Audio Engineering Society
Convention Paper 5290
Presented at the 110th Convention
2001 May 12–15 Amsterdam, The Netherlands

Estimation of Modal Decay Parameters from Noisy Response Measurements

Matti Karjalainen[1], Poju Antsalo[1], Aki Mäkivirta[2], Timo Peltonen[3], and Vesa Välimäki[1]

[1]Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, P.O. Box 3000, FIN-02015 HUT, Finland

[2]Genelec Oy, Iisalmi, Finland, and [3]Akukon Oy, Helsinki, Finland

ABSTRACT
Estimation of modal decay parameters from noisy measurements of reverberant and resonating systems is a common problem in audio and acoustics, e.g., in room and concert hall measurements or musical instrument modeling. In this paper, reliable methods to estimate the initial response level, decay rate, and noise floor level from noisy measurement data are studied and compared. A new method, based on nonlinear optimization of a model for exponential decay plus stationary noise floor, is presented. Comparison with traditional decay parameter estimation techniques using simulated measurement data shows that the proposed method outperforms in accuracy and robustness, especially in extreme SNR conditions. Three cases of practical applications of the method are demonstrated.

0 INTRODUCTION

Parametric analysis, modeling, and equalization (inverse modeling) of reverberant and resonating systems find many applications in audio and acoustics. These include room and concert hall acoustics, resonators in musical instruments, and resonant behavior in audio reproduction systems. Estimation of reverberation time or modal decay rate are important measurement problems in room and concert hall acoustics [1], where S/N ratios of only 30-50 dB are common. The same problems can be found for example in the estimation of parameters in model-based sound synthesis of musical instruments, such as vibrating strings or body modes of string instruments [2]. Reliable methods to estimate parameters from noisy measurements are thus needed.

In an ideal case of modal behavior, after a possible initial transient, the decay is exponential until a steady state noise floor is encountered. The parameters of primary interest to be estimated are:

- Initial level of decay (L_{I})
- Decay rate/time or reverberation time (T_{D})
- Noise floor level (L_{N})

In a more complex case there can be two or more modal frequencies, whereby the decay is not simple anymore, but shows additional fluctuation (beating) or a two-stage (or multiple-stage) decay behavior. In a diffuse field (room acoustics) the decay of a noise-like response is approximately exponential in rooms with compact geometry. The noise floor may also be non-stationary. In this article we primarily discuss a simple mode (i.e., a complex conjugate pole pair in transfer function) or a dense set of modes with exponential reverberant decay, together with a stationary noise floor.

Methods presented in literature and common-sense or ad-hoc methods will first be reviewed. Techniques based on energy-time curve analysis of signal envelope are known as methods where the noise floor can be found and estimated explicitly. Backwards integration of energy, so called Schroeder integration [3]-[4], is often applied first to obtain a smoothed envelope for decay rate estimation. AR modeling of modes by estimating the transfer function poles and group delay analysis are examples of straightforward methods which are not particularly robust against background noise.

The effect of background noise floor is known to be problematic, and techniques have been developed to compensate the effect of envelope flattening when the noise floor in a measured response is reached, including limiting the period of integration [5], subtracting an estimated noise floor energy level from a response [6], or using two separate measurements to reduce the effect of noise [7]. The iterative method by Lundeby *et al.* [8] is of particular interest since it addresses with care the case of noisy data. This technique, as most other methods, analyzes the initial level L_{I}, decay time T_{D}, and noise floor L_{N} parameters separately, typically starting from a noise floor estimate. Iterative procedures are common in accurate estimation.

A different approach was taken by Xiang [9] where a parametrized signal-plus-noise model is fitted to Schroeder-integrated measurement data by searching for a least squares (LS) optimal solution. In this study we have elaborated a similar method of nonlinear LS optimization further to make it applicable to a wide range of situations, showing good convergence properties. A specific parameter and/or a weighting function can be used to further fine-tune the method for specific problems. The technique is compared with the Lundeby *et al.* method by applying them to simulated cases of exponential decay plus stationary noise floor where the exact parameters are known. The improved nonlinear optimization technique is found to outperform traditional methods in accuracy and robustness, particularly in difficult conditions with extreme signal-to-noise ratios.

Finally, the applicability of the improved method is demonstrated by three examples of real measurement data: (a) reverberation time of a concert hall, (b) low-frequency mode analysis of a room, and (c) parametric analysis of guitar string behavior for model-based sound synthesis. Possibilities for further generalization of the technique to more complex problems, such as two-stage decay, will be discussed briefly.

1 DEFINITION OF PROBLEM DOMAIN

A typical property of resonant acoustic systems is that their impulse response is a decaying function after a possible initial delay and the onset. In the simplest case the response of a single mode resonator system is

$$h(t) = Ae^{-\tau(t-t_0)}\sin[\omega(t-t_0)+\phi]u(t-t_0) \tag{1}$$

where $u(t-t_0)$ is a step function with value 1 for $t \geq t_0$ and zero elsewhere, A is initial response level, t_0 response latency for example due to propagation delay of sound, τ decay rate parameter, $\omega = 2\pi f$ angular frequency, and ϕ initial phase of sinusoidal response. In practical measurements, when there are multiple modes in the system and noise (acoustic noise plus measurement system noise), a measured impulse response is of the form[1]

$$h(t) = \sum_{i=1}^{N} A_i e^{-\tau_i(t-t_0)}\sin[\omega_i(t-t_0)+\phi_i] + A_{\mathrm{n}}n(t) \tag{2}$$

where A_{n} is the rms value of background noise and $n(t)$ is unity level noise signal. Figure 1 illustrates a single delayed mode response corrupted by additive noise.

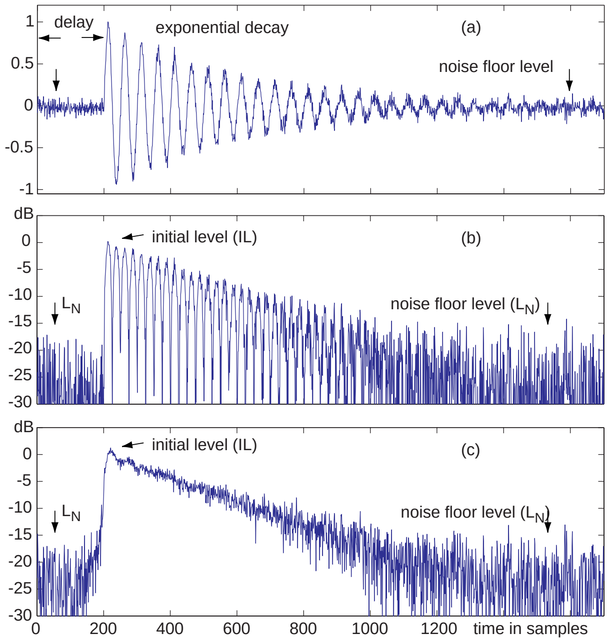

Fig. 1: (a) Single mode impulse response (sinusoidal decay) with initial delay and additive measurement noise. (b) Absolute value of response on the dB scale to illustrate the decay envelope. (c) Hilbert envelope, otherwise same as (b).

The problem of this study is defined as a task to find reliable estimates for the parameter set $\{A_i, \tau_i, t_0, \omega_i, \phi_i, A_{\mathrm{n}}\}$, given a noisy measured impulse response of the form (2). The main interest here is focused on systems of (a) separable single modes of type (1) including additive noise floor or (b) dense (diffuse, noiselike) set of modes resulting also in exponential decay similar to Fig. 1. In both cases the parameters of primary interest are A, τ, A_{n}, and t_0.

[1] In a more general case the initial delays may differ and there can be simple non-modal exponential terms, but these cases are of less importance here.

Often the decay time T_D is of main interest, for example in room acoustics where the reverberation time[2] of 60 dB decay, T_{60}, is related to τ by

$$T_{60} = -\frac{1}{\tau}\ln(10^{-3}) \approx \frac{6.908}{\tau} \quad (3)$$

Modern measurement and analysis techniques of system response are carried out by digital signal processing whereby discrete-time formulation for modal decay (without initial delay) with sampling rate f_s and sample period $T_s = 1/f_s$ becomes

$$h(n) = Ae^{-\tau_d n}\sin(\Omega n + \phi) \quad (4)$$

where n is sample index, $\tau_d = T_s\tau$, and $\Omega = 2\pi T_s f$.

2 DECAY PARAMETER ESTIMATION

In this section an overview of known techniques for decay parameter estimation will be presented. Initial delay and level estimation is discussed first briefly. The main problem, i.e., decay rate estimation, is the second topic. Methods to smooth the decay envelope from a measured impulse response are presented shortly. Noise floor estimation, an important subproblem, is discussed next. Finally, techniques for combined noise floor and decay rate estimation are reviewed.

2.1 Initial delay and initial level estimation

In most cases these two parameters are relatively easy to estimate. The initial delay may be short, not needing any attention, or the initial bulk delay can be cut off easily up to the edge of response onset. Only when the onset is relatively irregular or the S/N ratio is low, the detection of onset time moment can be difficult.

A simple technique to eliminate initial delay is to compute the minimum-phase component $h_{\mathrm{mphase}}(t)$ of the measured response [10]. An impulse response can be decomposed as a sum of minimum-phase and excess-phase components: $h(t) = h_{\mathrm{mphase}}(t) + h_{\mathrm{ephase}}(t)$. Since the excess-phase component will have allpass properties manifested as delay, computation of the minimum-phase part will remove the initial delay.

The initial level in the beginning of the decay can be detected directly from the peak value of the onset. For improved robustness, however, it may be better to estimate it from the matched decay curve, particularly its value at the onset time moment.

In the case of a room impulse response, the onset corresponds to direct sound from the sound source. It may be of special interest for the computation of source-to-receiver distance or in estimating the impulse response of the sound source itself by windowing the response before the first room reflection.

2.2 Decay rate estimation

Decay rate or time estimation is in practice based on fitting a line to the decay envelope, such as the energy-time curve, mapped on a logarithmic (dB) scale. Before computerized age this was done graphically on paper. The advantage of manual inspection is that an expert can avoid data interpretation errors in pathological cases. However, automatic determination of the decay rate or time is, however, highly desirable in practice.

Line fitting (linear regression to log envelope)

Fitting a line in a logarithmic decay curve is a conceptually and computationally simple way of decay rate estimation. The decay envelope $y(t)$ can be computed simply as a dB-scaled energy-time curve

$$y(t) = 10\lg\{x^2(t)\} \quad (5)$$

where $x(t)$ is the measured impulse response or a band-pass filtered part of it, such as an octave or 1/3-octave band. It is common to apply techniques such as Schroeder integration and Hilbert envelope computation (to be described below) in order to smooth the decay curve before line fitting. Least squares line fitting (linear regression) is done by finding the optimal decay rate k as

$$\min_{k,a}\int_{t_1}^{t_2}\{y(t) - [k(t) + a]\}^2 dt \quad (6)$$

for example using the Matlab function `polyfit` [11].

Practical problems with line fitting are related to the selection of interval $[t_1, t_2]$ and cases where the decay of the measured response is inherently nonlinear. The first problem is avoided by excluding onset transients in the beginning and noise floor biasing at the end of interval $[t_1, t_2]$. The second problem is related to such cases as a two-stage decay (initial decay rate or early reverberation and late decay rate or reverberation) or beating (fluctuation) of the envelope because of two modes very close in frequency (see Fig. 9b).

Nonlinear regression (Xiang's method)

Xiang [9] formulated a method where a measured and Schroeder-integrated energy-time curve is fitted to a parametric model of a linear decay plus a constant noise floor. Since the model is not linear in parameters, nonlinear curve fitting (nonlinear regression) is needed. Mathematically this is done by iterative means such as starting from a set of initial values for the model parameters and applying gradient descent to search for a least squares optimum

$$\min_{x_1,x_2,x_3}\int_{t_1}^{t_2}\left\{y_{\mathrm{sch}}(t) - [x_1e^{-x_2t} + x_3(L-t)]\right\}^2 dt \quad (7)$$

where $y_{\mathrm{sch}}(t)$ is the Schroeder-integrated energy envelope, x_1 is the initial level, x_2 the decay rate parameter, x_3 a noise floor related parameter, L the length of response, and (t_1, t_2) the time interval of nonlinear regression. Notice that the last term of noise floor effect is a descending line instead of a constant due to backward integration of the noise energy [9].

Nonlinear optimization is mathematically more complex than linear fitting and care should be taken to guarantee convergence. Even with convergence, the result may be only a local optimum, and generally the only way to know that a global optimum is found is to apply exhaustive search over possible value combinations of model parameters which, in a multi-parameter case, is often computationally too expensive.

Nonlinear optimization techniques will be studied in more detail later in this paper by introducing generalizations to the method of Xiang and by comparing the performance of different techniques in decay parameter estimation.

Autoregressive modeling (linear prediction)

For a single mode of Eq. (1) the response can be modeled as an impulse response of a resonating second-order all-pole filter.

[2] See reverberation time definition notes on page 10.

More generally, a combination of N modes can be modeled as a $2N$-order all-pole filter. Auto-regressive (AR) modeling [12] is a way to derive parameters for such a model. In many technical applications this method is called linear prediction [13]. For example the function `lpc` in Matlab [14] processes a signal frame through autocorrelation coefficient computation and solving the normal equations by Levinson recursion, resulting in a N^{th} order z-domain transfer function $1/(1+\sum_{i=1}^{N}\alpha_i z^{-i})$. Poles are obtained by solving the roots of the denominator polynomial. Each modal resonance appears as a complex-conjugate pole pair $(z_i,\, z_i^*)$ in the complex z-plane with angle $\phi = \arg(z_i) = 2\pi f/f_{\mathrm{s}}$ and radius $r = |z_i| = e^{-\tau/f_{\mathrm{s}}}$, where f is modal frequency, f_{s} is sampling rate, and τ is the decay parameter of the mode in Eq. (1).

Decay parameter analysis by AR modeling is not robust with long decay times (poles close to the unit circle) and background noise. Additive noise flattens the power spectrum of a modal resonance and moves the poles from unit circle towards the origin, thus resulting in shortened decay time estimates, which is contrary to the result in linear curve fitting. AR modeling works for decay parameter estimation only if it can spectrally resolve each mode. Thus it is not applicable to high modal densities, such as typical reverberation time measurements.

Group delay analysis

A complementary method to AR modeling is to use the group delay, i.e., phase derivative $T_{\mathrm{g}}(\omega) = -d\varphi(\omega)/d\omega$, as an estimate of the decay time for separable modes of an impulse response. While AR modeling is sensitive to power spectrum only, the group delay is based on phase properties only. For a minimum-phase single mode response the group delay at the modal frequency is inversely proportional to the decay parameter, i.e., $T_{\mathrm{g}} = 1/\tau$. Group delay computation is somewhat critical due to phase unwrapping needed, and the method is sensitive to measurement noise.

2.3 Decay envelope smoothing techniques

In the methods of linear or nonlinear curve fitting it is desirable to obtain a smooth decay envelope before the fitting operation. The following techniques are often used to improve the regularity of the decay ramp.

Hilbert envelope computation

In this method, signal $x(t)$ is first converted to an analytic signal $x_{\mathrm{a}}(t)$ so that $x(t)$ is the real part of $x_{\mathrm{a}}(t)$ and the complex part of $x_{\mathrm{a}}(t)$ is the Hilbert transform (90^{o} phase shift) [10] of $x(t)$. For a single sinusoid this results in an entirely smooth energy-time envelope. An example of Hilbert envelope for a noisy modal response is shown in Fig. 1c.

Schroeder integration

A monotonic and smoothed decay curve can be produced by 'backward integration' of the impulse response $h(t)$ over measurement interval $[0,T]$ and converting it to a logarithmic scale

$$L(t) = 10\lg\left(\frac{\int_t^T h^2(\tau)d\tau}{\int_0^T h^2(\tau)d\tau}\right) \text{ [dB]} \tag{8}$$

This process is commonly known as the Schroeder integration [3, 4]. Based on its superior smoothing properties it is used routinely in modern reverberation time measurements. A known problem with it is that if the background noise floor is included within the integration interval, the process produces a raised ramp that biases upwards the late part of decay. This is shown in Fig. 2 for the case of noisy single-mode decay of (a) for full response integration shown as curve (d).

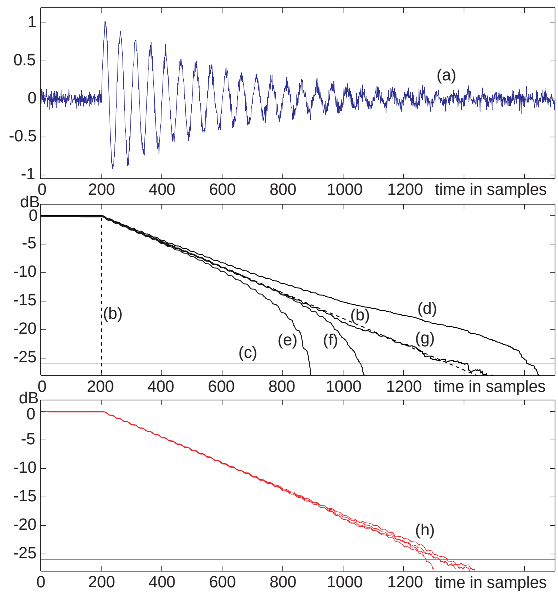

Fig. 2: Results of Schroeder integration applied to noisy decay of a mode: (a) measured noisy response including initial delay, (b) true decay of noiseless mode (dashed straight line), (c) noise floor (-26 dB), (d) Schroeder integration of total measured interval, (e) integration over a short interval (0,900), (f) integration over interval (0,1100), (g) integration after subtracting noise floor from energy-time curve, and (h) a few decay curves integrated by the Hirata's method.

The tail problem of Schroeder integration has been addressed by many authors, for example in [15, 8, 5, 6], and techniques to reduce slope biasing have been proposed. In order to apply these improvements, a good estimate of the noise floor level is needed first.

2.4 Noise floor level estimation

The limited signal-to-noise ratio inherent in practically all acoustical measurements, and especially measurements performed under field conditions, call for attention concerning the upper time limit of decay curve fitting or Schroeder integration. Theoretically this limit is set to infinity, but in practical measurements it is naturally limited to the length of the measured impulse response data. In practice, measured impulse responses must be long enough to accommodate for large enough dynamic range or the whole system decay down to the background noise level[3].

Thus the measured impulse response typically contains not only the decay curve under analysis, but also a steady level of background noise, which dominates for some time at the end of the response. Fitting the decay line over this part of envelope or Schroeder integrating this steady energy level along

[3]This is needed to avoid time aliasing with MLS and other cyclic impulse response measurement methods.

with the exponential decay curve causes an error both in the resulting decay rate (see Fig. 2) and in the time-windowed energies (i.e., energy parameters).

To avoid bias by noise, analysis must be performed on the impulse response data to find the level of background noise and the point where the room decay meets the noise level. This way it is possible to effectively truncate the impulse response at the noise level, minimizing the noise energy mixed with the actual decay.

The determination of noise floor level is difficult without using iterative techniques. The method by Lundeby *et al.* that will be outlined below is a good example of iterative techniques, integrated with decay rate estimation.

A simple way to have a reasonable estimate of background noise floor is to average a selected part of the measured response tail or to fit a regression line to it [16]. The level is certainly overestimated if the noise floor is not reached, but this is not necessarily problematic contrary to underestimating it. Another technique is to look at the background level before the onset of main response. This works if there is enough initial latency in the system response under study.

2.5 Decay estimation with noise floor reduction

In addition to determining the response starting point, it is thus essential to find an end point where the decay curve meets background noise, and to truncate the noise from the end of the response. Figure 2 illustrates the effect of limiting the Schroeder integration interval. If the interval is too short as in (e), the curve is biased downwards. Curve (f) shows a case where the bias due to noise is minimized by considering the decay only downto 10 dB above the noise floor.

There are no standardized exact methods for determining the limits for Schroeder integration and decay fitting or noise compensation techniques. Such methods are discussed below.

Limited integration or decay matching interval

There are several recommendations about dealing with noise floor and the point where the decay meets noise. For example, according to the ISO 3382 standard [1] for room reverberation determination, the noise floor must be 10 dB below the lowest decay level used for calculation of the decay slope. Morgan [5] recommends to truncate at the knee point and then to measure the decay slope of the backward integrated response down to a level about 5 dB above the noise floor.

Faiget *et al.* [16] propose a simple but systematic method for post-processing noisy impulse responses. The latter part of a response is used for the estimation of the background noise level by means of a regression line. Another regression line is used for the decay, and the end of the useful response is determined at the crossing point of the decay and background noise regression lines. The decay parameter fitting interval ends at 5dB above the noise floor.

Lundeby's method

Lundeby *et al.* [8] presented an algorithm for automatically determining the background noise level, the decay-noise truncation point, and the late decay slope of an impulse response. The steps of the algorithm are:

1. The squared impulse response is averaged into local time intervals in the range of 10–50 ms, to yield a smooth curve without losing short decays.
2. A first estimate for the background noise level is determined from a time segment containing the last 10 % of the impulse response. This gives a reasonable statistical selection without a large systematic error, if the decay continues to the end of the response.
3. The decay slope is estimated using linear regression between the time interval containing the response 0 dB peak, and the first interval 5–10 dB above the background noise level.
4. A preliminary crosspoint is determined at the intersection of the decay slope and the background noise level.
5. A new time interval length is calculated according to the calculated slope, so that there are 3–10 intervals per 10 dB of decay.
6. The squared impulse is averaged into the new local time intervals.
7. The background noise level is determined again. The evaluated noise segment should start from a point corresponding to 5–10 dB of decay after the crosspoint, or a minimum of 10 % of the total response length.
8. The late decay slope is estimated for a dynamic range of 10–20 dB, starting from a point 5–10 dB above the noise level.
9. A new crosspoint is found.

Steps 7–9 are iterated until the crosspoint is found to converge (max. 5 iterations).

Response analysis may be further enhanced by estimating the amount of energy under the decay curve after the truncation point. The measured decay curve is artificially extended beyond the point of truncation by extrapolating the regression line on the late decay curve to infinity. The total compensation energy is formed as an ideal exponential decay process, the parameters of which are calculated from the late decay slope.

Subtraction of noise floor level

Chu [15] proposed a subtraction method in which the mean square value of the background noise is subtracted from the original squared impulse response before the backward integration. Curve (g) in Fig. 2 illustrates this case. If the noise floor estimate is good and the noise is stationary, the resulting backward integrated curve is close to the ideal decay curve.

Hirata's method

Hirata [7] has proposed a simple method for improving the signal-to-noise ratio by replacing the squared single impulse response $h^2(t)$ with the product of two impulse responses measured separately at the same position:

$$\begin{aligned}\int_t^\infty h^2(t)dt \Leftarrow & \int_t^\infty [h_1(t)+n_1(t)][h_2(t)+n_2(t)]dt \\ = & \int_t^\infty [h_1(t)h_2(t)+h_1(t)n_2(t)+ \\ & + h_2(t)n_1(t)+n_1(t)n_2(t)]dt \\ = & \int_t^\infty \Big\{h_1(t)h_2(t)+ \\ & + n_1(t)n_2(t)[1+\frac{h_1(t)}{n_1(t)}+\frac{h_2(t)}{n_2(t)}]\Big\}dt \\ \approx & \int_t^\infty h^2(t)dt + K(t)\end{aligned} \tag{9}$$

The measured impulse responses consist of the decay terms $h_1(t)$, $h_2(t)$ and the noise terms $n_1(t)$, $n_2(t)$. The highly correlated decay terms $h_1(t)$ and $h_2(t)$ yield positive values corresponding to squared response $h^2(t)$, whereas the mutually uncorrelated noise terms $n_1(t)$ and $n_2(t)$ are seen as a random fluctuation $K(t)$ superposed on the first term.

Curves (h) in Fig. 2 illustrate a few decay curves obtained by backward integration with Hirata's method. In this simulated case they correspond to the case of (g), the noise floor subtraction technique.

Other methods

Under adverse noise conditions, a direct determination of the T_{30} decay curve from the squared and time-averaged impulse response has been noted to be more robust than the backward integration method (Satoh *et al.* [17]).

3 NONLINEAR OPTIMIZATION OF A DECAY-PLUS-NOISE MODEL

The nonlinear regression (optimization) method proposed by Xiang [9] was shortly described above. In the present study we have worked along similar ideas, using nonlinear optimization for improved robustness and accuracy. Below we introduce the nonlinear decay-plus-noise model and its application in several cases.

Let us assume that in noiseless conditions the system under study results in simple exponential decay of response envelope, corrupted by additive stationary background noise. We will study two cases that fit into the same model category. In the first case there is a single mode (i.e., a complex conjugate pole pair in transfer function) that in the time domain corresponds to an exponential decay function

$$h_{\mathrm{m}}(t) = A_{\mathrm{m}} e^{-\tau_{\mathrm{m}} t} \sin(\omega_{\mathrm{m}} t + \phi_{\mathrm{m}}) \qquad (10)$$

Here A_{m} is the initial envelope amplitude of the decaying sinusoidal, τ_{m} is a coefficient that defines the decay rate, ω_{m} is the angular frequency of the mode, and ϕ_{m} is the initial phase of modal oscillation.

The second case that leads to a similar formulation is where we have a high density of modes (diffuse sound field) with exponential decay, resulting in an exponentially decaying noise

$$h_{\mathrm{d}}(t) = A_{\mathrm{d}} e^{-\tau_{\mathrm{d}} t} n(t) \qquad (11)$$

where A_{d} is the initial rms level of response, τ_{d} is a decay rate parameter, and $n(t)$ is stationary Gaussian noise with rms level of 1 (= 0 dB).

In both Eqs. (10) and (11) we assume that a practical measurement of the system impulse response is corrupted with additive stationary noise

$$n_{\mathrm{b}}(t) = A_{\mathrm{n}} n(t) \qquad (12)$$

where A_{n} is the rms level of Gaussian measurement noise on the analysis bandwidth of interest, and it is assumed to be uncorrelated with the decaying system response. Statistically the rms envelope of the measured response is then

$$a(t) = \sqrt{h^2(t) + n_{\mathrm{b}}^2(t)} = \sqrt{A^2 e^{-2\tau t} + A_{\mathrm{n}}^2} \qquad (13)$$

This is a simple decay model that can be used for parametric analysis of noise-corrupted measurements. If the amplitude envelope of a specific measurement is $y(t)$, then optimized least squares (LS) error estimate for parameters $\{A, \tau, A_{\mathrm{n}}\}$ can be achieved by minimizing the following expression over a time span $[t_0, t_1]$ of interest

$$\min_{A,\tau,A_{\mathrm{n}}} \int_{t_0}^{t_1} [a(t) - y(t)]^2 dt \qquad (14)$$

Since the model of Eq. (13) is nonlinear in parameters $\{A, \tau, A_{\mathrm{n}}\}$, nonlinear LS optimization is needed to search for the minimum LS error.

By numerical experimentation with real measurement data it is easy to observe that LS fitting of the model of Eq. (14) places emphasis on large magnitude values, whereby noise floors well below the signal starting level are estimated poorly. In order to improve the optimization, a generalized form of model fitting can be formulated by minimization

$$\min_{A,\tau,A_{\mathrm{n}}} \int_{t_0}^{t_1} \Big\{ f(a(t), t) - f(y(t), t) \Big\}^2 dt \qquad (15)$$

where $f(y, t)$ is a mapping to balance weight for different envelope level values and time moments.

The choice of $f(y, t) = 20 \log(y(t))$ results in fitting on the decibel scale. It turns out that low-level noise easily has a dominating role in this formulation. A better result in model fitting can be achieved by using a power law scaling $f(y, t) = y^s(t)$ with factor $s < 1$, which is a compromise between amplitude vs. logarithmic scaling. A value of $s \approx 0.5$ has been found a useful default value[4].

A time-dependent part of mapping $f(y, t)$, if needed, can be separated as a temporal weighting function $w(t)$. A generalized form of the entire optimization is now to find

$$\min_{A,\tau,A_{\mathrm{n}}} \int_{t_0}^{t_1} \Big\{ w(t) a^s(t) - w(t) y^s(t) \Big\}^2 dt \qquad (16)$$

There is no clear physical motivation for the magnitude compression factor s. Specific temporal weighting functions $w(t)$ can be applied case by case, based on extra knowledge on the behavior of the system under study and goals of the analysis, such as focusing on the early decay time (early reverberation) of a room response.

The strengths of the nonlinear optimization method are apparent especially under extreme SNR conditions where all three parameters, $\{A, \tau, A_{\mathrm{n}}\}$, are needed with the best accuracy. This occurs both at very low SNR conditions where the signal is practically buried in background noise and at the other extreme where the noise floor is not reached within the measured impulse response but still an estimate of the noise level is desired. The necessary assumption for the method to work in such cases is that the decay model is valid, implying an exponential decay and stationary noise floor. Experiments show the method to work both for single mode decay and reverberant acoustic field decay models.

Figure 3 depicts three illustrative examples of decay model fitting of a single mode at an initial level of 0 dB and different noise floor levels. Because of simulated noisy responses it is easy to evaluate the estimation accuracy of each parameter. White curves show the estimated behavior of the decay-plus-noise model. In Fig. 3a the SN-ratio is only 6 dB. Errors in parameters in this case are a 0.5 dB underestimation of A,

[4] Interestingly enough, this is close to the *loudness* scaling in auditory perception known from psychoacoustics [18].

3.5 % underestimate in decay time related to parameter τ, and 1.8 dB overestimate of noise floor A_{n}. In Fig. 3b a similar case is shown with a moderate 30 dB SN-ratio. Estimation errors of parameters are +0.2 dB for A, -2.8 % for decay time, and +1.2 dB for A_{n}. In the third case of Fig. 3c the SN-ratio is -60 dB so that the noise floor is hardly reached within the analysis window. In this case the estimation errors are +0.002 dB for A, -0.07 % for decay time, and -1.0 dB for A_{n}. This shows that the noise floor is estimated with high accuracy also in this extreme case.

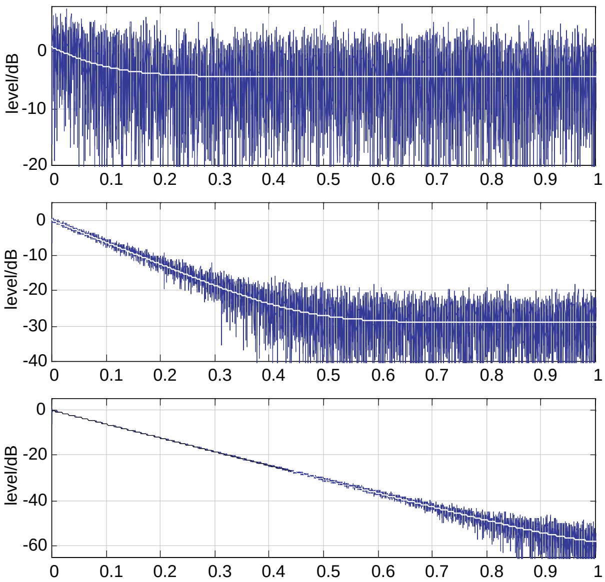

Fig. 3: Nonlinear optimization of decay-plus-noise model for three synthetic noisy responses with initial level of 0 dB and noise levels of (a) -6 dB, (b) -30 dB), and (c) -60 dB. Black curves show the Hilbert envelopes of the simulated responses and white curves depict the estimated decay behavior.

The nonlinear optimization used in this study is based on using the Matlab function `curvefit`[5] and the functions which implement the weighting by parameter s and weighting function $w(t)$ can be found at:
`http://www.acoustics.hut.fi/software/decay`

The optimization routines are found robustly converging in most cases, including such extreme cases as Fig. 3a and 3c, and the initial values of parameters for iteration are not critical. However, it is possible that in rare cases the optimization diverges and no (even a local) optimum is found[6]. It would be worth working out a dedicated optimization routine guaranteeing a result in minimal computation time.

Our experience in the nonlinear decay parameter fitting described here is that it still needs some extra information or top-level iteration for best results. It is advantageous to select the analysis frame so that the noise floor is reached not too early or late. If the noise floor is reached in the very beginning of the frame, the decay may be missed. Not reaching the noise floor in the frame is a problem only if the estimate of this level is important. A rule for an optimal value of the scaling parameter s is to use $s \approx 1.0$ for very low SN-ratios, such as in Fig. 3a, and let it approach a value of 0.4–0.5 when the noise floor is low as in 3c (see also Fig. 4 below).

4 COMPARISON OF DECAY PARAMETER ESTIMATION METHODS

The accuracy and robustness of methods for decay parameter estimation can be evaluated by using synthetic decay signals or envelope curves, computed for sets of parameters $\{A, \tau, A_{\mathrm{n}}\}$. By repeating the same for different methods, their relative performances can be compared. In this section we present results from the comparison of the proposed nonlinear optimization and the method of Lundeby *et al.*

The accuracy of the two methods was analyzed in the following setting. A decaying sinusoid of 1 kHz with a 60 dB decay time ('reverberation time') of 1 second was contaminated with white noise of Gaussian distribution and zero mean. Initial sinusoidal level to background noise ratio was varied from 0 dB to 80 dB by steps of 10 dB. Each method under study was applied to analyze the decay parameters and the error to the 'true' value was computed in decibels for the initial and the noise floor levels and as a percentage for decay time.

Figure 4 depicts the results of evaluation for the nonlinear optimization proposed in this paper. The accuracy of decay time estimation in Fig. 4a is excellent for SN-ratios above 30 dB and useful (below 10 % typically) even for SN-ratios of 0–10 dB. Initial level is accurate within 0.1 dB for SNR above 20 dB and about 1 dB for SNR of 0 dB. Noise floor estimate is within approximately 1–2 dB up to SNR of 60 dB and gives better than a guess up to 70–80 dB of SNR (notice that SNR alone is not important here but rather if the noise floor is reached in the analysis window or not).

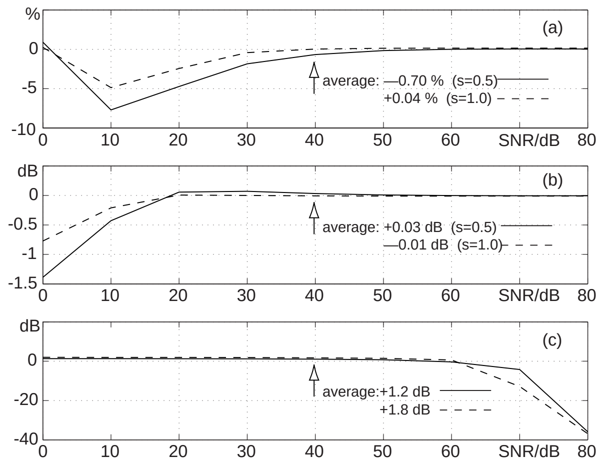

Fig. 4: Sine-plus-noise decay parameter estimation errors (average of 20 trials) for the nonlinear optimization method proposed in this paper as a function of SN-ratio: (a) decay time estimation error in percentage, (b) initial level estimation error in dB, (c) noise floor estimation error in dB. Solid line: $s = 0.5$, dotted line: $s = 1.0$.

[5] In new versions of Matlab, function `curvefit` is recommended to be replaced by function `lsqcurvefit`.
[6] Function `curvefit` also prints warnings of computational precision problems even when optimization results are excellent.

Figure 5 plots the same information for decay parameter estimation using the method of Lundeby *et al.* without noise compensation, implemented by us in Matlab. Since this iterative technique is not developed for extreme SN-ratios, such as 0 dB, it cannot deal with these cases without extra tricks, and even then it may have severe problems. We have used safety settings whereby it did not try to yield decay time values for SNR below 20 dB, and low SNR parts of the decay parameter estimate curves are omitted.

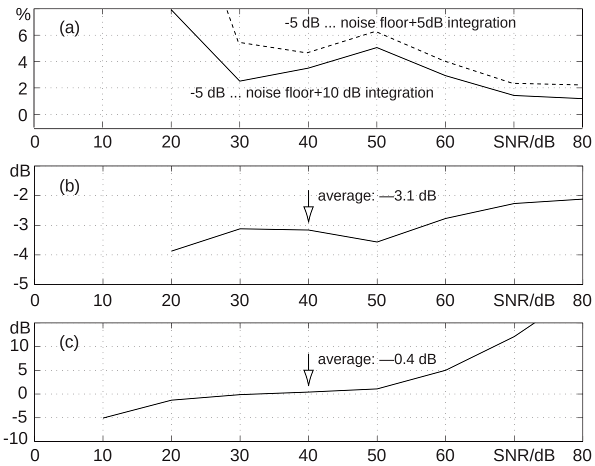

Fig. 5: Decay parameter estimation errors for the Lundeby *et al.* method as a function of SN-ratio: (a) decay time estimation error percentage with truncated Schroeder integration but without noise compensation, (b) initial level estimation error in dB, and (c) noise floor estimation error in dB.

For moderate SN-ratios the results of the method are fairly good and robust. Decay time shows a positive bias of few per cent except. The noise floor estimate is reliable in this case only up to about 50 dB SNR. Notice that the method is designed for practical reverberation time measurements rather than this test case where it could be tuned to perform better.

5 EXAMPLES OF DECAY PARAMETER ESTIMATION BY NONLINEAR OPTIMIZATION

In this section we present examples of applying the nonlinear estimation of a decay-plus-noise model to typical acoustic and audio applications including reverberation time estimation, analysis and modeling of low-frequency modes of a room response, and decay rate analysis of plucked string vibration for model-based synthesis applications.

5.1 Reverberation time estimation

Estimation of the reverberation time of a room or a hall is relatively easy if the decay curve behaves regularly and noise floor is low enough. Often in practice the case is quite different. Here we demonstrate the behavior of the nonlinear optimization method inan example where the measured impulse response includes an initial delay, irregular initial part, and a relatively high measurement noise floor. Figure 6 depicts three different cases of fitting the decay-plus-noise to this case of control room with a short reverberation time. in Fig. 6a the fitting is applied to the entire decay curve including the initial delay, and the resulting model is clearly biased towards too long reverberation time. In Fig. 6b the initiall delay is excluded from model fitting and the result is better. However, after the direct sound there is a period of only little energy due to the first reflections before the range of dnese reflections and diffuse response. If the reverberation time estimate should describe the decay of this diffuse part, the case of Fig. 6c with fitting, starting from about 30 ms, yields the best match to reverberation decay, and the approaching noise floor is also estimated well[7].

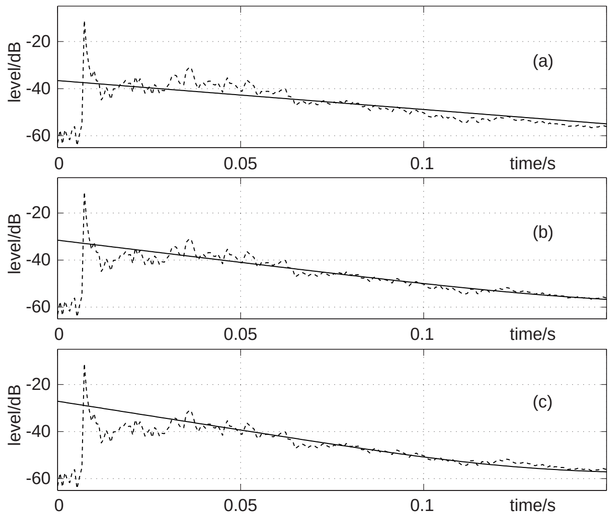

Fig. 6: Decay-plus-noise model fitting by nonlinear optimization to a room impulse response: (a) fitting range includes initial delay, transient phase, and decay, (b) fitting includes transient phase and decay, and (c) fitting includes only the decay phase.

5.2 Modeling of low-frequency room modes

The next case deals with the modeling of low-frequency modes of a room. Below a critical frequency (so-called Schroeder frequency) the mode density is low and individual modes can be decomposed from the measured room impulse response. The task here was to find most prominent modes and to analyze their modal parameters f_{m} and τ_{m}, frequency and decay parameter, respectively. The case studied was a hard-walled, partially damped room with moderate reverberation time ($\approx$ 1 sec) at mid and high frequencies, but much longer decay times at lowest modal frequencies. The following procedure was applied:

- A short-time Fourier analysis of the measured impulse response was computed to yield a time-frequency representation, shown in Fig. 7 as a waterfall plot.
- At each frequency bin (1.3 Hz spacing is used), the dB-scaled energy-time decay trajectory was fitted to the decay-plus-noise model with the nonlinear optimization technique to obtain the optimal decay parameter τ.

[7] In this example, decay parameter analysis is applied to the entire impulse response for demonstration purposes. In reality it should be done as a function of frequency, i.e., to octave or 1/3-octave band decay curves.

- Based on decay parameter values and spectral levels, a rule was written to pick up the most prominent modal frequencies and the related decay parameter values.

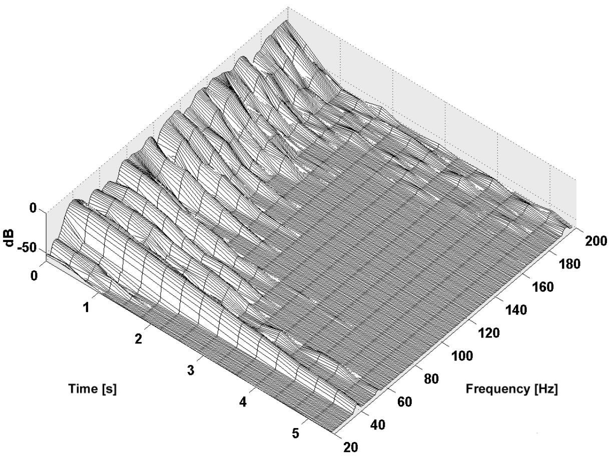

Fig. 7: Time-frequency plot of room response at low frequencies. Lowest modes (especially 40 Hz) show long decay times.

In this context we are interested in how well the decay parameter estimation works with noisy measurements. Application of the nonlinear optimization resulted in decay parameter estimates, some of which are illustrated in Fig. 8 by a comparison of the original decay and decay-plus-noise model behavior. For all frequencies in the vicinity of a mode the model fits robustly and accurately.

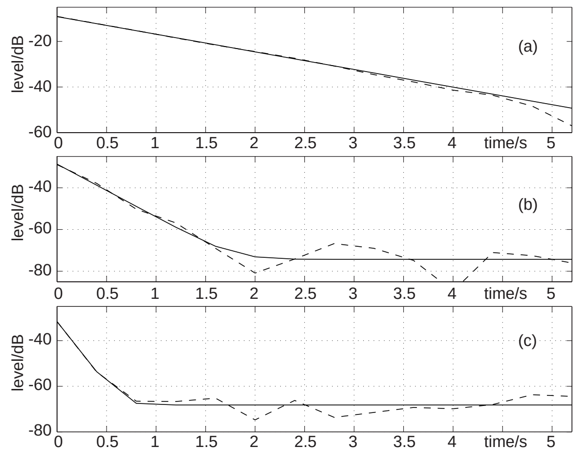

Fig. 8: Fitting of decay-plus-noise model to low-frequency modal data of a room (see Fig. 7): (a) at 40 Hz mode, (b) at 104 Hz, and (c) at 117 Hz (off-mode fast decay). Dotted line: measured, dashed line: modeled.

5.3 Analysis of decay rate of plucked string tones

Model-based synthesis of string tones can produce realistic guitar-like tones if the parameter values of the synthesis model are calibrated based on recordings [2]. The main properties of tones that need to be analyzed are their fundamental frequency and the decay time of all harmonic partials that are audible. While the estimation of fundamental frequency is quite easy, the measurement of decay times of harmonics (= modes of the string) is complicated by the fact that they all have a different rate of decay and also the initial level can vary within a range of 20-30 dB. There may also not be information about the noise floor level for all harmonics.

One method used for measuring the decay times is based on the short-time Fourier analysis. A recorded single guitar tone is sliced into frames with a window function in the time domain. Each window function is then Fourier transformed with the FFT using zero-padding to increase the spectral resolution, and harmonic peaks are hunted from the magnitude spectrum using a peak-picking algorithm. The peak values from the consecutive frames are organized as tracks, which correspond to the temporal envelopes of the harmonics. Then it becomes possible to estimate the decay rate of each harmonic mode. In the following, we show how this works with the proposed algorithm. Finally, the decay rate of each harmonic is converted into a corresponding target response, which is used for designing the magnitude response of a digital filter that controls the decay of harmonics in the synthesis model.

Figure 9 plots three examples of modal decay analysis of guitar string harmonics (string 5, open string plucking). Harmonic envelope trajectories were analyzed as described above. The decay-plus-noise model was fitted in a time window that started from the maximum value position of the envelope curve. In case (a), the second harmonic shows a highly regular decay after initial transient of plucking whereby decay fitting is almost perfect. Case (b), harmonic number 24, depicts a strongly beating decay where probably the horizontal and vertical polarizations have a frequency difference that after summation results in beating. Case (c), harmonic 54, shows a harmonic trajectory where the noise floor is reached within the analysis window. In all shown cases, the nonlinear optimization works as perfectly as a simple decay model can do.

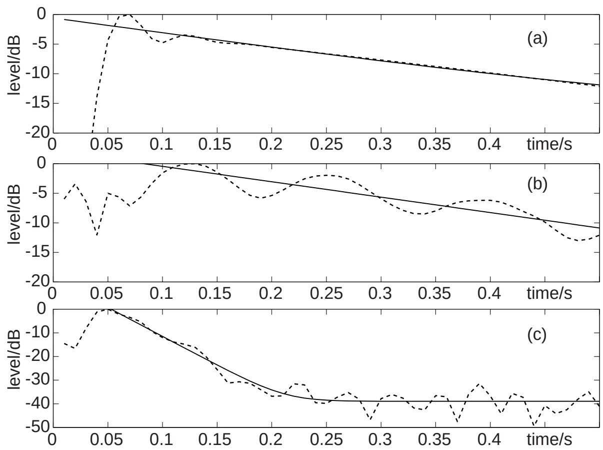

Fig. 9: Examples of modal decay matching for harmonic components of a guitar string: (a) regular decay after initial transient, (b) strongly beating decay (double mode), and (c) fast decay that reaches the noise floor. Dotted line: measured envelope, solid line: optimized fit.

As can be concluded from case (b) of Fig. 9, a string can exhibit more complicated behavior than a simple exponential decay. Even more complex is the case of piano tones because there are 2–3 strings slightly off-tuned, and the envelope fluctuation can be more irregular. Two-stage decay is also common where the initial decay is faster than later decay.

In all such cases a more complex decay model is needed to achieve a good match with the measured data. It remains a future problem to investigate fitting such models using linear optimization techniques.

6 SUMMARY AND CONCLUSIONS

In this paper, an overview of modal decay analysis methods for noisy impulse response measurements of reverberant acoustic systems is presented, and further improvements are introduced. The problem of decay time determination is important for example in room acoustics for characterizing the reverberation time. Another application where a similar problem is encountered is when estimating string model parameters for model-based synthesis of plucked string instruments.

It is shown in this article that the further developments of the decay-plus-noise model yield highly accurate decay parameter estimates, outperforming traditional methods especially under extreme SN-ratio conditions. Challenges for further studies remain to make the method (with increased number of parameters) able to analyze complex decay characteristics, such as double decay behavior and strongly fluctuating responses due to two or more modes very close in frequency.

The Matlab code for nonlinear optimization of decay parametres, including data examples, can be found at:
`http://www.acoustics.hut.fi/software/decay`

ACKNOWLEDGMENTS

This study is part of the VÄRE technology program, project TAKU (Control of Closed Space Acoustics), funded by Tekes (National Technology Agency). The work of Vesa Välimäki has been financed by the Academy of Finland.

Notes on reverberation time definitions

Reverberation time T_{60} or just T is defined as the time in seconds it takes for the energy in the steady-state sound field in a room to decay 60 dB after the source of sound excitation is suddenly turned off [19]. According to definition by Beranek [20], the reverberation time might be defined in any of the following ways:

1. the interval between turning off the sound source and the time when the instantaneous value of the pressure level first falls to 60 dB below its steady state value
2. the time interval until the average value of the fluctuating pressure first falls to this value
3. the time until the average value of the fluctuating log p decay curve first falls to 60 dB below its steady state value, or
4. the time required for the evaluated decay slope to drop off 60 dB.

In practice, the reverberation time is often determined from the slope of a decay curve using only the first 25 or 35 dB of decay and extrapolating the result to 60 dB. For the recommended practice of reverberation time determination, see standard [1].

REFERENCES

[1] ISO 3382-1997, *Acoustics – Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters.* Geneva, Switzerland: International Standards Organization, 1997. 21 p.

[2] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical Modeling of Plucked String Instruments with Application to Real-Time Sound Synthesis," *J. Audio Eng. Soc.*, vol. 44, no. 5, pp. 331–353, 1996 May.

[3] M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, pp. 409–412, 1965.

[4] M. R. Schroeder, "Integrated-impulse Method Measuring Sound Decay Without Using Impulses," *J. Acoust. Soc. Am.*, vol. 66, no. 2, pp. 497–500, 1979 Aug.

[5] D. Morgan, "A parametric error analysis of the backward integration method for reverberation time estimation," *J. Acoust. Soc. Am.*, vol. 101, no. 5, pp. 2686–2693, 1997 May.

[6] L. Faiget, C. Legros, and R. Ruiz, "Optimization of the Impulse Response Length: Application to Noisy and Highly Reverberant Rooms," *J. Audio Eng. Soc.*, vol. 46, no. 9, pp. 741–749, 1998 Sept.

[7] Y. Hirata, "A Method of Eliminating Noise in Power Responses," *J. Sound Vib.*, vol. 82, no. 4, pp. 593–595, 1982.

[8] A. Lundeby, T. E. Vigran, H. Bietz, and M. Vorländer, "Uncertainties of Measurements in Room Acoustics," *Acustica*, vol. 81, pp. 344–355, 1995.

[9] N. Xiang, "Evaluation of Reverberation Times Using a Nonlinear Regression Approach," *J. Acoust. Soc. Am.*, vol. 98, no. 4, pp. 2112–2121, 1995 Oct.

[10] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Englewood Cliffs, N.J.: Prentice-Hall, 1975. (Chapter 10, pp. 480–531).

[11] MathWorks, Inc., *MATLAB User's Guide*, 2001.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing: Volume I: Estimation Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1993.

[13] J. D. Markel and J. A. H. Gray, *Linear Prediction of Speech.* Berlin, Germany: Springer-Verlag, 1976.

[14] MathWorks, Inc., *MATLAB Signal Processing Toolbox, User's Guide*, 2001.

[15] W. T. Chu, "Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method," *J. Acoust. Soc. Am.*, vol. 63, no. 5, pp. 1444–1450, 1978 May.

[16] L. Faiget, R. Ruiz, and C. Legros, "Estimation of Impulse Response Length to Compute Room Acoustical Criteria," *Acustica*, vol. 82, no. Suppl. 1., p. S148, 1996 Sept.

[17] F. Satoh, Y. Hidaka, and H. Tachibana, "Reverberation Time Directly Obtained from Squared Impulse Response Envelope," in *Proc. Int. Congr. Acoust.*, vol. 4, (Seattle, WA), pp. 2755–2756, 1998 June.

[18] E. Zwicker and H. Fastl, *Psychoacoustics — Facts and Models.* Berlin: Springer-Verlag, 1990.

[19] W. C. Sabine, *Architectural Acoustics.* 1900. Reprint by Dover, 1964, New York.

[20] L. Beranek, *Acoustic Measurements.* 1949. Reprint by Acoust. Soc. Am., 1988.