IMPULSE RESPONSE MEASUREMENT TECHNIQUES AND THEIR APPLICABILITY IN THE REAL WORLD

Martin Holters, Tobias Corbach, Udo Zölzer

Dept. of Signal Processing and Communications,
Helmut Schmidt University
Hamburg, Germany
hol@hsu-hh.de, corbach@hsu-hh.de, zoelzer@hsu-hh.de

ABSTRACT

Measurement of impulse responses is a common task in audio signal processing. In this paper three common measurement techniques are reviewed: Maximum length sequences, exponentially swept sines and time delay spectrometry. The aim is to give the reader a brief tutorial of the methods with a special focus on deficiencies of the algorithms, aiding in the choice of the best algorithm for a task at hand. Additionally, for time delay spectrometry, a novel improvement is presented, lifting its restriction to relatively short impulse responses.

1. INTRODUCTION

Measurement of impulse responses is a common task in audio signal processing. Usual applications include measurement of speakers and measurement of room impulse responses. Typically PC hardware is used due to the ever decreasing prices for ever faster processors and the availability of high quality audio devices.

When discussing impulse response measurement techniques, it therefore makes sense to assume a setup as depicted in figure 1. Inside the PC, a measurement signal $x(n)$ is generated. This is converted to a continuous-time signal $\bar{x}(t)$ and fed to the system under test with the (unknown) impulse response $\bar{h}(t)$. The resulting output $\bar{y}(t)$ is sampled to obtain $y(n)$ which is used together with the known $x(n)$ to determine the sampled impulse response $h(n)$. The following assumptions are made:

- The whole system between $x(n)$ and $y(n)$ shows time-invariant behavior. This holds with sufficient accuracy when the system to be measured is time-invariant, the jitter of the ADC and DAC clocks is low and the ADC and DAC run from the same clock so that the sampling rates match perfectly. The latter is usually the case for typical PC equipment.
- All involved components are sufficiently linear. Usually reduction of the measurement level and hence the maximum amplitude of the measurement signal is sufficient to restore almost linear behavior in case non-linear distortion is observed. However, a small amount of distortion may remain, recommending a measurement technique which is tolerant to slight non-linearities.
- The impulse response may be long (e.g. a reverberant room), nevertheless there is an N such that $h(n)$ is practically zero for $n > N$, so that is is sufficient to determine $h(n)$ for n up to that N.
- A very accurate measurement is desired. Especially for artificial reverberation with a measured room impulse response, its signal to noise ratio (SNR) has to be very high.

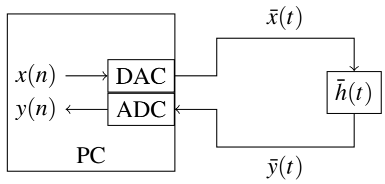


Figure 1: *Assumed setup for measurement of impulse responses.*

2. IDEAL IMPULSE RESPONSE MEASUREMENT

Theoretically, any measurement signal $x(n)$ may be used for estimating the unknown impulse response $h(n)$. Ignoring any noise terms, the measured signal $y(n)$ is the convolution

$$y(n) = \sum_{k=0}^{N} x(n-k)h(k), \tag{1}$$

where N is the order of the (significant part of the) impulse response. By rewriting this to matrix form as

$$\begin{pmatrix} y(0) \\ y(1) \\ \vdots \\ y(N+L-1) \end{pmatrix} = \begin{pmatrix} x(0) & 0 & \cdots & 0 \\ x(1) & x(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x(L-1) \end{pmatrix} \begin{pmatrix} h(0) \\ h(1) \\ \vdots \\ h(N) \end{pmatrix}, \tag{2}$$

where L denotes the length of the measurement signal $x(n)$, it is apparent that given $x(n)$ and $y(n)$, $h(n)$ may be computed using a linear least squares approach. In the presence of noise, this will give the minimum mean square error estimate, i.e. maximize the SNR.

Unfortunately, in practice, very long impulse responses may have to be measured and to attain high suppression of measurement noise, long measurement signals may be needed, resulting in extremely large problems which may turn out to become computationally intractable. Therefore, measurement signals are required that allow a simpler computation of the impulse response.

3. PSEUDO-NOISE MEASUREMENT SEQUENCES

One approach to simplify the problem is to instead look at the equivalent for stochastic signals, the well-known Wiener-Hopf

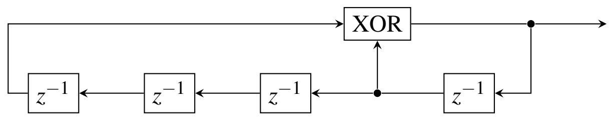


Figure 2: *Linear feedback shift register for generation of a maximum length sequence (of length* $2^4 - 1 = 15$ *in this case).*

equations

$$\begin{pmatrix} r_{xy}(0) \\ r_{xy}(1) \\ \vdots \\ r_{xy}(N) \end{pmatrix} = \begin{pmatrix} r_{xx}(0) & r_{xx}(1) & \cdots & r_{xx}(N) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(N) & r_{xx}(N-1) & \cdots & r_{xx}(0) \end{pmatrix} \begin{pmatrix} h(0) \\ h(1) \\ \vdots \\ h(N) \end{pmatrix}, \tag{3}$$

where $r_{xy}(n)$ and $r_{xx}(n)$ denote cross- and auto-correlation, respectively. By choosing the measurement signal such that

$$r_{xx}(n) = \begin{cases} \xi \gg 0 & \text{for } n = 0 \\ 0 & \text{for } 1 \leq n \leq N, \end{cases} \tag{4}$$

equation (3) simplifies to $h(n) = r_{xy}(n)/\xi$, so that computation of the impulse response reduces to cross-correlating $x(n)$ and $y(n)$.

While equation (4) holds for white noise sequences, it only approximately does for finite length excerpts of white noise. In practice, it is usually preferred to use pseudo-noise sequences with known properties, such as the commonly used maximum length sequences [1], which repeat with period L and have

$$r_{xx}(n) = \begin{cases} 1 & \text{for } n = 0 \\ -1/L & \text{for } 1 \leq n \leq L. \end{cases} \tag{5}$$

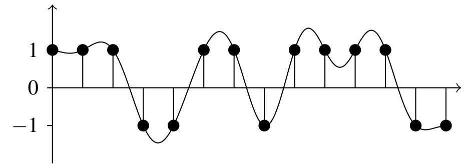
Thus, for sufficiently large L, which may be obtained without high computational demands, the auto-correlation can be made arbitrarily close to ideal. The remaining deviation from the ideal is only a DC offset, which is often irrelevant in practice anyway.

Another reason to use large L is that with a periodic measurement signal, the measured impulse response will also be periodic and hence, L has to be longer than the effective length N of the impulse response to avoid time aliasing effects.

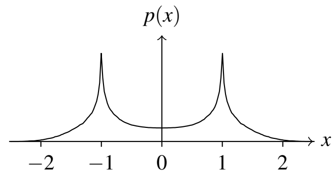
Maximum length sequences may be generated computationally cheap by employing linear feedback shift registers as depicted in figure 2, where a simple shift register and few exclusive-or gates are sufficient. The resulting bit sequence is remapped to a bipolar sequence with values -1 and 1. To compute the correlation, no multiplications are necessary, only changes of the sign and additions. Furthermore, the number of additions may be reduced by using the fast Hadamard transform [2].

In addition to the low computational costs, another alleged benefit of maximum length sequences is their ideal crest factor, as they only take on values -1 and $+1$, thus maximizing the total energy of the measurement signal when the peak amplitude is constrained. However, this is only true for the discrete time signal; after digital-to-analog conversion, the crest factor is severely deteriorated.

For the common case of a standard digital-to-analog converter with an output band-limited to half the sampling rate, the results depend on the low-pass, but in general will be similar to figure 3. For the example shown, the crest factor is increased to about 2.5.



(a) Discrete maximum length sequence and continuous signal obtained by band-limited reconstruction.



(b) Typical distribution of amplitude values after reconstruction.

Figure 3: *Example of a maximum length sequence and the continuous signal obtained by band-limited reconstruction.*

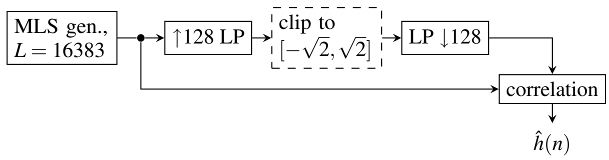
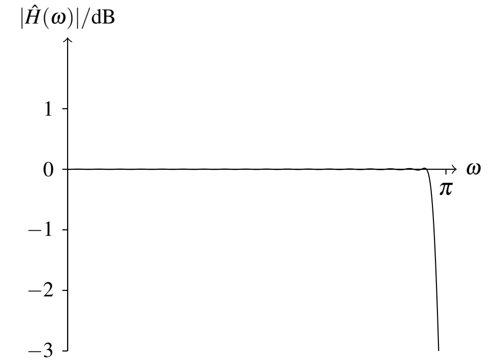


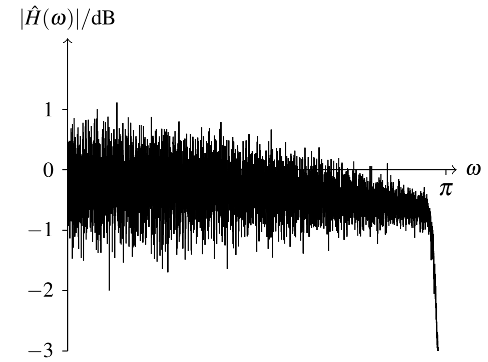
Figure 4: *Simulation setup for evaluation of the effect of insufficient headroom on MLS measurements.*

It should be noted that this increase is due to an increase of the maximum amplitude value, making careful calibration of the playback level necessary [3]. Calibrating with a bipolar sequence is preferable, but determining whether clipping occurs may not be easy as non-linear distortion is is not easily discernible. If instead a sinusoid is used for calibration, non-linearities become apparent by the occurrence of harmonic distortion, but sufficient headroom (8 dB for this example) has to be granted to allow for the higher peak amplitudes of the bipolar sequence. Otherwise, non-linear distortion will occur which has a severe impact on the measurement performance [4].

To undermine the adverse effect of insufficient headroom, a simulation experiment as depicted in figure 4 was conducted. To simulate the analog domain, the discrete signal is up-sampled by a factor of 128 and low-pass filtered with a Kaiser-windowed sinc. The obtained quasi-continuous signal is optionally clipped to the range $[-\sqrt{2}, \sqrt{2}]$, corresponding to 3 dB headroom. After another low-pass and down-sampling, the estimated impulse response $\hat{h}(n)$ is determined by correlating with the original sequence. Figure 5(a) shows the corresponding transfer function measured without clipping; as expected, the effect of the finite slope of the anti-aliasing low-pass becomes visible toward high frequencies. In comparison, the transfer function of figure 5(b) as measured with clipping looks severely distorted. The effect of clipping could easily be mistaken for measurement noise, but as it is completely deterministic, increasing measurement time will not improve the result.



(a) Without clipping.



(b) With clipping.

Figure 5: *Resulting transfer functions as measured with the setup of figure 4.*

4. EXPONENTIALLY SWEPT SINE

Another approach to impulse response measurement is to construct a measurement signal $x(n)$ for which a signal $x^{-1}(n)$ can easily be determined that is inverse to $x(n)$ in the sense that the convolution

$$\sum_{k=0}^{L-1} x(k) \cdot x^{-1}(n-k) = C \cdot \delta(n-n_0) \qquad (6)$$

yields a (potentially) scaled and time-shifted unit impulse. Then the convolution of the measured signal $y(n)$ and the inverse signal $x^{-1}(n)$ will give the scaled and time-shifted impulse response

$$C \cdot h(n-n_0) = \sum_{k=0}^{L-1} y(k) \cdot x^{-1}(n-k). \qquad (7)$$

Farina proposed [5, 6] to use an exponential sine sweep

$$x(n) = \sin\left(\frac{\omega_1 \cdot (L-1)}{\log(\omega_2/\omega_1)} \cdot \left(e^{\frac{n}{L-1}\log(\omega_2/\omega_1)} - 1\right)\right) \qquad (8)$$

with an instantaneous frequency increasing exponentially from ω_1 at $n=0$ up to ω_2 at $n=L-1$. While the crest factor $\sqrt{2} \approx 1.41$ is slightly worse than for an MLS signal, it is substantially better than for an MLS signal after band-limited reconstruction.

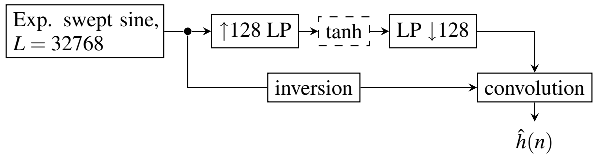


Figure 6: *Simulation setup for evaluation of the effect of non-linearities on exponential sweep measurements.*

The inverse signal can be obtained by time reversal and amplitude scaling according to

$$x^{-1}(n) = x(L-1-n) \cdot (\omega_2/\omega_1)^{\frac{-n}{L-1}} \qquad (9)$$

which approximates equation (6) with $n_0 = L-1$. The scaling factor is not given in the literature, but can be found to be

$$C = \frac{\pi L \cdot \left(\frac{\omega_1}{\omega_2} - 1\right)}{2(\omega_2 - \omega_1)\log\left(\frac{\omega_1}{\omega_2}\right)}. \qquad (10)$$

The approximation is in the sense that the convolution of $x(n)$ and $x^{-1}(n)$ results in a unit impulse (non-ideally) band-limited to the range (ω_1, ω_2), as can be seen in figure 7(a). While $\omega_2 = \pi$ may be chosen, due to the exponential sweep, we need $\omega_1 > 0$, precluding measurement of DC components and very low frequencies (in the example, $\omega_1 = 0.01$ and $\omega_2 = \pi$). The amount of overshoot and pass-band ripple may be traded off against steepness of the transition areas by applying a time-domain window on the sine-sweep, i.e. fading in and out.

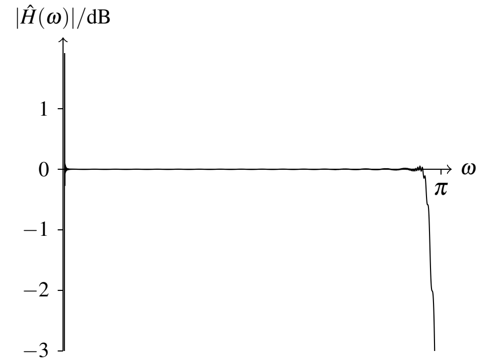
Using an exponential sweep has two major advantages: More energy is present in the measurement signal at low frequencies, and non-linearities of the system to be measured result in anti-causal components in the impulse response which can be easily separated. The concentration of energy at low frequencies is beneficial as the resulting higher accuracy is often desirable in audio applications.

The graceful handling of non-linearities stems from the fact that the time

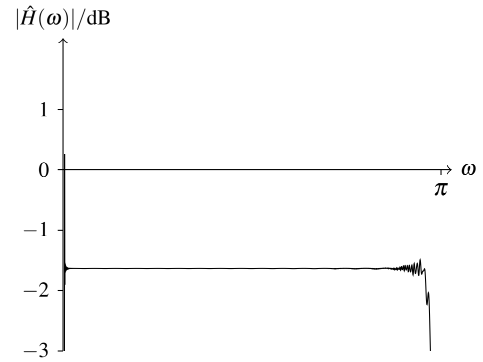
$$\Delta n(k) = \frac{\log(k+1)}{\log(\omega_2/\omega_1)}(L-1) \qquad (11)$$

for the instantaneous frequency of the sweep to rise by a factor of $k+1$ is constant. This means that if the output of the system contains the k-th harmonic of the excitation signal, that frequency will only be reached by the excitation signal $\Delta n(k)$ samples in the future. Therefore in the computed impulse response, effects of the k-th harmonic will become visible starting at $-\Delta n(k)$, i.e. in the anti-causal part.

To demonstrate the effect of non-linearities on the measurement using exponentially swept sines, an experiment similar to one for maximum length sequences was conducted, but with the clipping replaced by a tanh function, see figure 6. The tanh function results in a total harmonic distortion of about -25 dB for a full-scale sinusoid, representing a rather strong non-linearity for a system that should usually be assumed to be linear. Nevertheless, by limiting the impulse response to the causal part, the resulting transfer functions in figures 5(a) and 5(b) with and without non-linearity are very similar, mainly differing in the overall level. This loss in level is natural, as the maximum amplitude is reduced from 1 to about 0.76.



(a) Without non-linearity.



(b) With non-linearity.

Figure 7: *Resulting transfer functions as measured with the setup of figure 6.*

5. TIME-DELAY SPECTROMETRY

In time-delay spectrometry, also a swept sine is used, albeit a linearly swept one. What makes time-delay spectrometry stand out is the way the transfer function is computed [7].

Let us for a moment assume we could measure with a complex-valued measurement signal

$$x(n) = e^{\pi j n^2 r} \tag{12}$$

where r denotes the sweep rate. The instantaneous frequency is given by

$$\omega(n) = 2\pi n r. \tag{13}$$

Now when a linear system is excited with a complex sinusoid, its response is that sinusoid multiplied with $H(\omega)$, which in our case leads to

$$y(n) = H\big(\omega(n)\big)x(n) = H\big(\omega(n)\big)e^{\pi j n^2 r}. \tag{14}$$

By multiplying the received signal $y(n)$ with the complex conjugate measurement signal $x^*(n)$, the estimated transfer function is directly obtained as

$$\hat{H}\big(\omega(n)\big) = y(n)\cdot x^*(n) = H\big(\omega(n)\big)e^{\pi j n^2 r}e^{-\pi j n^2 r} = H\big(\omega(n)\big). \tag{15}$$

Time-windowing the impulse response may then be achieved by applying a low-pass to the measured transfer function.

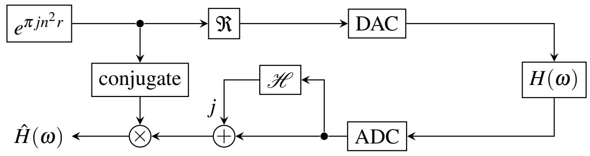


Figure 8: *Measurement system using time-delay spectrometry.*

Due to the symmetry of $H(\omega)$, it is sufficient to limit the sweep to positive frequencies, that is $0 \le \omega(n) \le \pi$ or $0 \le n \le \frac{1}{2r} = L-1$. By limiting the sweep to positive frequencies and observing that we hence have an analytic signal, the measurement signal may be made real by using

$$x(n) = \Re(e^{\pi j n^2 r}) = \cos(\pi n^2 r) \tag{16}$$

from which the original complex signal may be recovered using a Hilbert transform. The resulting time-delay spectrometry measurement system is depicted in figure 8.

Unfortunately, equation (14) is only an approximation, assuming a system which causes no significant spreading of the signal in time-domain. If the measured signal is instead expressed by the convolution sum (or by considering the Wigner distribution, as in [8]), we find

$$y(n) = \sum_{k=0}^{N} h(k)e^{\pi j(n-k)^2 r} = e^{\pi j n^2 r}\sum_{k=0}^{N} h(k)e^{\pi j k^2 r}e^{-2\pi j n k r} \tag{17}$$

and hence

$$\hat{H}\big(\omega(n)\big) = y(n)x^*(n) = \sum_{k=0}^{N} h(k)e^{\pi j k^2 r}e^{-2\pi j n k r}. \tag{18}$$

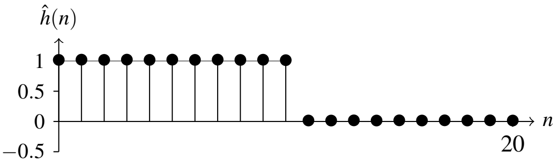
After substituting $n = \frac{\omega(n)}{2\pi r}$ this gives

$$\hat{H}\big(\omega(n)\big) = \sum_{k=0}^{N} h(k)e^{\pi j k^2 r}e^{-j\omega(n)k} \tag{19}$$

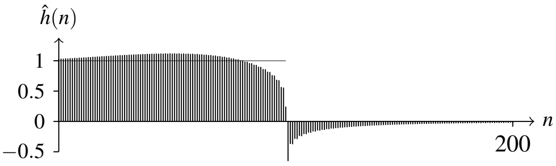
which obviously is the Fourier transform of $\hat{h}(n) = h(n)e^{\pi j n^2 r}$. In order for this to be a valid approximation of $h(n)$, $r \ll \frac{1}{N^2}$ must hold, which quickly results in very low sweep rates r and thus impractically long measurement durations. To exemplify the problem, an experiment was conducted naively employing the measurement system of figure 8 to measure a known system with an impulse response consisting of $N+1$ ones for different values of N. The sweep length was chosen as $L = 32\,768$ with $r = \frac{1}{2(L-1)} \approx 1.5259\times 10^{-05}$. As seen in figure 9(a), the measured impulse response closely resembles the original one for $N = 10$. For $N = 100$, the result shown in figure 9(b) is still a reasonable approximation, while the result for $N = 1000$ depicted in figure 9(c) is clearly unusable, as $r > \frac{1}{N^2}$.

It should be noted that originally, time delay spectrometry was proposed for measuring relatively short impulse responses, namely of loudspeakers; in fact, late parts of the impulse response were suppressed by low-pass filtering the transfer function, corresponding to windowing the impulse response.

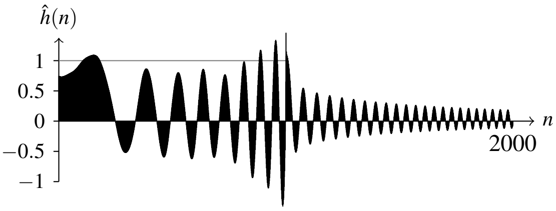
One obvious solution when long impulse responses are to be measured is to calculate $\hat{h}(n)$ and multiply by $x^*(n)$ to obtain $h(n) = \hat{h}(n)\cdot x^*(n)$. However, as the $\hat{H}(\omega)$ of equation (17) is not symmetric anymore, it is not sufficient to measure only for



(a) $N = 10$.



(b) $N = 100$.



(c) $N = 1000$.

Figure 9: *Resulting impulse responses* $\hat{h}(n)$ *obtained using time-delay spectrometry for an original impulse response consisting of* $N+1$ *ones.*

$0 \leq \omega \leq \pi$ and hence, a real-valued measurement signal may no longer be used. In practice, this means that the real and imaginary part of the measurement signal have to be applied to the system individually and recombined to a complex signal afterwards. The start and end points need further attention, the easiest solution being to extend the sweep beyond the original measurement length by at least the effective length of the impulse response. By employing these techniques, it is possible to exactly recover the original impulse responses in the above mentioned experimental setup.

6. CONCLUSIONS

Three widely used impulse response measurement techniques were critically reviewed for their applicability to real-word problems: maximum length sequences, exponentially swept sines and time delay spectrometry.

The measurement using maximum length sequences has very low computational demands as both generation of the measurement sequence and the necessary cross-correlation may be performed with very efficient algorithms. However, the method is quite sensitive to non-linear distortion so great care must be taken to stay within the linear amplitude range of the analog signal path. In particular, enough head-room to allow for the over-shoots that may result from the reconstruction low-pass must be provided.

Measuring with an exponentially swept sine, on the other hand, handles non-linearities much more gracefully, allowing higher measurement amplitudes, reducing the time needed to achieve a desired accuracy. Furthermore, the exponential sweep provides better accuracy at lower frequencies at the cost of higher frequencies, which may often be beneficial as it matches human perception. The disadvantage of the exponential sweep is that it necessarily has a non-zero lowest frequency under which no measurement is possible. The computational demands are significantly higher, especially for the convolution which has to be computed.

Time delay spectrometry has the curious property that no correlation but only multiplication with the measurement signal is necessary — and a Hilbert transform if a real-valued measurement signal is desired. Furthermore, no Fourier transform is necessary if the transfer function is the goal of the measurement, which was a real advantage at the time the method was invented. Unfortunately, only relatively short impulse responses may be measured with the original method. While a remedy that allows measurement of long impulse responses is presented in this paper, the involved complications make the method questionable when compared to the others, as no real advantages remain.

7. REFERENCES

[1] M. R. Schroeder, *Number Theory in Science and Communication*, Springer, New York, 2nd enlarged edition, 1986.

[2] J. Borish and J. B. Angell, “An efficient algorithm for measuring the impulse response using pseudorandom noise,” *J. Aud. Eng. Soc.*, vol. 31, no. 7/8, pp. 478–488, Aug. 1983.

[3] G.-B. Stan, J.-J. Embrechts, and D. Archambeau, “Comparison of different impulse response measurement techniques,” *J. Aud. Eng. Soc.*, vol. 50, no. 4, pp. 249–262, Apr. 2002.

[4] J. Vanderkooy, “Aspects of MLS measuring systems,” *J. Aud. Eng. Soc.*, vol. 42, no. 4, pp. 219–231, Apr. 1994.

[5] A. Farina, “Simultaneous measurement of impulse response and distortion with a swept-sine technique,” in *108th AES Convention*, Paris, France, Feb. 19-24 2000.

[6] A. Farina, “Advancements in impulse response measurements by sine sweeps,” in *122nd AES Convention*, Vienna, Austria, May 5-8 2007.

[7] R. C. Heyser, “Acoustical measurements by time delay spectrometry,” *J. Aud. Eng. Soc.*, vol. 15, no. 4, pp. 370–382, Oct. 1967.

[8] M. A. Poletti, “Linearly swept frequency measurements, time-delay spectrometry, and the Wigner distribution,” *J. Aud. Eng. Soc.*, vol. 36, no. 6, pp. 457–468, June 1988.