Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

A Novel Approach for Impulse Response Measurements in Environments with Time-Varying Noise

Aulis Telle and Peter Vary

Institute of Communication Systems and Data Processing, RWTH Aachen University, Aachen, Germany

PACS: 43.20.Ye, 43.58.Gn, 43.60.Cg

ABSTRACT

A typical measure taken against unavoidable noise during the measurement of an impulse response of an acoustical system is to repeat the measurement a couple of times and coherently average the distinct measurements. With stationary noise this is the optimal solution which gives an improvement of the signal-to-noise ratio of about 3 dB per doubling of the number of measurements averaged. If transient or time varying noise components are present the resulting accuracy suffers significantly. In this contribution, the authors propose a weighted averaging approach where the distinct measurements are weighted according to the present noise power. The noise power is estimated with the so-called SWiC method (sliding window correlation) recently published by the authors.

INTRODUCTION

The measurement of the impulse response of a time-invariant acoustical transmission path must often be performed in adverse noisy environments. The most common solution for improving the signal-to-noise ratio (SNR) of the measured impulse response is to repeat the excitation signal periodically and average the periods of the system response. By this means, the measurement SNR can be increased by up to 3 dB per doubling of the number of periods. To achieve the maximum increase of the SNR the noise signal must be white, stationary, and statistically independent of the excitation signal.

If these conditions are not met, the overall SNR can decrease dramatically. Specifically, in the presence of transient noise components averaging periods with greatly differing SNRs mostly results in an overall SNR being significantly lower than the SNR of the "best" period. This effect can render a long measurement with many periods useless.

In this paper a novel method for performing an impulse response measurement over multiple periods is presented which is robust against fluctuating and transient noise. It makes use of an algorithm recently published by the authors [7] for estimating the SNR during the measurement. With the knowledge of an accurate estimate of the noise power it is possible to introduce noise power dependent weighting factors per period into the averaging process. Such weighting factors are derived and proved to be the optimum factors in the sense of maximizing the SNR of the resulting impulse response. The common averaging approach is contained in the presented method as a special case when the weighting factors are equal for all periods. This happens when the noise power is constant throughout the whole measurement procedure.

First, the method and the excitation signals used throughout this paper for impulse response measurements are described. Then important aspects of the theory of coherent averaging are summarized and extended to weighted averaging. The optimality in terms of maximum SNR of the weighted averaging approach is proved for the case of perfect knowledge of the noise power in every period during the measurement. As a good estimate of the noise power is crucial for the method to be practical the theory of the proposed noise estimation method is summarized. Finally, the simulation and measurement setups and the respective results are presented.

IMPULSE RESPONSE MEASUREMENT

To describe an acoustic system, the impulse response of the transmission path between the sound source and the sound sink is of interest. Many measures like decay time and the magnitude spectrum can be deduced from it. There exist a great variety of measurement algorithms for determining the characteristics of an acoustical system.

Frequency domain deconvolution

One of the most widely used flexible and universal methods which will be used throughout this contribution is the frequency domain deconvolution. For a single shot measurement the excitation signal is emitted once and the response is recorded for at least the duration of the expected length of the system impulse response. The excitation and response signals are then transformed into the frequency domain. The response spectrum is divided by the excitation spectrum and finally transformed back into the time domain. In a noise free system this results in the impulse response of the system under test. This is depicted in Figure 1.

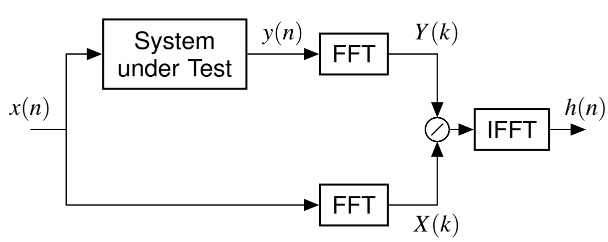

Figure 1: Impulse response measurement by frequency domain deconvolution.

To lower the impact of noise by means of coherent averaging, the single shot measurement can be repeated a couple of times.

The resulting impulse responses (or the measurement signals) are then averaged. If the excitation signal is at least as long as the expected impulse response the excitation signal can be emitted multiple times without a pause in between. After that, the system response is split into parts with the length of one period of the excitation. The first part is dropped, because here the system is not yet in a steady state. The remaining parts can be averaged over time, as will be discussed below.

Excitation signal

For a general purpose impulse response measurement without knowledge of the system under test, a broadband excitation signal has to be chosen, i.e., it should have significant energy at all frequencies. This is especially true if a deconvolution approach is used for system identification as otherwise the needed inverse signal does not exist or the approximation leads to large errors in the resulting impulse response.

For this contribution we evaluated the presented algorithms with many different excitation signals ranging from maximum-length sequences [5] over perfect sequences [3] to sweeps (logarithmic, linear [2, 4] and "perfect" [6]). As the results do not depend significantly on the actual excitation signal only the simulation and measurement results for perfect sweeps are shown.

Perfect Sweep

Perfect sweeps [6] are linear sweeps with a perfectly constant magnitude spectrum. They are constructed in the frequency domain by setting the magnitude to a constant and constructing the phase so that its absolute value increases quadratically. Thus, the perfect sweep has linear group delay $\tau_G(f)$

$$\tau_G(f) = \tau_G(0) + f \cdot \frac{\tau_G(f_g) - \tau_G(0)}{f_g} \tag{1}$$

with f_g being the upper frequency bound of the sweep and f the frequency in the range $0 \leq f \leq f_g$. For the phase $\Phi(f)$ this yields

$$\Phi(f) = -\int_0^f \tau_G(\zeta)\mathrm{d}\zeta \tag{2}$$

$$= -\left(f \cdot \tau_G(0) + \frac{1}{2} f^2 \cdot \frac{\tau_G(f_g) - \tau_G(0)}{f_g} \right). \tag{3}$$

Finally the constructed spectrum is transformed into the time domain.

In contrast to traditional sweep construction no countermeasures are needed to reduce temporal aliasing as it guarantees a smooth transition between successive periods when emitted periodically. In combination with the perfectly flat spectrum this is especially advantageous for the measurement of time variant systems. For the measurement of a static system, perfect sweeps have almost the same properties as normal linear sweeps.

COHERENT AVERAGING

Every impulse response measurement suffers from noise in one or the other way. Typically the microphones and the A/D-converters contribute to measurement noise. Additionally, there is normally some unavoidable background noise present at the location of the measurement. As long as such noise is stationary and statistically independent of the excitation signal and the system under test is sufficiently static, one can reduce its influence by emitting the excitation signal periodically and coherently averaging successive periods of the measured signal.

Let N be the period length, $y(n)$ the noiseless system response to the excitation signal $x(n)$, $r(n)$ the noise signal and $\tilde{y}(n) = y(n) + r(n)$ the noisy system response. When the excitation signal is periodic, the noiseless system response of a linear time-invariant system must also be periodic

$$y(n) = y(n-N). \tag{4}$$

For the noisy system response we get

$$\tilde{y}(n) = y(n) + r(n) \tag{5}$$

$$\tilde{y}(n-N) = y(n) + r(n-N). \tag{6}$$

With unweighted averaging of the distinct periods of the noisy system response $\tilde{y}(n)$ the averaged system response $\bar{y}(n)$ of length N is then given by

$$\bar{y}(n) = \frac{1}{M} \sum_{i=0}^{M-1} \tilde{y}(n+iM) \tag{7}$$

$$= \frac{1}{M} \sum_{i=0}^{M-1} y(n+iM) + r(n+iM) \tag{8}$$

$$= \frac{1}{M} \sum_{i=0}^{M-1} y(n) + r(n+iM) \tag{9}$$

$$= y(n) + \frac{1}{M} \sum_{i=0}^{M-1} r(n+iM) \tag{10}$$

with $0 \leq n < N$ and M the number of periods to average.

This means, that the energy of the clean system response being part of the noisy system response is not altered. The noise component is represented by the sum of the noise segments divided by the number of averaged periods.

Now let the noise signal $r(n)$ be an ergodic random signal where all the segments of length N are independent of each other. With $N \to \infty$ the power in the sum of the noise segments equals the power in one noise segment multiplied by the number of segments. With the scaling factor $1/M$ from the averaging operation the overall noise power Ψ_a in the averaged noisy system response results in

$$\Psi_\mathrm{a} = \frac{1}{M}\Psi \tag{11}$$

with

$$\Psi = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} r(k). \tag{12}$$

Thus, with an ergodic noise signal, i.e., an equal noise power in all noise segments, the noise power can be lowered by a factor of $\frac{1}{M}$. This is the well known rule of 3 dB gain in SNR per doubling the number of averaging periods.

WEIGHTED COHERENT AVERAGING

If the noise power fluctuates over time, the unweighted averaging is suboptimal, though. The SNR can even degrade dramatically if the noise power in the averaged periods differs significantly. To model changing noise power lets assume that the noise in each of the M periods has its source in a different ergodic noise process $r_i(n)$ with distinct constant power Ψ_i with $0 \leq i < M$. Introducing the weights c_i into (7) yields

$$\bar{y}_\mathrm{w}(n) = \sum_{i=0}^{M-1} c_i y(n) + c_i r(n+iM) \tag{13}$$

The weights are constrained by

$$\sum_{i=0}^{M-1} c_i = 1. \tag{14}$$

so that the weighted average $\bar{y}_w(n)$ becomes

$$\bar{y}_w(n) = y(n) + \sum_{i=0}^{M-1} c_i r(n+iM). \tag{15}$$

The power Ψ_w of the weighted sum of the distinct noise signals $r_i(n)$ with the weights c_i yields

$$\Psi_w = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \left(\sum_{i=0}^{M-1} c_i r_i(k) \right)^2 = \sum_{i=0}^{M-1} c_i^2 \Psi_i. \tag{16}$$

The goal is to find weights c_i that minimize the power of the summed noise signal

$$\Psi_w = \sum_{i=0}^{M-1} c_i^2 \Psi_i \to \min. \tag{17}$$

This results in a constrained optimization problem which can be solved by means of the method of Lagrange multipliers [1]. The Lagrangian Λ is given by

$$\Lambda(c_0, \ldots, c_{M-1}, \lambda) = \sum_{i=0}^{M-1} c_i^2 \Psi_i + \lambda \left(1 - \sum_{k=0}^{M-1} c_k \right) \tag{18}$$

with the partial derivatives

$$\frac{\partial \Lambda(c_0, \ldots, c_{M-1}, \lambda)}{\partial c_i} = 2c_i \Psi_i - \lambda \quad \text{with } 0 \le i < M \tag{19}$$

and

$$\frac{\partial \Lambda(c_0, \ldots, c_{M-1}, \lambda)}{\partial \lambda} = 1 - \sum_{i=0}^{M-1} c_i \tag{20}$$

Setting the partial derivatives to zero yields the system of equations:

$$c_i = \frac{\lambda}{2\Psi_i} \quad \text{with } 0 \le i < M \tag{21}$$

$$\sum_{i=0}^{M-1} c_i = 1 \tag{22}$$

Substituting (21) into (22) and solving the equation for λ results in

$$\lambda = \frac{2}{\sum_{i=0}^{M-1} \frac{1}{\Psi_i}} \tag{23}$$

For the weights c_i this gives

$$c_i = \frac{1}{\Psi_i \sum_{k=0}^{M-1} \frac{1}{\Psi_k}} \quad \text{with } 0 \le i < M \tag{24}$$

Substituting (24) into (16) gives the closed form for the resulting averaged noise power

$$\Psi_w = \sum_{i=0}^{M-1} \left(\frac{1}{\Psi_i \sum_{k=0}^{M-1} \frac{1}{\Psi_k}} \right)^2 \Psi_i \tag{25}$$

$$= \sum_{i=0}^{M-1} \frac{1}{\Psi_i \left(\sum_{k=0}^{M-1} \frac{1}{\Psi_k} \right)^2} \tag{26}$$

$$= \sum_{i=0}^{M-1} \frac{1}{\Psi_i \left(\frac{\sum_{k=0}^{M-1} \prod_{m=0, m\neq k}^{M-1} \Psi_m}{\prod_{k=0}^{M-1} \Psi_k} \right)^2} \tag{27}$$

$$= \frac{1}{\left(\sum_{i=0}^{M-1} \prod_{k=0, k\neq i}^{M-1} \Psi_k \right)^2} \sum_{i=0}^{M-1} \frac{\left(\prod_{k=0}^{M-1} \Psi_k \right)^2}{\Psi_i} \tag{28}$$

$$= \frac{\prod_{i=0}^{M-1} \Psi_i}{\left(\sum_{i=0}^{M-1} \prod_{k=0, k\neq i}^{M-1} \Psi_k \right)^2} \sum_{i=0}^{M-1} \prod_{k=0, k\neq i}^{M-1} \Psi_k \tag{29}$$

$$= \frac{\prod_{i=0}^{M-1} \Psi_i}{\sum_{i=0}^{M-1} \prod_{k=0, k\neq i}^{M-1} \Psi_k} \tag{30}$$

Because the second partial derivatives of (18)

$$\frac{\partial^2 \Lambda(c_0, \ldots, c_{M-1}, \lambda)}{\partial c_i^2} = 2\Psi_i \quad \text{with } 0 \le i < M \tag{31}$$

and

$$\frac{\partial^2 \Lambda(c_0, \ldots, c_{M-1}, \lambda)}{\partial \lambda^2} = 0 \tag{32}$$

are all greater or equal to zero, the previously found extremum really is an absolute minimum. Thus, the weighted sum (16) with the weights (24) results in the minimally achievable noise power.

NOISE POWER ESTIMATION

To actually perform the optimal weighted averaging, the noise power in the distinct periods has to be estimated as the actual noise signal is not known. The noise power estimation method recently published by the authors [7] is designed for measurement setups with periodically repeated excitation signal. It is used for noise power estimation in this paper and hence will be sketched in the following.

In order to derive a method for estimating the noise power we recall the properties (5) and (6) of the noisy system response $\tilde{y}(n)$

$$\tilde{y}(n) = y(n) + r(n)$$
$$\tilde{y}(n-N) = y(n) + r(n-N).$$

and define a signal $e(n)$ with

$$e(n) = \tilde{y}(n) - \tilde{y}(n-N) \tag{33}$$

$$= r(n) - r(n-N). \tag{34}$$

Now consider the expectation of the product of $e(n)$ and the N-shifted version of it $e(n+N)$

$$\begin{aligned} \mathrm{E}\{e(n)\cdot e(n+N)\} &= \\ &= \mathrm{E}\{(r(n)-r(n-N))\cdot(r(n+N)-r(n))\} \\ &= -\mathrm{E}\{r^2(n)\}+\mathrm{E}\{r(n)\cdot r(n+N)\} \\ &\quad +\mathrm{E}\{r(n)\cdot r(n-N)\} \\ &\quad -\mathrm{E}\{r(n+N)\cdot r(n-N)\}. \end{aligned} \tag{35}$$

For a stationary white noise signal all terms but the first become zero and we get

$$\mathrm{E}\{e(n)\cdot e(n+N)\} = -\mathrm{E}\{r^2(n)\} \tag{36}$$

This relation shows, that the autocorrelation of the signal $e(n)$ at shift N is equal to the negated power of the noise signal $r(n)$. Thus, the power Ψ_i of period i of the noise signal can be approximated by

$$\begin{aligned} \Psi_i &\approx -c_e^{(N,N)}(i\cdot N) \\ &= -\frac{1}{N}\sum_{k=0}^{N-1} e(i\cdot N+k)\cdot e(i\cdot N+k+N). \end{aligned} \tag{37}$$

where $c_x^{(W,D)}(n)$ is the so called SWiC (sliding window correlation) of the signal x with window length W and displacement D (see [7]).

SIMULATION

For evaluating the presented method for weighted averaging, simulations as well as real measurements were performed. In this section the simulation setup and results will be presented. The next section will deal with the real measurements.

The system impulse response of the simulated system, which is shown in Figure 2, was constructed from random white Gaussian noise with a length of $N = 10000$ samples fading out exponentially and normalized to have a power of one. Different excitation signals were used: a maximum length sequence, a logarithmic sweep, and a sweep with a perfectly flat spectrum, a so-called perfect sweep ([6]). As mentioned earlier only the results using the perfect sweep are presented.

The excitation signal was repeated $M = 17$ times and filtered with the system impulse response. Then a noise signal was added. To simulate ambient stationary noise random white gaussian noise with a power of -40 dB was added. For transient

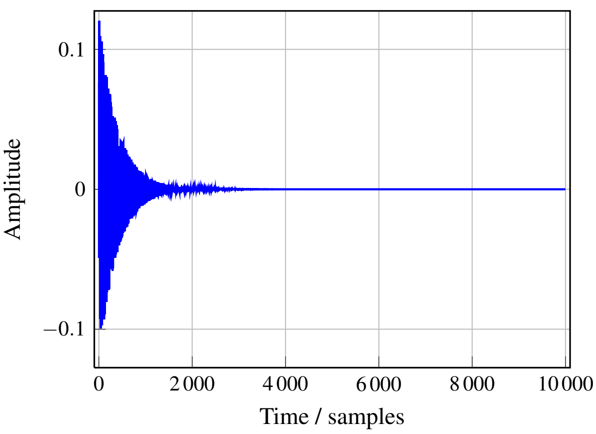

Figure 2: System impulse response for the simulation.

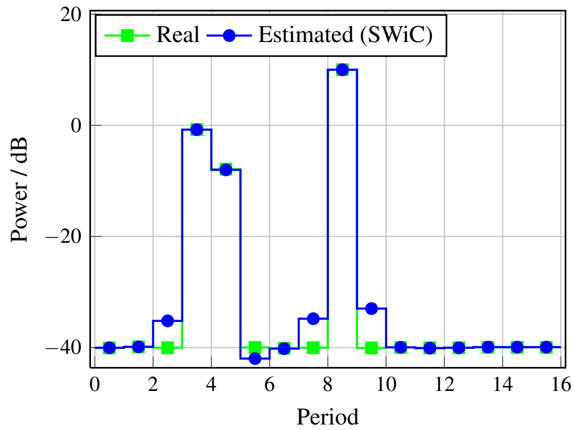

Figure 3: Real and estimated noise power.

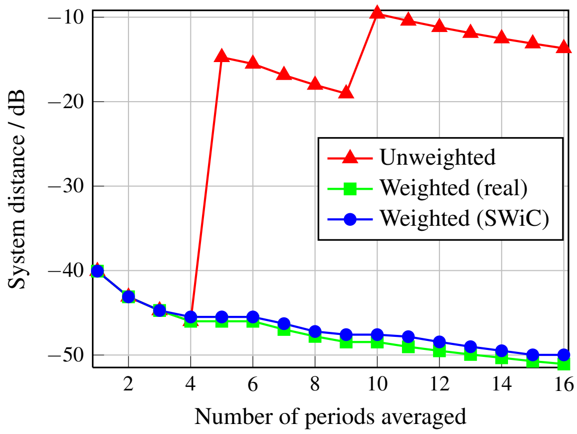

Figure 4: Comparison of different averaging approaches by means of system distance between the real impulse response and the estimated impulse response averaged over an increasing number of periods.

noise, short bursts of random white Gaussian noise with increased power were added. Figure 3 shows the real and the estimated power of the noise signal.

The estimated impulse responses were determined for every period by frequency domain deconvolution as described earlier in this paper. Because the system is not yet stabilized in the first excitation period, that period was dropped. For normal averaging the impulse responses were summed and divided by the number of impulse responses used for averaging. For weighted averaging the impulse responses were weighted according to the real and the estimated noise power present in the corresponding period of the noisy system response.

In order to see the impact of transient noise, the averaged impulse responses were calculated for the averaging of 1, 2, 3, and so on periods.

Simulation results

The performance of the weighting algorithms was measured by means of the system distance D between the estimated impulse response $\tilde{h}(n)$ and the real impulse response $h(n)$

$$D = \frac{\sum_{k=0}^{N-1}\left(h(k)-\tilde{h}(k)\right)^2}{\sum_{k=0}^{N-1} h^2(k)}. \tag{38}$$

Figure 4 shows the results of normal averaging and weighted averaging with the real noise power and with the estimated noise power used to calculate the weights.

MEASUREMENT

To evaluate the benefits of the proposed averaging method in practice, a set of real measurements were performed. The acoustical transfer path to be measured was from a loudspeaker in a dummy head's mouth to a microphone located near the ear of the dummy head. The acoustical components were connected to professional A/D- and D/A-converters (manufactured by RME). The environment was a normal office with typical background noise (e.g., computer fans and hard disk drive). Transient noise was produced by dropping a ruler and knocking at the door.

The excitation signal types and methods for calculating the system impulse response were the same as in the simulation setup described above. The length of one period of the excitation signals was approximately 4 seconds at a sample rate of 48 kHz. The number of periods was 16. For each excitation signal a set of two measurements was performed. First the noise level was kept as low as possible to get a reference measurement for calculating a reference impulse response by averaging over periods 2 through 16. During the second measurement transient noises were produced as described above.

Measurement results

Like in the simulation setup the averaging results for the different excitation signals are quite similar. Thus, only the results for the perfect sweep are presented here.

Figure 5 shows the noise power during the measurement estimated by means of the SWiC. The true noise power is not known. The measurements are evaluated in the same way as the simulation results only that instead of the true impulse response, which is not known, the reference impulse response is used. The resulting system distances are shown in Figure 6.

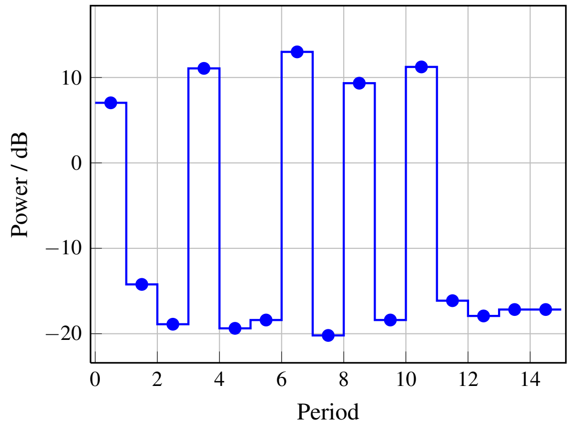

Figure 5: Estimated (SWiC) noise power during real measurement.

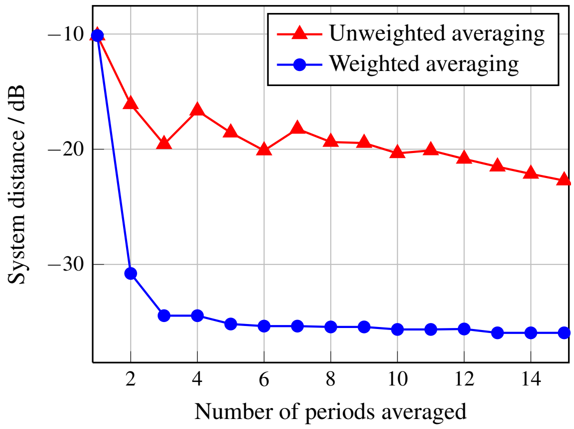

Figure 6: System distance to reference impulse response for unweighted and weighted averaging for an increasing number of periods averaged.

CONCLUSIONS

The simulations as well as the measurement results show clearly, that weighted averaging is by far superior to unweighted averaging in terms of achievable system distance. In the case of stationary background noise the weights are set to be all equal, i.e., the traditional averaging is included as a special case.

If the noise has transient components with high power, the traditional averaging fails. The transient components render the resulting impulse responses nearly useless. By weighting the distinct periods so that periods with high noise power have less impact, the resulting impulse responses can be improved significantly.

To achieve a superior SNR in the resulting system impulse response, a precise estimate of the noise power is needed. The noise estimation method based on the SWiC (Sliding window correlation) follows the real noise power quite closely and thus is well suited for the weighted averaging approach. Compared to the results with the real noise power the system distance degrades only slightly.

Thus, the weighted averaging method proposed in this paper makes it possible to retrieve high quality impulse responses with high SNR even in changing noise conditions.

REFERENCES

[1] Dimitri P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, 1996. ISBN: 1-886529-04-3.

[2] Angelo Farina. "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique". *AES 108th Convention*. Paris, France, Feb. 2000.

[3] Hans Dieter Lüke. "Sequences and Arrays with Perfect Periodic Correlation". *IEEE Transactions on Aerospace and Electronic Systems* 24.3 (May 1988), pp. 287–294. ISSN: 0018-9251. DOI: 10.1109/7.192096.

[4] Swen Müller and Paulo Massarani. "Transfer-Function Measurement with Sweeps". *Journal of the Audio Engineering Society* 49.6 (June 2001), pp. 443–471.

[5] Douglas D. Rife and John Vanderkooy. "Transfer-Function Measurement with Maximum-Length Sequences". *Journal of the Audio Engineering Society* 37.6 (June 1989), pp. 419–444.

[6] Aulis Telle, Christiane Antweiler, and Peter Vary. "Der perfekte Sweep - Ein neues Anregungssignal zur adaptiven Systemidentifikation zeitvarianter akustischer Systeme". *Proceedings of German Annual Conference on Acoustics (DAGA)* (Mar. 2010).

[7] Aulis Telle and Peter Vary. "A new Method for SNR-Estimation in Impulse Response Measurements". *Proceedings of European Signal Processing Conference*. Glasgow, Scotland, Aug. 2009.