Estimation of Modal Decay Parameters from Noisy Response Measurements*

MATTI KARJALAINEN,[1] *AES Fellow*, **POJU ANTSALO,[1] AKI MÄKIVIRTA,[2]** *AES Member*, **TIMO PELTONEN,[3]** *AES Member*, **AND VESA VÄLIMÄKI,[1,4]** *AES Member*

[1] *Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland*
[2] *Genelec Oy, Iisalmi, Finland*
[3] *Akukon Oy, Helsinki, Finland*
[4] *Pori School of Technology and Economics, Tampere University of Technology, Pori, Finland*

The estimation of modal decay parameters from noisy measurements of reverberant and resonating systems is a common problem in audio and acoustics, such as in room and concert hall measurements or musical instrument modeling. 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 the nonlinear optimization of a model for exponential decay plus stationary noise floor, is presented. A 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. Estimating the reverberation time or the modal decay rate are important measurement problems in room and concert hall acoustics [1], where signal-to-noise ratios (SNRs) 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 or reverberation time T_D
- Noise floor level L_N.

* Presented in the 110th Convention of the Audio Engineering Society, Amsterdam, The Netherlands, 2001 May 12–15; revised 2001 November 8 and 2002 September 6.

In a more complex case there can be two or more modal frequencies, whereby the decay is no longer simple, but shows additional fluctuation (beating) or a two-stage (or multiple-stage) decay behavior. In a diffuse field (room acoustics) the decay of a noiselike response is approximately exponential in rooms with compact geometry. The noise floor may also be nonstationary. In this paper we primarily discuss a simple mode (that is, a complex conjugate pole pair in the transfer function) or a dense set of modes with exponential reverberant decay, together with a stationary noise floor.

Methods presented in the literature and commonsense or ad hoc methods will first be reviewed. Techniques based on energy–time curve analysis of the signal envelope are known as methods where the noise floor can be found and estimated explicitly. Backward integration of energy, so-called Schroeder integration [3], [4], is often applied first to obtain a smoothed envelope for decay rate estimation.

The effect of the 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 the case of noisy data with care.

This technique, as most other methods, analyzes the initial level L_I, the decay time T_D, and the noise floor L_N 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 parameterized 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 fine-tune the method further for specific problems. The technique is compared with the Lundeby et al. method by it applying to simulated cases of exponential decay plus a 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 SNRs.

Finally, the applicability of the improved method is demonstrated by three examples of real measurement data: 1) the reverberation time of a concert hall, 2) the low-frequency mode analysis of a room, and 3) the 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, are 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

$$h(t) = A e^{-\tau (t - t_0)} \sin\left[\omega (t - t_0) + \phi\right] u(t - t_0) \qquad (1)$$

where $u(t - t_0)$ is a step function with value 1 for $t \geq t_0$ and 0 elsewhere, A is the initial response level, t_0 the response latency, for example, due to the propagation delay of sound, τ the decay rate parameter, $\omega = 2\pi f$ the angular frequency, and ϕ the initial phase of the 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[5]

$$h(t) = \sum_{i=1}^{N} A_i e^{-\tau_i (t - t_0)} \sin\left[\omega_i (t - t_0) + \phi_i\right] A_n n(t) \qquad (2)$$

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

The task of this study is defined as finding reliable estimates for the parameter set $\{A_i, \tau_i, t_0, \omega_i, \phi_i, A_n\}$, given a noisy measured impulse response of the form of Eq. (2).

[5] In a more general case the initial delays of signal components may differ and there can be simple nonresonant exponential terms, but these cases are of less importance here.

The main interest here is focused on systems of 1) separable single modes of type (1), including additive noise floor, or 2) the 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.

Often the decay time T_D is of main interest, for example, in room acoustics where the reverberation time [10], [11] of 60-dB decay[6] T_{60} is related to τ by

$$T_{60} = -\frac{1}{\tau} \ln\left(10^{-3}\right) \approx \frac{6.908}{\tau} . \qquad (3)$$

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

$$h(n) = A e^{-\tau_d n} \sin(\Omega n + \phi) \qquad (4)$$

with n being the sample index, $\tau_d = T_s \tau$, and $\Omega = 2\pi T_s f$.

2 DECAY PARAMETER ESTIMATION

In this section an overview of the known techniques for decay parameter estimation will be presented. Initial delay

[6] 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 a recommended practice of reverberation time determination, see [1].

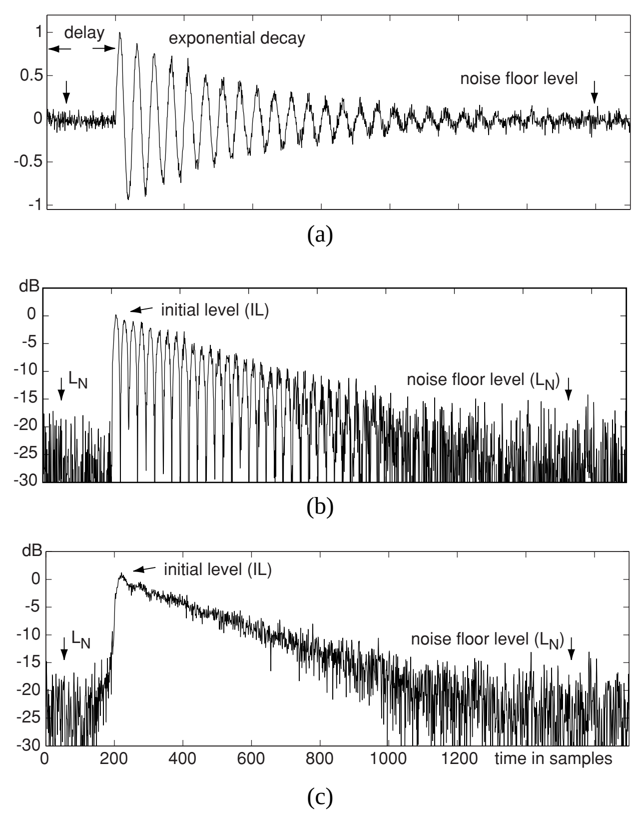

Fig. 1. (a) Single-mode impulse response (sinusoidal decay) with initial delay and additive measurement noise. (b) Absolute value of response on dB scale to illustrate decay envelope. (c) Hilbert envelope; otherwise same as (b).

and level estimation are first discussed briefly. The main problem, decay rate estimation, is the second topic. Methods to smooth the decay envelope from a measured impulse response are presented. 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 the initial delay and the initial level 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 SNR is low, can the detection of onset time be difficult.

A simple technique to eliminate initial delay is to compute the minimum-phase component $h_{\mathrm{mphase}}(t)$ of the measured response [12]. 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 all-pass properties manifested as a 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.

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 the source-to-receiver distance or in estimating the impulse response of the sound source itself by windowing the response prior to the first room reflection.

2.2 Decay Rate Estimation

Decay rate or time estimation is in practice based on fitting a straight line to the decay envelope, such as the energy–time curve, mapped on a logarithmic (dB) scale. Before the 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, in practice the automatic determination of the decay rate or time is highly desirable.

2.2.1 Straight-Line Fit 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,

$$y(t) = 10 \log_{10}\left[x^2(t)\right] \tag{5}$$

where $x(t)$ is the measured impulse response or a band-pass filtered part of it, such as an octave or one-third-octave band. It is common to apply techniques such as Schroeder integration and Hilbert envelope computation (to be described later) 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 and the initial level a,

$$\min_{k,a} \int_{t_1}^{t_2} \left[y(t) - (kt + a)\right]^2 \mathrm{d}t \tag{6}$$

using, for example, the Matlab function polyfit [13].

Practical problems with line fitting are related to the selection of the 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 the noise floor at the end of the measurement interval. The second problem is related to such cases as 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 close in frequency [see Fig. 9(b)].

2.2.2 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 its 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,

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

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

Nonlinear optimization is mathematically more complex than linear fitting, and care should be taken to guarantee convergence. Even when converging, 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 multiparameter 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.

2.2.3 AR and ARMA Modeling

For a single mode of Eq. (1) the response can be modeled as an impulse response of a resonating second-order all-pole or pole–zero filter. More generally, a combination of N modes can be modeled as a $2N$-order filter. AR (autoregressive) and ARMA (autoregressive moving average) modeling [14] are ways to derive parameters for such models. In many technical applications the AR method is called linear prediction [15]. For example, the function lpc in Matlab [16] processes a signal frame through autocorrelation coefficient computation and solving normal equations by Levinson recursion, resulting in the Nth-order z-domain transfer function $1/(1 + \Sigma_{i=1}^{N} \alpha_i z^{-i})$. 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 pole angle $\phi = \arg(z_i) = 2\pi f/f_s$ and pole radius $r = |z_i| = e^{-\tau/f_s}$, where f is the modal frequency, f_s the sampling rate, and τ the decay parameter of the mode in Eq. (1). ARMA modeling requires an iterative solution for a pole–zero filter.

Decay parameter analysis by AR and ARMA modeling is an important technique and attractive, for example, in cases where modes overlapping or very close to each other have to be modeled, which is often difficult by other means. For reverberation with high modal density the order of AR modeling may become too high for accurate modeling. Such accuracy is also not necessary for analyzing the overall decay rate (reverberation time) only. AR and ARMA modeling of modal behavior in acoustic systems are discussed in detail, for example, in [17].

2.2.4 Group Delay Analysis

A complementary method to AR modeling is to use the group delay, that is, the phase derivative $T_g(\omega) = -d\varphi(\omega)/d\omega$, as an estimate of the decay time for separable modes of an impulse response. While AR modeling is sensitive to the 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, that is, $T_g = 1/\tau$. Group delay computation is somewhat critical due to the phase unwrapping needed, and the method can be 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 prior to the fitting operation. The following techniques are often used to improve the regularity of the decay ramp.

2.3.1 Hilbert Envelope Computation

In this method the signal $x(t)$ is first converted to an analytic signal $x_a(t)$ so that $x(t)$ is the real part of $x_a(t)$ and the Hilbert transform (90° phase shift) [12] is the imaginary part of $x_a(t)$. For a single sinusoid this results in an entirely smooth energy–time envelope. An example of a Hilbert envelope for a noisy modal response is shown in Fig. 1(c).

2.3.2 Schroeder Integration

A monotonic and smoothed decay curve can be produced by backward integration of the impulse response $h(t)$ over the measurement interval $[0, T]$ and converting it to a logarithmic scale,

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

This process is commonly known as 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 upward the late part of the decay. This is shown in Fig. 2 for the case of noisy single-mode decay [curve (a)] for full response integration [curve (d)].

The tail problem of Schroeder integration has been addressed by many authors (for example, in [18], [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 SNR inherent in practically all acoustical measurements, and especially measurements performed under field conditions, calls 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 a large enough dynamic range or the whole system decay down to the background noise level.[7]

Thus the measured impulse response typically contains not only the decay curve under analysis, but also a steady level of background noise, which dominates at the end of

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

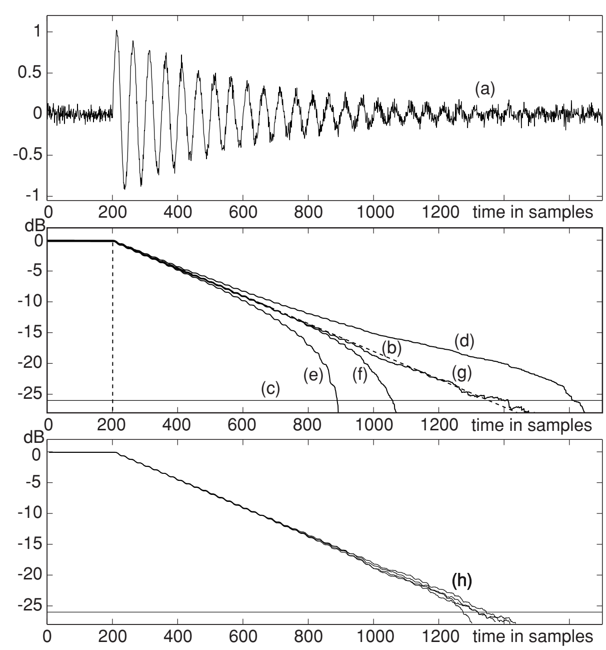

Fig. 2. Results of Schroeder integration applied to noisy decay of a mode. Curve (a) measured noisy response including initial delay; curve (b) true decay of noiseless mode (dashed straight line); curve (c) noise floor (−26 dB); curve (d) Schroeder integration of total measured interval; curve (e) integration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

the response. Fitting the decay line over this part of the envelope or Schroeder integrating this steady energy level along with the exponential decay curve causes an error both in the resulting decay rate (see Fig. 2) and in the time-windowed energies (energy parameters).

To avoid bias by noise, an 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.

Determination of the noise floor level is difficult without using iterative techniques. The method by Lundeby et al., which will be outlined later, is a good example of iterative techniques integrated with decay rate estimation.

A simple way to obtain a reasonable estimate of a background noise floor is to average a selected part of the measured response tail or to fit a regression line to it [19]. The level is certainly overestimated if the noise floor is not reached, but this is not necessarily problematic, as opposed to underestimating it. Another technique is to look at the background level before the onset of the 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 the background noise, and to truncate the noise from the end of the response. Fig. 2 illustrates the effect of limiting the Schroeder integration interval. If the interval is too short, as in curve (e), the curve is biased downward. Curve (f) shows a case where the bias due to noise is minimized by considering the decay only down to 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. The methods are discussed next.

2.5.1 Limited Integration or Decay Matching Interval

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

Faiget et al. [19] propose a simple but systematic method for postprocessing 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 the background noise regression lines. The decay parameter fitting interval ends 5 dB above the noise floor.

2.5.2 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 as follows.

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 (maximum five iterations).

The 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.

2.5.3 Subtraction of Noise Floor Level

Chu [18] 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 accurate and the noise is stationary, the resulting backward integrated curve is close to the ideal decay curve.

2.5.4 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,

$$\begin{aligned}\int_t^\infty h^2(t)\,\mathrm{d}t &\Leftarrow \int_t^\infty \left[h_1(t)+n_1(t)\right]\left[h_2(t)+n_2(t)\right]\mathrm{d}t\\ &= \int_t^\infty \left[h_1(t)h_2(t)+h_1(t)n_2(t)+h_2(t)n_1(t)+n_1(t)n_2(t)\right]\mathrm{d}t\\ &= \int_t^\infty \left\{h_1(t)h_2(t)+n_1(t)n_2(t)\left[1+\frac{h_1(t)}{n_1(t)}+\frac{h_2(t)}{n_2(t)}\right]\right\}\mathrm{d}t\\ &\approx \int_t^\infty h^2(t)\,\mathrm{d}t + K(t)\,. \end{aligned} \tag{9}$$

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 the 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. Hirata's method relies on the impulse response at large time values to be stationary. This condition is often not met in practical concert hall measurements.

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

2.5.5 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. [20]).

3 NONLINEAR OPTIMIZATION OF A DECAY-PLUS-NOISE MODEL

The nonlinear regression (optimization) method proposed by Xiang [9] was briefly described earlier. In the present study we worked along similar ideas, using nonlinear optimization for improved robustness and accuracy. In the following 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 a simple exponential decay of the response envelope, corrupted by additive stationary background noise. We will study two cases that fit into the same modeling category. In the first case there is a single mode (a complex conjugate pole pair in transfer function) that in the time domain corresponds to an exponential decay function,

$$h_\mathrm{m}(t) = A_\mathrm{m}\,\mathrm{e}^{-\tau_\mathrm{m} t}\,\sin\left(\omega_\mathrm{m} t + \phi_\mathrm{m}\right)\,. \tag{10}$$

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 signal,

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

where A_d is the initial rms level of the response, τ_d is a decay rate parameter, and $n(t)$ is stationary Gaussian noise with an 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,

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

where A_n is the rms level of the Gaussian measurement noise in 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

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

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 an optimized least-squares (LS) error estimate for the parameters $\{A, \tau, A_\mathrm{n}\}$ can be achieved by minimizing the following expression over a time span $[t_0, t_1]$ of interest:

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

Since the model of Eq. (13) is nonlinear in the 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 as

$$\min_{A,\tau,A_\mathrm{n}} \int_{t_0}^{t_1} \left[f\left(a(t), t\right) - f\left(y(t), t\right)\right]^2 \mathrm{d}t \tag{15}$$

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

The choice of $f(y, t) = 20\log_{10}[y(t)]$ results in fitting on the dB 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 the exponent $s < 1$, which is a compromise between amplitude and logarithmic scaling. A value of $s = 0.5$ has been found to be a useful default value.[8]

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

$$\min_{A,\tau,A_n} \int_{t_0}^{t_1} \left[w(t)\, a^s(t) - w(t)\, y^s(t) \right]^2 \mathrm{d}t \,. \tag{16}$$

There is no clear physical motivation for the magnitude compression exponent s. A specific temporal weighting function $w(t)$ can be applied case by case, based on extra knowledge of 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 $\{A, \tau, A_n\}$ are needed with greatest 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 an estimate of the noise level is nevertheless desired. A necessary assumption for the method to work in such cases is that the decay model is valid, implying an exponential decay and a stationary noise floor.

Experiments show that the model is useful for both single-mode decay and reverberant acoustic field decay models. Fig. 3 depicts three illustrative examples of decay model fitting to a single mode plus noise 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. 3(a) the SNR is only 6 dB. Errors in the parameters in this case are a 0.5-dB underestimate of A, a 3.5% underestimate in the decay time related to parameter τ, and a 1.8-dB overestimate of the noise floor A_n. In Fig. 3(b) a similar case is shown with a moderate 30-dB SNR. Estimation errors of the parameters are $+0.2$ dB for A, -2.8% for decay time, and $+1.2$ dB for A_n. In the third case [Fig. 3(c)] the SNR is -60 dB so that the noise floor is barely 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, even in this extreme case.

The nonlinear optimization used in this study is based on using the Matlab function curvefit,[9] and the functions that implement the weighting by parameter s and the weighting function $w(t)$ can be found at http://www.acoustics.hut.fi/software/decay.

The optimization routines are found converging robustly in most cases, including such extreme cases as Fig. 3(a) and (c), and the initial values of the parameters for iteration are not critical. However, it is possible that in rare cases the optimization diverges and no (not even a local) optimum is found.[10] 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 the very best results. It is advantageous to select the analysis frame so that the noise floor is reached neither too early nor too 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 SNRs such as in Fig. 3(a), and let it approach a value of 0.4–0.5 when the noise floor is low, as in Fig. 3(c) (see also Fig. 4).

4 COMPARISON OF DECAY PARAMETER ESTIMATION METHODS

The accuracy and robustness of the methods for decay parameter estimation can be evaluated by using synthetic

[10] The function curvefit also prints warnings of computational precision problems even when optimization results are excellent.

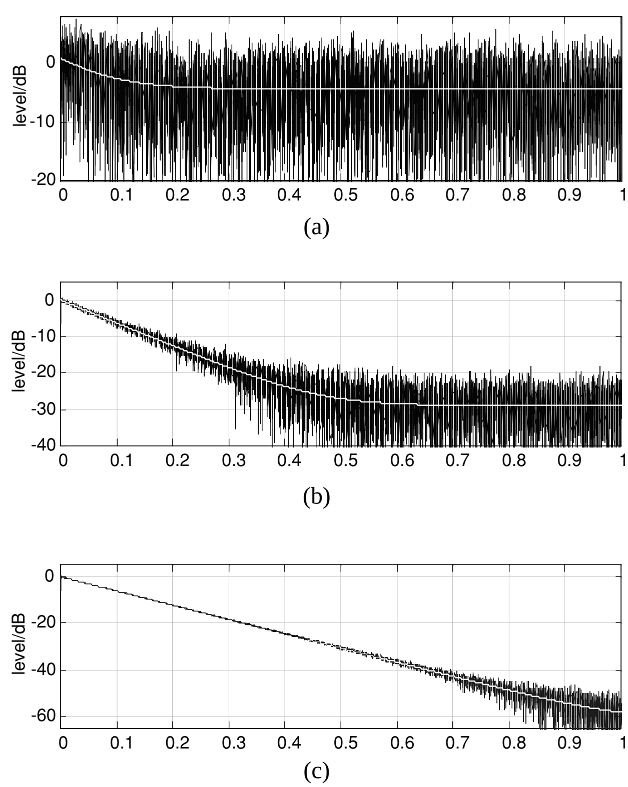

Fig. 3. Nonlinear optimization of decay-plus-noise model for three synthetic noisy responses with initial level of 0 dB and various noise levels. (a) -6 dB. (b) -30 dB. (c) -60 dB. Black curves—Hilbert envelopes of simulated responses; white curves—estimated decay behavior.

[8] Interestingly enough, this resembles the loudness scaling in auditory perception known from psychoacoustics [21].

[9] In new versions of Matlab, the function curvefit is recommended to be replaced by the function lsqcurvefit.

decay signals or envelope curves, computed for sets of the parameters $\{A, \tau, A_n\}$. By repeating the same for different methods, their relative performances can be compared. In this section we present results from a comparison of the proposed nonlinear optimization and the method of Lundeby et al. [8].

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. The initial sinusoidal level to background noise ratio was varied from 0 to 80 dB in 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 dB for the initial and the noise floor levels and as a percentage of decay time.

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

Fig. 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 SNRs, such as 0 dB, it cannot deal with these cases without extra tricks, and even then it may have severe problems. We used safety settings whereby we did not try to obtain decay time values for SNRs below 20 dB, and low SNR parts of the decay parameter estimate curves are omitted.

For moderate SNRs the results of the method are fairly good and robust. The decay time shows a positive bias of a few percent, except for an SNR below 30 dB. 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 for 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

Estimating the reverberation time of a room or a hall is relatively easy if the decay curve behaves regularly and the noise floor is low enough. In practice the case is often quite different. Here we demonstrate the behavior of the nonlinear optimization method in an example where the measured impulse response includes an initial delay, an irregular initial part, and a relatively high measurement noise floor.

Fig. 6 depicts three different cases of fitting the decay-plus-noise model to this case of a control room with a short reverberation time. In Fig. 6(a) the fitting is applied

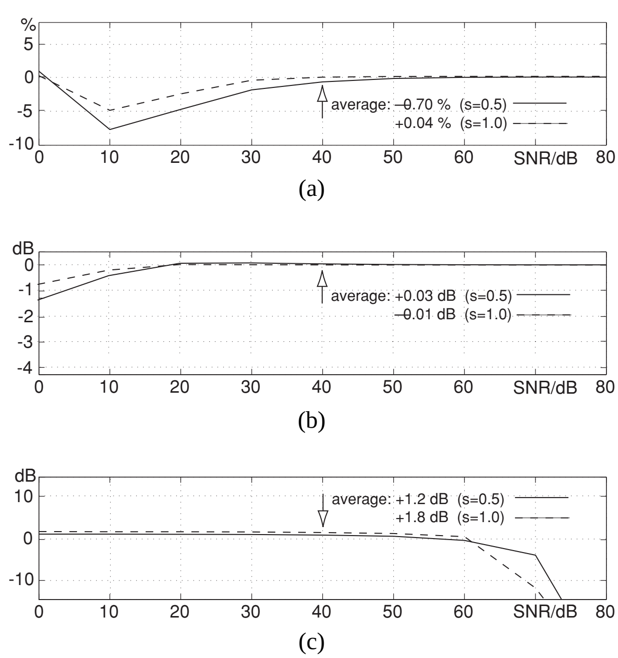

Fig. 4. Sine-plus-noise decay parameter estimation errors (average of 20 trials) for proposed nonlinear optimization method as a function of SNR. (a) Decay time estimation error in %. (b) Initial level estimation error in dB. (c) Noise floor estimation error in dB. —— $s = 0.5$; – – – $s = 1.0$.

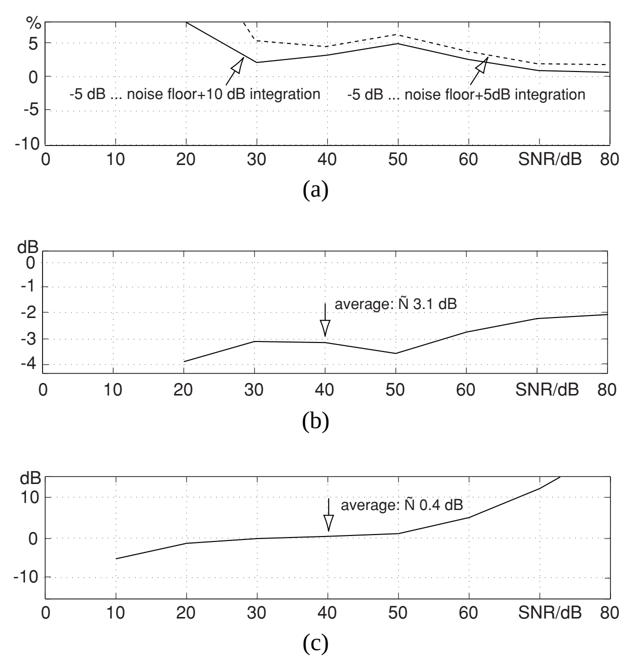

Fig. 5. Decay parameter estimation errors for Lundeby et al. method as a function of SNR. (a) Decay time estimation error in % with truncated Schroeder integration but without noise compensation. (b) Initial level estimation error in dB. (c) Noise floor estimation error in dB.

to the entire decay curve, including the initial delay, and the resulting model is clearly biased toward too long a reverberation time. In Fig. 6(b) the initial delay is excluded from model fitting, and the result is better. However, after the direct sound there is a period of only little energy during the first reflections prior to the range of dense reflections and diffuse response. If the reverberation time estimate is to describe the decay of this diffuse part, the case of Fig. 6(c), with a fitting starting from about 30 ms, yields the best match to reverberation decay,[11] and the approaching noise floor is also estimated well.

5.2 Modeling of Low-Frequency Room Modes

The next case deals with the modeling of the low-frequency modes of a room. Below a critical frequency (the 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 the most prominent modes and to analyze their modal parameters f_m and τ_m, the frequency and decay parameters, respectively. The case studied was a hard-walled, partially damped room with moderate reverberation time ($\approx$ 1 second) at mid and high frequencies, but much longer decay times at the lowest modal frequencies. The following procedure was applied:

- A short-time Fourier analysis of the measured impulse response was computed to yield the 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.

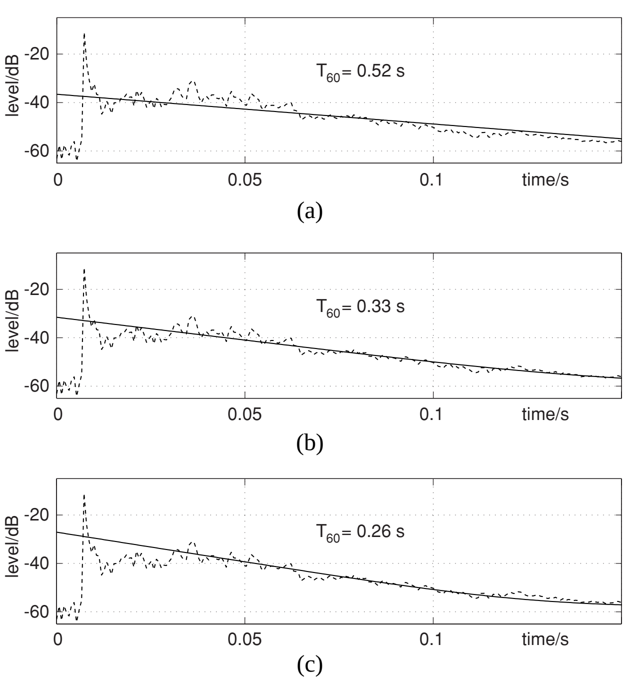

Fig. 6. Decay-plus-noise model fitting by nonlinear optimization to a room impulse response. (a) Fitting range includes initial delay, transient phase, and decay. (b) Fitting includes transient phase and decay. (c) Fitting includes only decay phase. Estimated T_{60} values are given.

- 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.

In this context we are interested in how well the decay parameter estimation worked with noisy measurements. Application of the nonlinear optimization resulted in decay curve fits, some of which are illustrated in Fig. 8, by comparing the original decay and the decay-plus-noise model behavior. For all frequencies in the vicinity of a mode the model fits robustly and accurately.[12]

5.3 Analysis of Decay Rate of Plucked String Tones

A 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 estimating the fundamental frequency is quite easy, measurement of the 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 be no 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 fast Fourier transform 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 decay parameter estimation 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.

Fig. 9 plots three examples of modal decay analysis of guitar string harmonics (string 5, open string). Harmonic envelope trajectories were analyzed as described. The decay-plus-noise model was fitted in a time window that started from the maximum value position of the envelope curve. In Fig. 9(a) the second harmonic shows a highly regular decay after an initial transient of plucking, whereby decay fitting is almost perfect. Fig. 9(b), harmonic 24,

[11] In this example, decay parameter analysis is applied to the entire frequency range of the impulse response. In practice it is computed as a function of frequency, that is, applied to octave or one-third-octave band decay curves.

[12] To obtain the best frequency resolution it may be desirable to replace short-time spectral analysis with energy-decay curves obtained by backward integration and the model of Eq. (16) with a corresponding formulation derived from Xiang's formula, Eq. (7).

depicts a strongly beating decay, where probably the horizontal and vertical polarizations have a frequency difference that after summation results in beating. Fig. 9(c), harmonic 54, shows a trajectory where the noise floor is reached within the analysis window. In all cases shown the nonlinear optimization works as perfectly as a simple decay model can do.

As can be concluded from Fig. 9(b), a string can exhibit more complicated behavior than simple exponential decay. Even more complex is the case of piano tones because there are two to three strings slightly off tune, and the envelope fluctuation can be more irregular. Two-stage decay is also common where the initial decay is faster than later decay [22].

In all such cases a more complex decay model is needed to achieve a good match with measured data. Such techniques are studied in [17].

6 SUMMARY AND CONCLUSIONS

An overview of modal decay analysis methods for noisy impulse response measurements of reverberant acoustic systems has been presented, and further improvements were 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 the estimation of string model parameters for model-based synthesis of

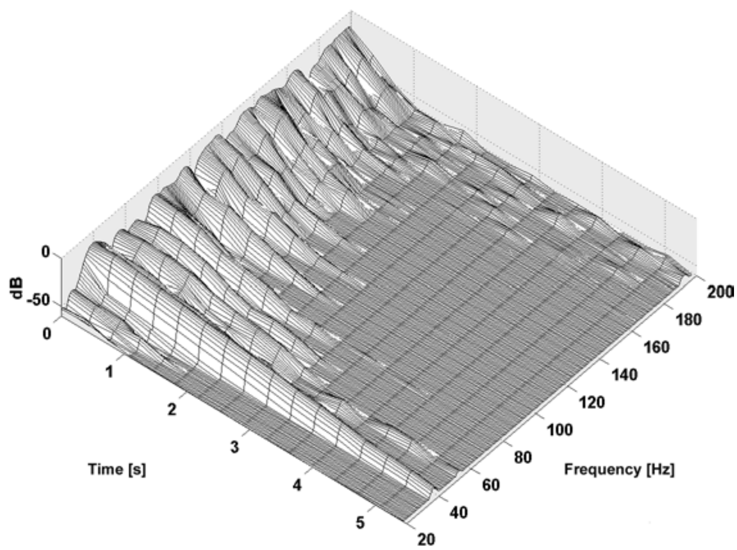

Fig. 7. Time–frequency plot of room response at low frequencies. Lowest modes (especially 40 Hz) show long decay times.

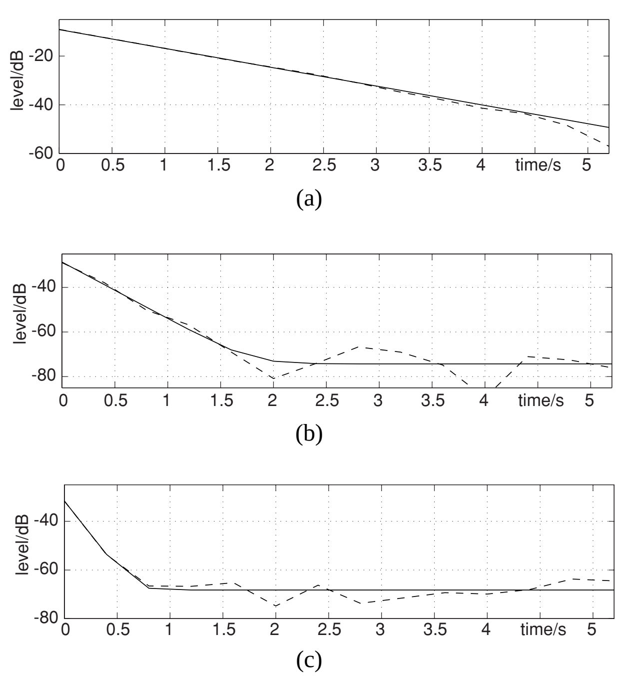

Fig. 8. Fitting of decay-plus-noise model to low-frequency modal data of room (see Fig. 7). (a) At 40-Hz. (b) At 104 Hz. (c) At 117 Hz (off-mode fast decay). – – – measured; —— modeled.

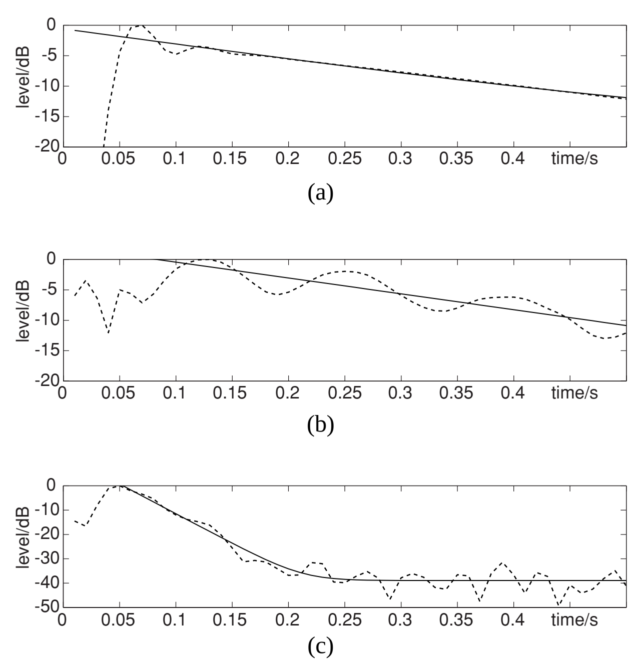

Fig. 9. Examples of modal decay matching for harmonic components of guitar string. (a) Regular decay after initial transient. (b) Strongly beating decay (double mode). (c) Fast decay that reaches noise floor. – – – measured envelope; —— optimized model fit.

plucked string instruments.

It is shown that the developed decay-plus-noise model yields highly accurate decay parameter estimates, outperforming traditional methods, especially under extreme SNR conditions.

There exist other methods, such as AR modeling, that show potential in specific applications. Challenges for further research are to make modal decay methods (with an 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.

A Matlab code for the nonlinear optimization of decay parameters, including data examples, can be found at http://ww.acoustics.hut.fi/software/decay.

7 ACKNOWLEDGMENT

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.

8 REFERENCES

[1] ISO 3382-1997, "Acoustics—Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters," International Standards Organization, Geneva, Switzerland (1997), 21 pp.

[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, 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, 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, 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, pp. 741–750 (1998 Sept.).

[7] Y. Hirata, "A Method of Eliminating Noise in Power Responses," *J. Sound Vib.*, vol. 82, 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, pp. 2112–2121 (1995 Oct.).

[10] W. C. Sabine, *Architectural Acoustics* (1900; reprinted by Dover, New York, 1964).

[11] L. Beranek, *Acoustic Measurements* (1949; reprinted by Acoustical Society of America, 1988).

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

[13] MathWorks, Inc., *MATLAB User's Guide* (2001).

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing*, vol. I, *Estimation Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1993).

[15] J. D. Markel and J. A. H. Gray, *Linear Prediction of Speech* (Springer, Berlin, 1976).

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

[17] M. Karjalainen, P. A. A. Esquef, P. Antsalo, A. Mäkivirta, and V. Välimäki, "Frequency-Zooming ARMA Modeling of Resonant and Reverberant Systems," to be published, vol. 50 (2002 Dec.).

[18] W. T. Chu, "Comparison of Reverberation Measurements Using Schroeder's Impulse Method and Decay-Curve Averaging Method," *J. Acoust. Soc. Am.*, vol. 63, pp. 1444–1450 (1978 May).

[19] L. Faiget, R. Ruiz, and C. Legros, "Estimation of Impulse Response Length to Compute Room Acoustical Criteria," *Acustica*, vol. 82, suppl. 1, p. S148 (1996 Sept.).

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

[21] E. Zwicker and H. Fastl, *Psychoacoustics—Facts and Models* (Springer, Berlin, 1990).

[22] G. Weinreich, "Coupled Piano Strings," *J. Acoust. Soc. Am.*, vol. 62, pp. 1474–1484 (1977 Dec.).

THE AUTHORS

P. Antsalo

T. Peltonen

Poju Pietari Antsalo was born in Helsinki, Finland, in 1971. He studied electrical and communications engineering at the Helsinki University of Technology, and obtained a master's degree in acoustics and audio signal processing in 1999. He has carried out research on room acoustics and audio reproduction, particularly at low frequencies. E-mail: antsalo@cc.hut.fi.

●

Timo Peltonen was born in Espoo, Finland, in 1972. He received an M.Sc. degree in electrical engineering from the Helsinki University of Technology in 2000. His master's thesis addressed the measurement and room acoustic analysis of multichannel impulse responses.

He is the author of IRMA, the Matlab-based measurement and analysis software. Currently he is working as an acoustical consultant at Akukon Consulting Engineers Ltd., Helsinki, Finland.

Mr. Peltonen's interests include acoustical measurement techniques and signal processing methods for room acoustical analysis.

He is a member of the AES, a board member of the AES Finnish Section, and a member of the Acoustical Society of Finland. He is the author of several AES conference papers.

The biographies of Matti Karjalainen and Vesa Välimäki were published in the 2002 April issue of the *Journal*.

The biography of Aki Mäkivirta was published in the 2002 July/August issue of the *Journal*.