.

The first anechoic recording (MOcla) presents most part of its energy in the octave band of 500 Hz, while the energy of the second anechoic recording (BRcla) is concentrated in the octave band of 1000 Hz.

The impulse responses were measured in two Italian opera houses: the Masini Theatre in Faenza and the Goldoni Theatre in Bagnacavallo. Six couples of impulse responses (IR_a, IR_b) have been selected using these constraints relative to the target octave band:

(1) Same reverberation time T_{30} for both IR_a and IR_b calculated with the Schroeder integral.
(2) Same reverberation time T_{30} for both IR_a and IR_b calculated with the PPED method.
(3) Same early decay time EDT for both IR_a and IR_b calculated with the Schroeder integral.
(4) Different (lower) EDT only for one IR_b calculated with the PPED method, while the EDT value for IR_a is the same as the one calculated with the Schroeder integral.

The values of T_{30} and EDT were considered equal when the JND thresholds (5%) were not exceeded. In Table III all the values are shown.

As previously explained, the IRs couples 1–2, 3–4, and 5–6 have been convolved with the first anechoic recording (MOcla), while the other three couples 7–8, 9–10, and 11–12 with the second anechoic recording (BRcla).

B. Participants

The type I error (α) and the type II error (β) were fixed at 0.05 and the percentage of assessors detecting the difference p_d at 40% For these parameters ISO 5495[53] recommended a minimum number of assessors of 67. The authors recruited 74 assessors. They were 43 males and 31 females, aged 16 to 56, the mean age being 32. None of them ever participated to sensory tests. Their attitude toward music and audio quality, self-rated on a five-point scale, shows that a comparable percentage of assessors located at level 4 or 5 (high or very high) and 1 or 2 (absent or poor). Five of them are professional musicians and the other have a moderate confidence with music. Two assessors declared to have hearing impairments: they were not considered in the final results.

C. Procedure

Stimuli were presented in a listening room compliant with ITU-R BS 1116[56] and having a measured reverberation time of 0.2 s in the 1 kHz octave band. Preliminarily, a training session was proposed to the assessors. The three trials were ordered by increasing difficulty in telling the attribute; the first two motifs had very different reverberation time, while the third couple respected the same constraints of the test. At each assessor six couples of stimuli were proposed and for each couple it was asked: *Which of the two stimuli*

TABLE IV. Number of correct responses of the listening test for each stimulus fed.

Anechoic	IR_{a-b}	Number correct	Number required	Passed
MOcla	1–2	54	44	Yes
	3–4	48	44	Yes
	5–6	52	44	Yes
BRcla	7–8	39	44	No
	9–10	65	44	Yes
	11–12	62	44	Yes

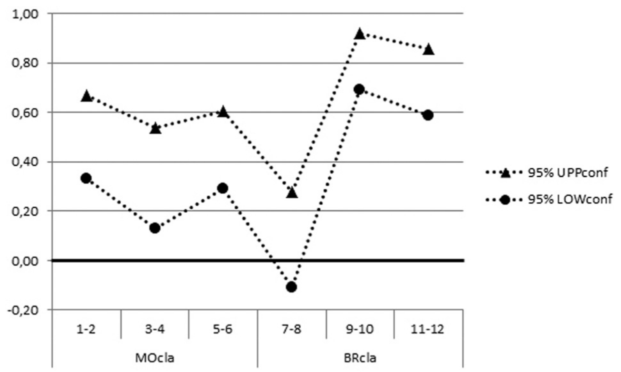


FIG. 7. Results of the listening tests: upper and lower confidence interval for each stimulus fed.

presents greater reverberance? Calling A the stimulus with IR_a and B the stimulus with IR_b, at each assessor was proposed a mixture of AB and BA couples of stimuli.

D. Results

In this section, *correct response* corresponds to the case when the assessor declares to perceive a greater reverberance for the motif with greater EDT value calculated with the PPED method (EDT values calculated with the Schroeder integral were equal).

Considering the total amount of responses, the test returns 74% of correct responses that is to compare with the minimum percentage of 61% of correct responses required to conclude that a perceptible difference exists.[53]

Considering each stimulus separately, the number of correct responses (among a total of 72 assessors) are shown in Table IV. Figure 7 shows the confidence interval separately for each stimulus. Both of them show that five out of six stimuli passed the test.

Thus, these preliminary results show that with a significance level of 5% it is possible to declare that the difference between the EDT values detected by the PPED method exists.

[1]B. S. Atal, M. R. Schroeder, and G. M. Sessler, "Subjective reverberation time and its relation to sound decay," in *Proceedings of the 5th ICA*, Liege, Paper G32 (1965).
[2]M. R. Schroeder, B. S. Atal, G. M. Sessler, and J. E. West, "Acoustical measurements in philharmonic hall (New York)," J. Acoust. Soc. Am. **40**(2), 434–440 (1966).
[3]V. L. Jordan, "Acoustical criteria for auditoriums and their relation to model techniques," J. Acoust. Soc. Am. **47**(2), 408–412 (1970).
[4]M. Barron, "Interpretation of early decay time in concert auditoria," Acustica **81**, 320–331 (1995).
[5]ISO 3382-1:2009: *Acoustics—Measurement of room acoustic parameters* (International Organisation for Standardization, Geneva, Switzerland, 2009).
[6]M. Barron and L. J. Lee, "Using the standard on objective measures for concert auditoria, ISO3382, to give reliable results," Acoust. Sci. Technol. **26**, 162–169 (2005).
[7]C. C. J. M. Hak, R. H. C. Wenmaekers, and L. C. J. van Luxemburg, "Measuring room Impulse responses: Impact of the decay range on derived room acoustic parameters," Acta Acust. Acust. **98**, 907–915 (2012).
[8]A. Cocchi, M. Garai, and C. Tavernelli, "Boxes and sound quality in an Italian opera house," J. Sound Vib. **232** 171–191 (2000).
[9]M. Garai, S. De Cesaris, and D. D'Orazio, "Spatial distribution of monaural acoustical descriptors in historical Italian theatres," in *Proceedings of ISRA 2013 International Symposium on Room Acoustics*, Toronto, Canada (2013), Vol. 81, pp. 1–13.
[10]M. Garai, F. Morandi, D. D'Orazio, S. De Cesaris, and L. Loreti, "Acoustic measurements in eleven Italian opera houses: Correlations between room criteria and considerations on the local evolution of a typology," Build. Environ., in press (2015).
[11]M. Schroeder, "New method of measuring reverberation time," J. Acoust. Soc. Am. **37**, 409–412 (1965).
[12]M. Schroeder, "Complementarity of sound buildup and decay," J. Acoust. Soc. Am. **40**, 549–551 (1966).
[13]W. T. Chu, "Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method," J. Acoust. Soc. Am. **63**(5), 1444–1450 (1978).
[14]Y. Hirata, "A method of eliminating noise in power response," J. Sound Vib. **82**(4), 593–595 (1982).
[15]N. Xiang, "Evaluation of reverberation times using a nonlinear regression approach," J. Acoust. Soc. Am. **98**, 2112–2121 (1995).
[16]A. Lundeby, T. E. Vigran, H. Bietz, and M. Vorländer, "Uncertainties of measurements in room acoustics," Appl. Acoust. **81**, 344–355 (1995).
[17]M. Guski and M. Vorländer, "Comparison of noise compensation methods for room acoustic impulse response evaluations," Acta Acoust. **100**, 320–327 (2014).
[18]Y. Tahara and T. Miyajima, "A new approach to optimum reverberation time characteristics," Appl. Acoust. **54**, 113–129 (1998).
[19]J. Nannariello and F. Fricke, "The prediction of reverberation time using neural network analysis," Appl. Acoust. **58**, 305–325 (1999).
[20]T. Cox, F. F. Li, and P. Darlington, "Extracting room reverberation time from speech using artificial neural networks," J. Audio Eng. Soc. **49**, 219–230 (2001).
[21]K. Lebart, J. M. Boucher, and P. N. Denbigh, "A new method based on spectral subtraction for speech dereverberation," Acta Acoust. **87**(3), 359–366 (2001).
[22]R. Ratnam, D. L. Jones, B. C. Wheeler, W. D. O'Brien, Jr., C. R. Lansing, and A. S. Feng, "Blind estimation of reverberation time," J. Acoust. Soc. Am. **114**(5), 2877–2892 (2003).
[23]R. Ratnam, D. L. Jones, and W. D. O'Brien, Jr., "Fast algorithm for blind estimation time," IEEE Signal Processing Lett. **11**(6), 537–540 (2004).
[24]S. Vesa and A. Harma, "Automatic estimation of the reverberation time from binaural signals," in *Proceedings of IEEE International Acoustics, Speech, and Signal Processing*, 18–23 March (2005), Vol. 3.
[25]F. Satoh, Y. Hidaka, and H. Tachibana, "Reverberation time directly obtained from squared impulse response envelope," in *Proceedings of the International Congress on Acoustics*, Seattle, WA (1998), pp. 2755–2756.
[26]P. Luizard and B. F. G. Katz, "Coupled volume multi-slope impulse responses: A quantitative analysis method," Proc. IOA **33**(2), 169–176 (2011).
[27]M. Meissner, "Evaluation of decay times from noisy room responses with pure-tone excitation," Arch. Acoust. **38**, 47–54 (2013).
[28]K. Bodlund, "On the use of the integrated impulse response method for laboratory reverberation measurements," J. Sound Vib. **55**(3), 341–362 (1978).
[29]J. L. Davy, "The variance of decay rates in reverberation rooms," Acustica **43**, 12–25 (1979).
[30]M. Vorländer and H. Bietz, "Comparison of methods for measuring reverberation time," Acustica **80**, 205–215 (1994).
[31]H. Kuttruff, "A simple iteration schemes for the computation of decay constants in enclosures with diffusely reflecting boundaries," J. Acoust. Soc. Am. **98**(1), 288–293 (1995).
[32]N. Xiang and P. M. Goggans, "Evaluation of decay times in coupled spaces: Bayesian parameter estimation," J. Acoust. Soc. Am. **110**(3), 1415–1424 (2001).
[33]N. Xiang and P. M. Goggans, "Evaluation of decay times in coupled spaces: Bayesian decay model selection," J. Acoust. Soc. Am. **113**, 2685–2697 (2003).
[34]N. Xiang, P. M. Goggans, T. Jasa, and M. Kleiner, "Evaluation of decay times in coupled spaces: Reliability analysis of Bayesian decay time estimation," J. Acoust. Soc. Am. **117**(6), 3707–3715 (2005).
[35]W. D. Joyce, "Sabine's reverberation time and ergodic auditoria," J. Acoust. Soc. Am. **58**, 643–655 (1975).
[36]O. Legrand and D. Sornette, "Test of Sabine's reverberation time in ergodic auditoria within geometrical acoustics," J. Acoust. Soc. Am. **88**, 865–870 (1990).
[37]J. D. Polack, "Modifying chambers to play Billiards: The foundation of reverberation theory," Acustica **76**, 257–272 (1992).

[38]J. M. Jot, L. Cerveau, and O. Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *Proceedings of the 103rd AES Convention*, New York (1997).
[39]W. C. Sabine, "Collected papers on acoustics" (Harvard University Press, Cambridge, 1922).
[40]H. Arau-Puchades, "An improved reverberation formula," Acustica **65**, 163–180 (1988).
[41]M. Karjalainen, P. Antsalo, A. Makivirta, T. Peltonen, and V. Valimaki, "Estimation of modal decay parameters from noisy response measurements," J. Audio Eng. Soc. **50**, 867–879 (2002).
[42]H. Kuttruff and M. Jusofie, "Nachhallmessungen nach dem Verfahren der integrierten Impulsantwort," Acustica **19**, 56–58 (1967).
[43]M. Schroeder, "Statistical parameters or the frequency response curves of large rooms," Acustica **4**, 594–600 (1954).
[44]S. T. Neely and J. B. Allen, "Invertibility of a room impulse response," J. Acoust. Soc. Am. **66**(1), 165–169 (1979).
[45]R. Kumaresan and A. Rao, "Model-based approach to envelope and positive instantaneous frequency estimation of signals with speech application," J. Acoust. Soc. Am. **105**(3), 1912–1924 (1999).
[46]H. Urkowitz, "Energy detection of unknown deterministic signal," Proc. IEEE **55**, 523–531 (1967).
[47]D. D'Orazio, S. De Cesaris, and M. Garai, "A comparison of methods to compute the 'effective duration' of the autocorrelation function and an alternative proposal," J. Acoust. Soc. Am. **130**(4), 1954–1961 (2011).
[48]D. D'Orazio, S. De Cesaris, and M. Garai, "Measuring reverberation time using preprocessed energy detection," in *Proceedings of Internoise 2012*, New York (2012).
[49]F. N. Frisch and R. E. Carlson, "Monotone piecewise cubic interpolation," SIAM J. Numer. Anal. **17**(2), 238–246 (1980).
[50]B. F. G. Katz, "International round robin on room acoustical impulse response analysis software," Acoust. Res. Lett. **5**, 158–164 (2004).
[51]K. Hoshi, H. Okubo, J. Kanda, T. Asakura, A. Marui, H. Miyazaki, M. Yairi, and Y. Yasuda, "A study on establishing benchmarks for acoustical parameters derived from impulse responses," in *Proceedings of International Symposium on Room Acoustics*, 9–11 June, Toronto (2013).
[52]P. Dietrich, M. Guski, M. Pollow, B. Masiero, M. Müller-Trapet, R. Scharrer, and M. Vorländer, "ITA-toolbox—An open source MATLAB toolbox for acousticians," in *Fortschritte der Akustik – DAGA* (2012). pp. 151–152.
[53]ISO 5495:2007: *Sensory analysis—Methodology—Paired comparison test* (International Organization of Standardization, Geneva, Switzerland, 2007).
[54]J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," Acta Acust. Acust. **94**(6), 856–865 (2008).
[55]S. Sato, H. Sakai, and N. Prodi, "Subjective preference for sound sources located on the stage and in the orchestra pit of opera house," J. Sound Vib. **258**(3), 549–561 (2002).
[56]ITU-R BS 1116-1:1997: "Methods for the subjective assessment of small impairments in audio system including multichannel sound systems," (International Telecommunication Union, Geneva, Switzerland, 1997).