PERFECT-SWEEP NLMS FOR TIME-VARIANT ACOUSTIC SYSTEM IDENTIFICATION

C. Antweiler, A. Telle, P. Vary

Institute of Communication Systems
and Data Processing, RWTH Aachen University,
52056 Aachen, Germany
antweiler@ind.rwth-aachen.de

G. Enzner

Institute of Communication Acoustics
Ruhr-University Bochum,
44780 Bochum, Germany
gerald.enzner@rub.de

ABSTRACT

Fast and robust acoustic system identification is still a research topic of interest, because of the typically time-variant nature of acoustic systems and the natural performance limitation of electroacoustic measurement equipment. In this paper, we propose NLMS-type adaptive identification with perfect-sweep excitation. The perfect-sweep is derived from the more general class of perfect sequences and, thus, it inherits periodicity and especially the desired decorrelation property known from perfect sequences. Moreover, the perfect-sweep shows the desirable characteristics of swept sine signals regarding the immunity against non-linear loudspeaker distortions. On this basis, we first demonstrate the fast tracking ability of the perfect-sweep NLMS algorithm via computer generated simulation of a time-variant acoustic system. Then, the robustness of the perfect-sweep NLMS algorithm against non-linear characteristics of real measurements in a time-invariant case is presented. By finally addressing the measurement of quasi-continuous head-related impulse responses, we face the combined challenge of time-variant and possibly non-linear distorted acoustic system identification in a real application scenario and we can demonstrate the superiority of the perfect-sweep NLMS algorithm.

Index Terms— Adaptive filters, system identification, perfect sequences, acoustics, head-related transfer functions

1. INTRODUCTION

The problem of tracking a time-variant acoustic transmission path arises in many application areas. One approach relies on an NLMS-type (*normalized least mean square*) adaptive filtering algorithm in combination with its *optimal* excitation signal, i.e., a *perfect sequence* (PSEQ) [1, 2], enabling maximum convergence speed. It rests upon the assumption that the changes of the acoustic system under test are slow in comparison to the time available for its identification. In this context, besides the rapid convergence speed, an appropriate choice of system parameters is essential. Owing to its simplicity and beneficial properties we used this technique to track time-variant systems in diverse applications [3, 4, 5, 6].

It is well understood that the choice of the excitation signal takes immense influence on the performance of the adaptation process. To achieve a sufficiently high signal-to-noise ratio (SNR), the excitation signal must have a high energy uniformly spread over the frequency range of interest. Furthermore, in practice, besides a certain amount of noise, there are typically also distortions due to loudspeaker non-linearities, non-ideal converters, amplifiers, etc., and this restricts the power of the excitation signal [7]. To reduce this limitation in acoustics, *sweep* signals, avoiding non-linear distortions to a large extent, have been established for a wide range of mea-

surement tasks, e.g., [8, 9]. In this paper, we propose NLMS-type adaptive identification with *perfect-sweep* excitation – in this context a new excitation signal. With this specific deterministic excitation the NLMS algorithm exhibits the same rapid convergence speed as the *recursive least squares* (RLS) algorithm and, in addition, the extremely low complexity of the NLMS algorithm. Furthermore, the use of perfect-sweeps instead of other perfect sequences such as, e.g., *ternary* PSEQs [10], which have been used so far, results in a certain robustness against non-linear loudspeaker distortion.

The basic measurement setup and the corresponding tracking algorithm are presented in Sec. 2. In Sec. 3 the common characteristics of perfect sequences (PSEQs) are introduced. We motivate the use of the new subclass of PSEQs, the perfect-sweeps and show how to construct them. In Sec. 4 the advantages of a perfect-sweep excitation are demonstrated with computer-generated data and more realistic with data from acoustic measurements. Furthermore, we are dealing with an actual application, the acquisition of head-related impulse responses (HRIRs), where high-fidelity measurement of the binaural transmission paths is desired. An interesting approach for spatially continuous acquisition was presented in [5, 6], relying on dynamical identification of the HRIRs with a rotating dummy head and the use of adaptive filters. In this approach, we therefore require both fast and robust adjustment of the filter coefficients. In Sec. 5 we will demonstrate to which extent perfect-sweeps meet these requirements.

2. TIME-VARIANT SYSTEM IDENTIFICATION

The discrete time model in Fig. 1 depicts a system for the identification of an unknown time-variant acoustic transmission path by means of an adaptive filter. The recorded signal $y(k)$ at time instant k can be expressed as

$$y(k) = \mathbf{g}^T(k)\,\mathbf{x}(k) + n(k)$$

with the real-valued excitation vector

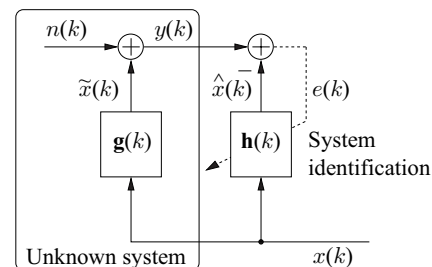$$\mathbf{x}(k) = (x(k), x(k-1), \ldots, x(k-N+1))^T, \qquad (1)$$

Fig. 1. Discrete time model for the NLMS-based identification

the vector representation of the time-varying acoustic system,

$$\mathbf{g}(k) = (g_0(k), \, g_1(k), \, \ldots, \, g_{(N-1)}(k))^T \,,$$

and the observation noise $n(k)$ at the microphone.

The system identification method relies on the *normalized least mean-square* (NLMS) algorithm which is a linear adaptive filtering algorithm that consists of an *adaptive process* performing the adjustment of the filter taps, i.e.,

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu_0 \frac{e(k)\,\mathbf{x}(k)}{||\mathbf{x}(k)||^2} \,,$$

and of a *filtering process* calculating the estimation error between the recorded response $y(k)$ and the adaptive filter ouput $\hat{x}(k)$, i.e.,

$$e(k) = y(k) - \mathbf{h}^T(k)\,\mathbf{x}(k) \,.$$

In most applications, the acoustic systems under test are of infinite length. As the impulse response of the adaptive filter is restricted in length, the resulting truncation of the last samples of the acoustic path leads to a systematic error due to the periodic excitation signal $x(k)$. This effect has been intensively discussed in [6] and will not be considered in this paper, as we provide sufficient length N for vector $\mathbf{h}(k)$.

The aim of the identification process is to achieve the best possible match between the adaptive filter represented by $\mathbf{h}(k)$ and the system under test $\mathbf{g}(k)$. The key to obtain a rapid convergence speed – especially in case of time-variant acoustic systems – is to use the NLMS algorithm in combination with its *optimal* excitation signal, i.e., with *perfect sequences* (PSEQs) [11, 12].

3. PSEQS: PSEUDO-NOISE VS. SWEEPS

Perfect sequences (PSEQs) [11, 12] are time-discrete signals of finite length M that are repeated periodically and posses a perfect, i.e., impulse-like periodic autocorrelation function

$$\tilde{\varphi}_{pp}(\lambda) = \sum_{m=0}^{M-1} p(m)p(m+\lambda) = \begin{cases} E_p & \lambda = 0 \pmod{M} \\ 0 & \lambda \neq 0 \pmod{M}, \end{cases} \quad (2)$$

where E_p is the energy of one period of the sequence.

3.1. Ternary Perfect Sequences

So far, in our applications we mostly used *ternary* PSEQs [10], belonging to the class of pseudo-noise sequences. They show a very high energy efficiency, i.e., a small crest factor.

However, Müller and Massarani [7] showed, that the theoretically high energy efficiency of pseudo-noise sequences cannot be reached in practical measurement setups, because the D/A-converters introduce significant distortion for high amplitudes. Furthermore, it is argued that pseudo-noise signals might induce additional distortions due to loudspeaker non-linearities. As a consequence, in acoustics *sweeps* are figured out as the preferable choice for the majority of measurement tasks. However, the sweep signals do usually not show the perfect impulse-like auto-correlation function (2) as needed for our application scenarios.

In [13], a new class of so-called *perfect-sweeps* is defined which combines both, the characteristics of a PSEQ and a sweep signal. Acoustical measurements using these perfect-sweeps as excitation signal for the NLMS-based identification process can thus exploit the maximum convergence speed needed to track time-variant systems *and* the higher immunity against distortions.

3.2. Construction of Perfect-Sweeps

Fortunately, constructing a perfect-sweep is quite easy [13]. Actually, a perfect-sweep is a time-stretched pulse [14] constructed in the frequency domain with an ideally flat magnitude spectrum and a linear group delay, i.e., a quadratic phase. The general construction
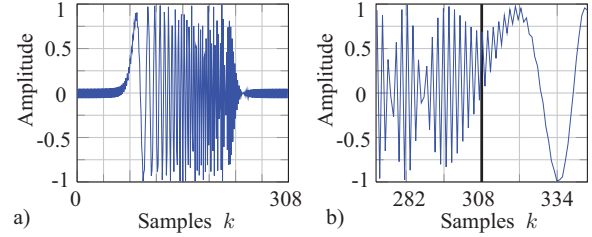

Fig. 2. a) Time-stretched pulse with $M = 308$ and $m = M/4$. b) Transition between contiguous periods of a periodically repeated perfect-sweep; $M = 308$, $m = M/2$.

formula is given by

$$P(\nu) = \begin{cases} \exp\left(\dfrac{-j4m\pi\nu^2}{M^2}\right); & 0 \leq \nu \leq \frac{M}{2} \\ P^*(M - \nu); & \frac{M}{2} < \nu < M \,, \end{cases}$$

with frequency index ν, length M of one period of the sequence, and the factor m which determines the stretch of the time-stretched pulse. As the magnitude spectrum is flat and the phase is odd-symmetric, in the time-domain we obtain a real-valued, perfect sequence. The term stretch is used to describe the relative portion of one period of the sequence in which the main energy of the sweep is concentrated. For example, a stretch factor of $m = M/2$ means that the sweep covers the whole period and the energy is equally distributed. A stretch factor of $m = M/4$ results in a sweep which starts approximately at $n = M/4$ and ends at $n = 3M/4$, i.e., the energy of the sweep is concentrated in only one half of the period length, see, e.g., Fig. 2-a).

Typically, in acoustical measurement setups the excitation signal is emitted only once or, if the excitation signal is emitted periodically, a pause is inserted between each repetition for the acoustical system to settle before it is excited again. For such a scenario, mostly a stretch factor of $m = M/4$ or smaller is used and the sequence is windowed to avoid discontinuities at the start and end of the sequence.

However, the crest factor of such a sequence is suboptimal and if it is windowed it isn't perfect anymore. For dynamic system identification with the NLMS-algorithm a perfect, periodically repeated excitation signal is needed and a high crest factor is desired. Thus, for the construction of the perfect-sweep, the stretch factor is set to $m = M/2$ so that the sweep covers the whole sequence. This increases the energy efficiency by 3 dB when compared to a time-stretched pulse with $m = M/4$ and there is a continuous transition between sucessive periods as can be seen in Fig. 2-b).

4. PSEQ-NLMS FILTERING

In order to verify the advantages of the new excitation signal, we show the results of the system identification process for a simulated transmission path, for actual room impulse response measurements, and finally, in Sec. 5, for a practical application scenario.

4.1. Simulation with Computer-Generated Data

In the PSEQ-NLMS algorithm we periodically apply the PSEQ of length M to the system under test. The *optimal* adaptation of $\mathbf{h}(k)$ relies on $N = M$ consecutive excitation vectors $\mathbf{x}(k)$ being exactly orthogonal to each other, see (2). As a consequence, the period M of the PSEQ has to match the length N of the adaptive filter $\mathbf{h}(k)$. Assuming a stepsize of $\mu_0 = 1$ and $N = M$, the PSEQ excitation enables the NLMS algorithm to identify a linear, time-invariant, simulated system within N iterations [1].
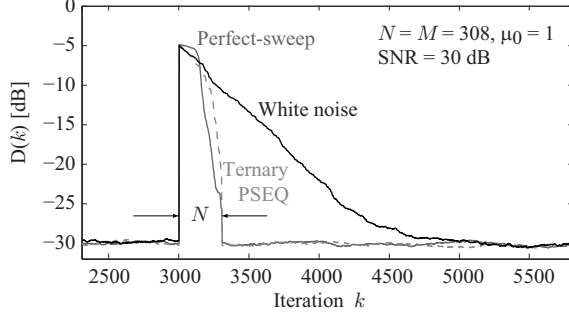
Fig. 3. System distance for different excitation signals; $\mathbf{g}$ time-invariant, except sudden change at $k = 3000$.

Fig. 3 illustrates the simulation results of an NLMS adaptation process in case of a ternary PSEQ, a perfect-sweep and additionally a *white noise* excitation signal in terms of the normalized *system distance*

$$\mathrm{D}(k) = ||\mathbf{g}(k) - \mathbf{h}(k)||^2 / ||\mathbf{g}(k)||^2 .$$

For the simulation, we chose an effective signal-to-noise ratio (SNR) of 30 dB at the microphones.

For the ternary PSEQ and the perfect-sweep excitation, Fig. 3 reflects the rapid convergence speed after a sudden change of the simulated impulse response $\mathbf{g}$ at $k = 3000$. In addition, the effect of the time-constant N can be observed. The direct comparison with the system distance achieved with white noise emphasizes that the NLMS benefits from the special correlation properties of the deterministic PSEQs.

4.2. System Identification from Real Measurements

Ususally, we have to deal with non-linearities due to the measurement equipment such as the converter, loudspeaker, microphones, and amplifier. In the experiments of this section, we perform room impulse response measurements in a professional studio box and focus on the distortions due to loudspeaker non-linearities. The effects of the other components are kept small by the careful choice of the saturation parameter for the D/A converter.

For the comparison, all considered excitation signals are normalized to the same signal power, i.e., for all excitation signals at the microphone the same SNR is guaranteed, which leads to different maximum amplitudes of the excitation signals.

Again, an instrumental objective measure of the achieved quality for the system identification is of interest. However, as in this experiment the actual room impulse response $\mathbf{g}(k)$ is not known, only the power of the identification error $\mathrm{E}\{e^2(k)\}$ can be used as an indicator for the quality of the adaptation process. It should be noted that a small value of $\mathrm{E}\{e^2(k)\}$ does not automatically correspond to a small system distance. As quality index, we define the normalized error signal attenuation

$$Q = \mathrm{E}\{y^2(k)\} / \mathrm{E}\{e^2(k)\} .$$

Fig. 4 depicts the results in terms of the quality measure Q for different *sound pressure levels* (SPLs), different excitation signals, and two loudspeakers. The observations can be summarized as follows:

- With computer-generated data, ternary PSEQ and the perfect-sweep excitation provide almost identical rapid convergence speed as illustrated in Fig. 3. However, in case of a system identification from acoustic measurements, major differences in the Q-values might be detected, as shown in Fig. 4.
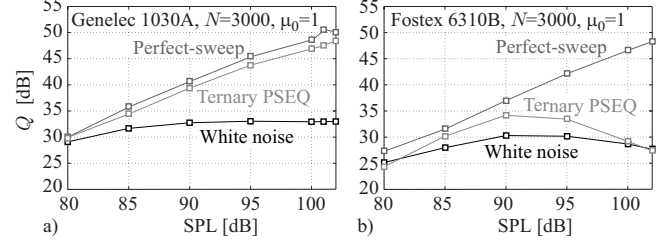

Fig. 4. Q-measure for different excitations and two loudspeakers.

- The curves in Fig. 4-a) corresponding to the loudspeaker of higher quality generally show a better performance than in Fig. 4-b).

- Increasing the sound pressure level by $\Delta_{\mathrm{SPL}} = 5$ dB goes along with an increased SNR at the microphone. Consequently, the quality index should also show an improvement of $\Delta_Q = 5$ dB. Obviously, only the curves refering to the perfect-sweep excitation prove the expected theoretic results for both loudspeakers in Fig. 4-a) and b).

- In contrast, if we lower the volume, two effects can be observed. On the one hand, the distortions due to loudspeaker non-linearities can be reduced and the quality indices Q of the different curves approach. On the other hand, the SNR at the microphone degrades, the influence of the background noise increases and the quality measure Q degrades, respectively.

As a result we can conclude that obviously, the perfect-sweep can be fed to the loudspeaker with considerably more power without introducing distortions into the adaptation process.

5. APPLICATION TO QUASI-CONTINUOUS ACQUISITION OF HRIRS

In this section we investigate the performance of the perfect-sweep NLMS algorithm in the context of quasi-continuous acquisition of *head-related impulse responses* (HRIRs) [5] as needed, e.g., for binaural rendering. For the continuous HRIR representation in the azimuth direction, the subject of interest is continuously rotating during the binaural recording. Using the input and output signals of the measurement setup, the NLMS-based identification process extracts the time-varying HRIRs at any azimuth. Besides the efficiency in terms of the time consumption for the acoustic measurement itself, the concept also benefits from avoiding the classical sampling and interpolation issue. In [6] the effect of a PSEQ excitation for this technique has been discussed. In this paper we will now focus on the influence of the proposed perfect-sweep excitation which combines the rapid tracking ability *and* a higher robustness against distortions due to loudspeaker non-linearities.

5.1. Measurement Setup

The measurements are performed with a dummy head-and-torso simulator placed in the middle of an anechoic chamber facing a loudspeaker in a distance of 2 m. The different excitation signals $x(k)$ are emitted via loudspeaker at a sampling rate of $f_s = 44.1$ kHz. The reaction of the system including anechoic chamber, outer ear, and torso is recorded with microphones located at the two ear canal entrances. A signal processing stage according to Fig. 1, provides for the left and right acoustic channel a representation of the actual binaural HRIRs in every single time instant.

Two different kinds of measurements are considered. First, a dynamic measurement for the continuous-azimuth aquisition of all
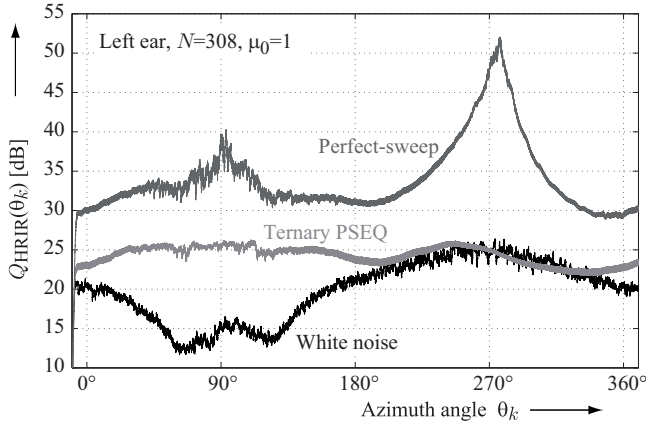
Fig. 5. Dynamic HRIR measurement for different excitations.

Fig. 6. Stationary HRIR measurement for different excitations.

HRIRs in the azimuthal plane is performed with a continuously rotating dummy head with a revolution time of $T_{360} = 20\,\text{s}$, see [5] for more details. Secondly, we present the results of a stationary measurement at $270°$ such that the left ear is exposed to maximal direct sound.

5.2. Instrumental Comparison

The results in terms of the Q-values are summarized in Fig. 5 and Fig. 6, respectively. For the experimental comparison, again all three excitation signals are normalized to the same power.

The curves belonging to the perfect-sweep excitation show the best performance in all experiments. Due to its high convergence rate and its robustness against loudspeaker non-linearities, the perfect-sweeps outperform the other excitation signals by far.

In the dynamic measurements in Fig. 5 the typical fluctuation of the Q-value over the azimuth angle can be seen. At $270°$ we detect the highest Q-values because the left ear picks up the most direct sound from the loudspeaker and thus the best SNR condition as well as the highest Q-value is achieved.

Stationary measurements in Fig. 6-a) can be compared to the position $\theta_k = 270°$ of Fig. 5 and due to symmetry reasons Fig. 6-b) to $\theta_k = 90°$, respectively. Obviously, in almost all conditions the results of the dynamic measurements range in the same order as in the static case. The only exception is detected for $\theta_k = 90°$ and *white noise* excitation. The Q-value degrades severly in the dynamic case as the NLMS algorithm does not show the rapid tracking ability as in case of a PSEQ excitation.

6. CONCLUSIONS

In this paper the perfect-sweep NLMS algorithm has been investigated for time-variant acoustic system identification. The advantage of using perfect-sweeps is that they combine the characteristics of PSEQs and sweeps, enabling a rapid convergence speed of the NLMS adaptation algorithm as well as a high robustness against distortions. The benefits have been demonstrated with simulations based on computer-generated data, with real acoustic measurements, and in an application-oriented context. The influence of different excitation signals, such as white noise, ternary PSEQs, and perfect-sweeps on the measurements has been investigated showing immense differences in the performance. The use of perfect-sweeps opens up the possibility to excite the unknown acoustic system with considerably more power without introducing non-linear distortions. Thus, the capabili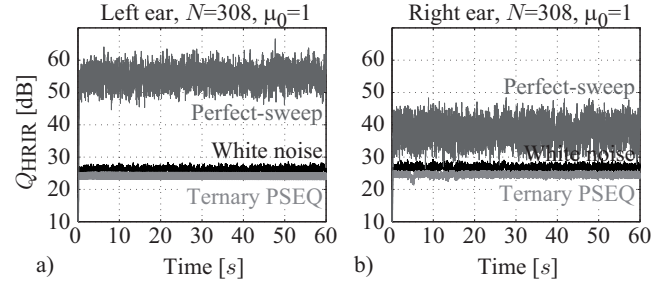ties of the given hardware components can be much better exploitet. As a consequence significant improvements in terms of the objective Q-values are gained which also correlates with performed informal listening results.

7. ACKNOWLEDGMENT

8. REFERENCES

[1] C. Antweiler and M. Antweiler, "System Identification with Perfect Sequences Based on the NLMS Algorithm," *International Journal of Electronics and Communications (AEÜ)*, vol. 3, pp. 129–134, 1995.

[2] C. Antweiler, "Multi-Channel System Identification with Perfect Sequences," in *Advances in Digital Speech Transmission (Martin, R., Heute, U. und Antweiler, C., eds.)*, Chichester, UK: John Wiley & Sons, Ltd., Jan 2008, pp. 171–198.

[3] C. Antweiler and H.-G. Symanzik, "Simulation of Time Variant Room Impulse Responses," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Detroit, USA, May 1995, pp. 3031–3034.

[4] C. Antweiler, A. Telle, P. Vary, and E. Di Martino, "Virtual Time-Variant Model of the Eustachian Tube," in *Proc. of the Int. Symposium on Circuits and Systems (ISCAS)*, Island of Kos, Greece, May 2006, pp. 5559–5562.

[5] G. Enzner, "Analysis and Optimal Control of LMS-Type Adaptive Filtering for Continuous-Azimuth Acquisition of Head Related Impulse Responses," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, USA, March 2008, pp. 393–396.

[6] C. Antweiler and G. Enzner, "Perfect Sequence LMS for Rapid Acquisition of Continuous-Azimuth Head Related Impulse Responses," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, October 2009, pp. 281–284.

[7] S. Müller and P. Massarani, "Transfer-Function Measurement with Sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, June 2001.

[8] P. Majdak, P. Balazs, and B. Laback, "Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions," *J. Audio Eng. Soc.*, vol. 55, no. 7/8, pp. 623–637, July/Aug. 2007.

[9] A. Novák, L. Simon, F. Kadlec, and P. Lotton, "Nonlinear System Identification Using Exponential Swept-Sine Signal," *Instrumentation and Measurement, IEEE Transactions on*, vol. 59, no. 8, pp. 2220 –2229, Aug. 2010.

[10] V.P. Ipatov, "Ternary Sequences with Ideal Periodic Autocorrelation Properties," *Radio Engineering and Electronic Physics*, vol. 24, pp. 75–79, 1979.

[11] H. D. Lüke, "Sequences and Arrays with Perfect Periodic Correlation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 3, pp. 287–294, May 1988.

[12] D. Jungnickel and A. Pott, "Perfect and Almost Perfect Sequences," *Discrete Applied Mathematics*, vol. 95, no. 1-3, pp. 331–359, July 1999.

[13] A. Telle, C. Antweiler, and P. Vary, "Der perfekte Sweep - ein neues Anregungssignal zur adaptiven Systemidentifikation zeitvarianter akustischer Systeme," in *Proceedings of German Annual Conference on Acoustics (DAGA)*, Berlin, Germany, Mar. 2010, DEGA, pp. 341–342.

[14] Y. Suzuki, F. Asano, H.-Y. Kim, and T. Sone, "An Optimum Computer-Generated Pulse Signal Suitable for the Measurement of Very Long Impulse Responses," *Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1119–1123, Feb. 1995.