

# PERFECT SEQUENCE LMS FOR RAPID ACQUISITION OF CONTINUOUS-AZIMUTH HEAD RELATED IMPULSE RESPONSES

Christiane Antweiler

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

GeraldENZner

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

## ABSTRACT

In recent publications, continuous-azimuth inference of *head related impulse responses* (HRIRs) was treated as a time-varying system identification problem on the basis of dynamical measurements. The system identification thus can be handled by LMS-type adaptive filters for which we have the freedom to choose the excitation signal in this application. In order to provide the perspective of reducing the measurement time to a minimum, we now suggest the *optimal* excitation signal in terms of the rate of convergence. This excitation signal is given by *perfect sequences* (PSEQs) out of the larger family of periodic pseudo-noise signals. After the discussion of specific implications of perfect sequences, we compare the performances of our perfect-sequence LMS algorithm (PSEQ-LMS) to the results of white noise processing. We demonstrate a uniform improvement by PSEQ-LMS in terms of instrumental mean-square error analysis as well as subjective listening to dynamic HRIRs. Both measures turn out to be consistent.

**Index Terms**— Adaptive filters, head related impulse responses, perfect sequences

## 1. INTRODUCTION

*Head related impulse responses* (HRIRs) are the key tool of binaural signal processing, e.g., in applications such as virtual acoustics or advanced teleconferencing. Just recently it was discovered that the acquisition of HRIRs for all azimuthal directions can be addressed by rotating human subjects or dummy heads during the measurement [1, 2]. In [1] the tracking of time-variant HRIRs from one time-instant to the next has been proposed with LMS-type adaptive filtering and a white noise stimulus. The validity of the HRIRs rests upon the assumption that the HRIR changes are slow in comparison to the time available for their identification. In this context, a rapid identification process is therefore of special interest.

In [3] we introduced the PSEQ-LMS as a rapid tracking algorithm. This algorithm relies on LMS-type filtering with *perfect sequences* (PSEQs), e.g., [4]. PSEQs are periodically repeated pseudo-noise signals. With their special correlation properties PSEQs represent the *optimal* excitation signal for the LMS algorithm in terms of the rate of convergence. Owing to its simplicity, the PSEQ-LMS measurement has been used in several applications, e.g., for the simulation of time-variant room impulse responses in acoustic echo cancellation [5]. The rapid tracking ability of the PSEQ-LMS in combination with the demand of high quality HRIRs makes the approach so attractive for HRIR acquisition, too.

The basic measurement setup and the corresponding tracking algorithm are introduced in Sec. 2. Subsequently, in Sec. 3 we discuss the principal characteristics of the PSEQ-LMS algorithm. The benefit of using *perfect sequences* for HRIR acquisition will be demonstrated with experimental results in terms of an instrumental quality measure and subjective listening tests in Sec. 4.

## 2. IDENTIFICATION OF TIME-VARYING SYSTEMS

We first introduce the mechanical measurement setup and then present the algorithmic principals.

### 2.1. Measurement Setup

The measurements were performed with a dummy head and torso placed in the middle of an anechoic chamber facing a loudspeaker in 1.5 m distance (Fig. 1). The excitation signal  $x(k)$ , either white noise or PSEQs, is emitted via loudspeaker at a sampling rate of  $f_s = 1/T_s = 44.1$  kHz. The reaction of the system including anechoic chamber, outer ear, and torso is recorded with microphones located at the two ear canal entrances of the continuously rotating dummy head, leading to the binaural signals  $y_1(k)$  and  $y_2(k)$ . As suggested in [1], we chose  $T_{360} = 20$  s for the revolution time.

Different physical sources of error can influence the quality of the recorded signals  $y_1(k)$  and  $y_2(k)$  and thus the quality of the corresponding HRIRs. The hardware components of the measurement setup, the A/D- and D/A-converter, the loudspeaker, and the microphones add to the overall noise and distortion. A certain amount of background noise caused by the turntable engine exists, too, but has been kept negligibly low. In our system model in Sec. 2.2, the observation noise at the microphones is represented by independent additive noise signals  $n_i(k)$ ,  $i \in \{1, 2\}$ .

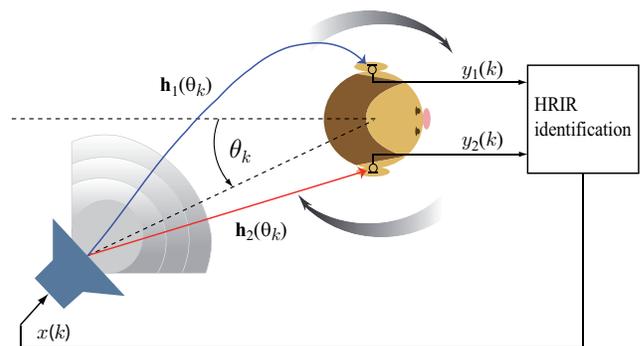


Figure 1: Measurement setup for the acquisition of HRIRs.

Besides  $n_i(k)$ , we observe a second principal source of error. The actual physical HRIRs can only be represented with an impulse response of finite length  $N$ . Depending on the level of  $n_i(k)$  and length  $N$ , the truncation of the last samples of the HRIR (“tail”) leads to a non-negligible error. This error is modeled by a noise signal  $n_{i,\text{tail}}(k)$ . In Sec. 3.2 we will show that especially this noise limits the performance of white noise LMS, while PSEQ-LMS is less sensitive.

## 2.2. Adaptive Filtering

The acquisition of HRIRs for all azimuthal directions is based on a linear time-variant system model [1]. The recorded signals  $y_i(k)$ ,  $i \in \{1, 2\}$ , at discrete time  $k$  can be expressed as

$$y_i(k) = \mathbf{h}_i^T(\theta_k) \mathbf{x}(k) + n_{i,\text{tail}}(k) + n_i(k) \quad (1)$$

with the excitation vector

$$\mathbf{x}(k) = (x(k), x(k-1), \dots, x(k-N+1))^T \quad (2)$$

and the vector representation of time-varying HRIRs,

$$\mathbf{h}_i(\theta_k) = (h_i(0, \theta_k), \dots, h_i(\kappa, \theta_k), \dots, h_i(N-1, \theta_k))^T, \quad (3)$$

where  $\theta_k = \omega k T_s$  (with  $\omega = 2\pi/T_{360}$ ) denotes the azimuth of the continuously rotating dummy head (cf. Fig. 1). According to Sec. 2.1, the signals  $n_i(k)$  and  $n_{i,\text{tail}}(k)$  represent independent noise and the undermodeled HRIR “tail”, respectively.

The system identification approach 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.,

$$\hat{\mathbf{h}}_i(\theta_{k+1}) = \hat{\mathbf{h}}_i(\theta_k) + \mu_0 \frac{e_i(k) \mathbf{x}(k)}{\|\mathbf{x}(k)\|^2}, \quad (4)$$

and of a *filtering process* generating an estimation error between the recorded response and the adaptive filter output, i.e.,

$$e_i(k) = y_i(k) - \hat{\mathbf{h}}_i^T(\theta_k) \mathbf{x}(k). \quad (5)$$

The aim of the identification process is to achieve the best possible match between the adaptive filter with impulse response  $\hat{\mathbf{h}}_i(\theta_k)$  and the HRIR represented by  $\mathbf{h}_i(\theta_k)$ . Three main factors determine the tracking performance of the NLMS algorithm: the stepsize  $\mu_0$ , the filter length  $N$ , and the correlation properties of the excitation signal  $x(k)$ :

- With a stepsize of  $0 < \mu_0 < 1$ , the NLMS algorithm performs inherently an averaging process, which provides more robustness in the presence of noise  $n_i(k)$ , but also reduces the convergence speed.
- The choice of the filter length  $N$  is always a trade off between convergence speed and the resulting error  $n_{i,\text{tail}}(k)$  and thus has to be chosen properly.
- Correlation properties of the excitation signal represent another major factor to be discussed separately in Sec. 3.

## 3. PSEQ-LMS FILTERING

From many applications it is well understood that a colored input signal  $x(k)$  severely reduces the convergence speed of the NLMS algorithm. In [3] we have shown that a special class of pseudo-noise sequences, so called *perfect sequences* (PSEQs) [4], represent the *optimal* excitation signals of the NLMS algorithm with respect to its convergence rate. PSEQs are discrete-time,

binary, ternary, or polyphase sequences  $p(k)$  of length  $N_P$ . The distinctive attribute is that they show a periodic impulse autocorrelation function according to

$$\varphi_{pp}(\lambda) = \sum_{\nu=0}^{N_P-1} p(\nu) p(\lambda + \nu) = \begin{cases} \|\mathbf{p}(\lambda)\|^2 & \lambda \bmod N_P = 0 \\ 0 & \text{otherwise.} \end{cases}$$

This special correlation property is the key to the rapid convergence of the NLMS algorithm.

In the PSEQ-LMS algorithm we periodically apply the PSEQ of length  $N_P$  to the system. As proven in [3], the period  $N_P$  has to match the length  $N$  of the adaptive filter  $\hat{\mathbf{h}}(\theta_k)$ . However, as PSEQs are available for a sufficient variety of lengths, this requirement represents no major limitation. Applying a PSEQ periodically as input signal  $x(k) = p(k)$  allows to identify a linear, time-invariant system within  $N$  iterations ( $\alpha = 1$ ,  $N = N_P$ ).

### 3.1. Rate of Convergence in Background Noise

Fig. 2 illustrates the results of a continuous NLMS adaptation process in case of a white noise and a PSEQ excitation in terms of the normalized *system distance*

$$D(k) = \|\mathbf{h}_i(\theta_k) - \hat{\mathbf{h}}_i(\theta_k)\|^2 / \|\mathbf{h}_i(\theta_k)\|^2 \quad (6)$$

in case of an effective signal-to-noise ratio

$$\text{SNR} = \frac{\text{E}\{(\mathbf{h}_i^T(\theta_k) \mathbf{x}(k))^2\}}{\text{E}\{(n_{i,\text{tail}}(k) + n_i(k))^2\}} \quad (7)$$

of 30 dB. The rapid identification process and the effect of the time-constant  $N$  can be observed for the PSEQ excitation. The direct comparison with the system distance achieved with white noise emphasizes that the NLMS benefits from the special correlation properties of the deterministic PSEQ. As the quality of the measured impulse responses  $\hat{\mathbf{h}}(\theta_k)$  depends on the ratio of physical HRIR variability to the rate of adaptation, the results of Fig. 2 clearly motivate the use of PSEQs. The effect of the different convergence rates will be further investigated in Sec. 4.

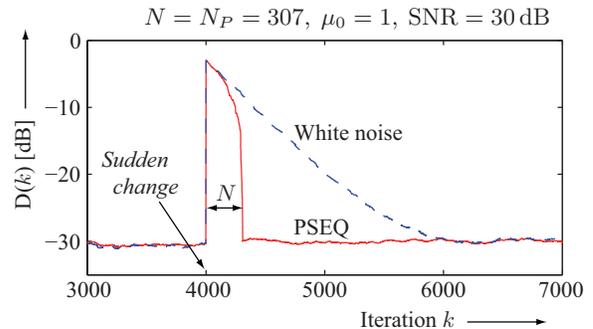


Figure 2: System distance for PSEQ and white noise excitation,  $\mathbf{h}_i$  time-invariant, except for sudden change at  $k = 4000$ .

### 3.2. Systematic vs. Non-Systematic Tail Error

The error signals  $n_i(k)$  and  $n_{i,\text{tail}}(k)$  influence the results of the identification process. For both excitation signals (white noise and PSEQ),  $n_i(k)$  causes the same disturbing effect on the adaptation process, however, the influence of  $n_{i,\text{tail}}(k)$  is different.

To visualize the principal differences, we set  $n_i(k) = 0$  and investigate the influence of  $n_{i,\text{tail}}(k)$  in one simulation example

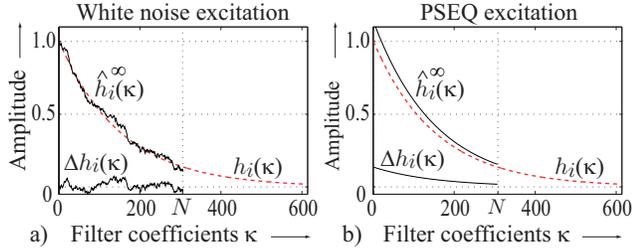


Figure 3: Distortions due to mismatched filter lengths

$\Delta h_i(\kappa) = \hat{h}_i^\infty(\kappa) - h_i(\kappa)$ ,  $\kappa = 0 \dots N-1$ ;  $\mu_0 = 1$ ;  $n_i(k) = 0$   
 $h_i(\kappa)$ : target impulse response of length 614, time-invariant  
 $\hat{h}_i^\infty(\kappa)$ : identified impulse response with  $N = 307$ , steady-state

(Fig. 3). For the “unknown” impulse response  $h_i(\kappa)$  we chose a special test function in the form of an exponentially decaying curve which is of length  $2N = 614$  and time-invariant, i.e., the angle  $\theta_k$  is irrelevant. For the NLMS algorithm a shorter adaptive filter of length  $N = 307$  is used to cause a certain noise level for exemplification. Figure 3 depicts the target impulse response  $h_i(\kappa)$ , the steady-state impulse response  $\hat{h}_i^\infty(\kappa)$  obtained after sufficiently long adaptation, and the difference  $\Delta h_i(\kappa)$ .

The NLMS algorithm principally minimizes the distance between  $y_i(k)$  and the adaptive filter output  $\hat{y}_i(k)$ , even in case of an undermodeled adaptive filter ( $N = 307 < 614$ ). Considering

$$n_{i,\text{tail}}(k) = \sum_{\kappa=N}^{2N-1} h_i(\kappa) x(k-\kappa) \quad (8)$$

in Eq. (1) and aiming at  $\hat{y}_i(k) = y_i(k)$ , we expand both sides

$$\sum_{\kappa=0}^{N-1} \hat{h}_i^\infty(\kappa) x(k-\kappa) = \sum_{\kappa=0}^{2N-1} h_i(\kappa) x(k-\kappa) \quad (9)$$

and rewrite (9) according to

$$\begin{aligned} & \sum_{\kappa=0}^{N-1} [h_i(\kappa) + \Delta h_i(\kappa)] x(k-\kappa) \\ &= \sum_{\kappa=0}^{N-1} [h_i(\kappa) x(k-\kappa) + h_i(\kappa+N) x(k-\kappa-N)] \\ \Rightarrow & \sum_{\kappa=0}^{N-1} \Delta h_i(\kappa) x(k-\kappa) = \sum_{\kappa=0}^{N-1} h_i(\kappa+N) x(k-\kappa-N) \\ \Rightarrow & \sum_{\kappa=0}^{N-1} [\Delta h_i(\kappa)] x(k-\kappa) \\ &= \sum_{\kappa=0}^{N-1} \left[ h_i(\kappa+N) \frac{x(k-\kappa-N)}{x(k-\kappa)} \right] x(k-\kappa). \quad (10) \end{aligned}$$

From the comparison of both sides of (10), it follows a solution for all  $\kappa = 0, \dots, N-1$  with

$$\Delta h_i(\kappa) = h_i(\kappa+N) \cdot x(k-\kappa-N)/x(k-\kappa). \quad (11)$$

In case of white noise excitation  $x(k)$ ,  $\Delta h_i(\kappa)$  is a *non-systematic* noise as, according to (11), the “tail”  $h_i(\kappa+N)$  is multiplied with a quotient of statistically independent noise samples, see also Fig. 3-a. With the choice of a smaller stepsize  $\mu_0$ , the noise can be reduced by the NLMS.

For a PSEQ excitation we get  $\Delta h_i(\kappa) = h_i(\kappa+N)$  by taking the periodicity of  $x(k)$  into account. The truncated “tail” of the target impulse response  $h_i(\kappa)$  is thus added onto the leading

part of  $\hat{h}_i(\kappa)$  (Fig. 3-b). The error  $\Delta h_i(\kappa)$  is a *systematic* error and cannot be reduced by an averaging process ( $\mu_0 < 1$ ). Note, however, that for both excitation signals the power of the resulting error  $E\{\Delta h_i^2(\kappa)\}$  is identical.

For the sake of simplicity in the example of Fig. 3 the target impulse response  $h_i(\kappa)$  was only twice as long as the adaptive filter  $\hat{h}_i(\kappa)$ . Longer target impulse responses will cause multiple superpositions of “tail”-sections.

In the following section we will investigate how the nature of the *systematic* and the *non-systematic* error influences the quality of the HRIR acquisition.

## 4. RESULTS

The results are based on recordings according to Sec. 2.1. and on the identification of time-varying HRIRs according to Sec. 2.2. We will first discuss the subjective quality of the HRIRs and then define an instrumental measure which proves to be nicely correlated with the subjective quality.

### 4.1. Subjective Listening Tests

In our experiments, “dry” speech signals were convolved with rotating HRIRs in order to give the impression of a virtually rotating speaker.

In case of white noise excitation, we observed that the stepsize  $\mu_0$  of the NLMS adaptive algorithm needs to be reduced below unity, e.g.,  $0.25 < \mu_0 < 0.5$ , in order to reject unsystematic noise  $n_{i,\text{tail}}(k)$  and related sound artefacts of the audio signal. If the NLMS algorithm is operated inappropriately with  $\mu_0 = 1$ , sound artefacts can be perceived typically at the shadow-side of the auditory circle (with low direct sound component). With  $\mu_0$  below unity, however, the tracking ability of the NLMS algorithm is reduced such that the spatial resolution of the rotating HRIRs is degraded.

Using PSEQ excitation, we have demonstrated that the HRIR “tail” will not appear as unsystematic noise in the system identification process. We can rise the stepsize of the NLMS algorithm to  $0.5 < \mu_0 < 1$  while not encountering significant sound artefacts in the audio signal. Furthermore, the large stepsize is chosen in favor of a small time-constant of the NLMS algorithm, which will better preserve the spatial resolution of the rotating sound source. The systematic HRIR-aliasing related to PSEQ excitation (see Sec. 3.2.) was not observed as an auditory degradation in the virtually rotating sound source.

### 4.2. Instrumental Measure of HRIR Quality

For an objective comparison of the different excitation signals an instrumental measure of HRIR quality is of interest. The aim of the identification process is to achieve the best possible estimation of the rotating unknown HRIR. A good match between the adaptive filter and the HRIR, i.e.,  $\hat{\mathbf{h}}_i(\theta_k) \approx \mathbf{h}_i(\theta_k)$ , will also cause a smaller estimation error  $e_i(k)$ . Thus, the power of the estimation error  $E\{e_i^2(k)\}$  can be used as an indicator for the quality of the adaptation process. However, it should be noted that a small value of  $E\{e_i^2(k)\}$  does not automatically correspond to a small system distance between  $\hat{\mathbf{h}}_i(\theta_k)$  and  $\mathbf{h}_i(\theta_k)$ . As quality index we define

$$Q_{\text{HRIR}}(\theta_k) = E\{y_i^2(k)\}/E\{e_i^2(k)\}. \quad (12)$$

The quality measure mainly relies on the estimation error  $e(k)$ . In case of PSEQ excitation, the “tail” of the truncated impulse response is superimposed onto the leading part of the estimated

HRIR (see Fig. 3). Due to the periodic nature of the stimulus, the HRIR has the potential to achieve a smaller error signal and thus higher  $Q_{\text{HRIR}}(\theta_k)$  as compared to the white noise case.

Obviously, the quality measure “hides” the *systematic* error, but reveals the *non-systematic* error. These properties of  $Q_{\text{HRIR}}(\theta_k)$  are consistent with the subjective impression in Sec. 4.1 and  $Q_{\text{HRIR}}(\theta_k)$  thus turns out to be particularly suitable as an instrumental measure for dynamical HRIR quality.

### 4.3. Instrumental Comparison

Figure 4 illustrates the quality measure  $Q_{\text{HRIR}}(\theta_k)$  in dB for four cases: white noise and PSEQ excitation each with filter length  $N = 307$  and  $614$ . We will focus only on the evaluation of the left ear because of the symmetry of the measurement setup.

In the curve referring to ‘WN,  $N = 307$ ’ we notice a significant decay of  $Q_{\text{HRIR}}(\theta_k)$ , when the loudspeaker is located at  $\theta_k = 90^\circ$ . This decay reflects the auditory impression of the listening tests and is caused by noise  $n_{i,\text{tail}}(k)$  due to the limited adaptive filter length  $N$ . Naturally,  $n_{i,\text{tail}}(k)$  occurs also for all other angles, however, at  $\theta_k = 90^\circ$ , the direct sound is missing. Thus, the influence of  $n_{i,\text{tail}}(k)$  increases significantly. As this noise signal is *non-systematic* in each iteration, a *statistically independent* impulse response error  $\Delta h_i(\kappa)$  interferes the HRIR according to Fig. 3-a in each angle. The convolution of an audio signal with consecutive HRIRs might result in a noise-like audible disturbance at  $90^\circ$ .

The curve referring to ‘PSEQ,  $N = 307$ ’ shows quite stable results in the order of  $Q_{\text{HRIR}}(\theta_k) = 23$  dB for all azimuths. It varies slightly during the rotation, but  $Q_{\text{HRIR}}(\theta_k)$  does not severely degrade at  $90^\circ$  as in case of the ‘WN,  $N = 307$ ’. As said before, this behavior of the quality measure resembles the results of subjective listening tests. The auditory quality benefits from the fact that the *systematic* error  $\Delta h_i(\kappa)$  does not fluctuate from one iteration to the next as much as in the white noise case.

If we enlarge the filter length, i.e., ‘WN,  $N = 614$ ’ and ‘PSEQ,  $N = 614$ ’, the noise  $n_{i,\text{tail}}(k)$  can be reduced to a certain degree, however, the convergence speed of the NLMS slows down in both cases. Lower tracking ability affects the results obtained around the front ( $180^\circ$ ) and rear ( $0^\circ/360^\circ$ ) loudspeaker position. For these angles,  $Q_{\text{HRIR}}(\theta_k)$  decreases significantly,

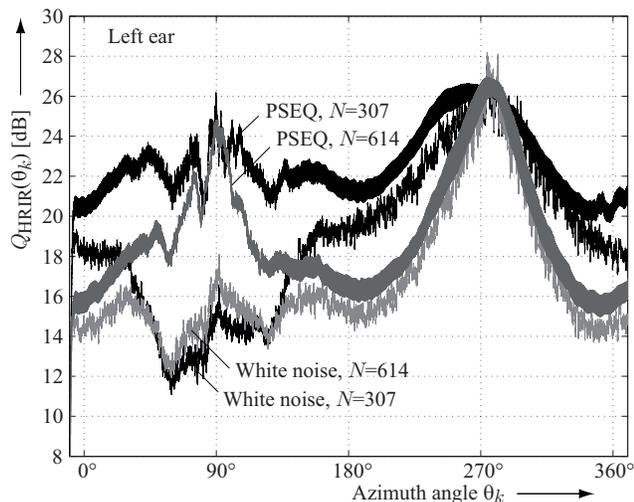


Figure 4: Quality measure for white noise and PSEQ,  $\mu_0 = 0.5$ .

while the results for the other angles remain somewhat unaffected. This means that rotations in the front and the back of the head cause more HRIR changes than in the lateral directions. This trend is independent of the stimulus and other parameters of the system.

At  $90^\circ$ , where the presence of  $n_{i,\text{tail}}(k)$  limits the performance of the identification process, an increased filter length generally reduces the *non-systematic* error in case of white noise. Thus, for ‘WN,  $N = 614$ ’ the measure  $Q_{\text{HRIR}}(\theta_k)$  increases slightly. The quality resulting from the PSEQ-LMS algorithm, however, can not further improve as  $Q_{\text{HRIR}}(\theta_k)$  anyhow “hides” the *systematic* error (Sec. 4.2).

At  $270^\circ$ , the direct sound component dominates the simulation results, i.e.,  $n_{i,\text{tail}}(k)$  is of minor influence and, thus,  $Q_{\text{HRIR}}(\theta_k)$  is similar in all cases.

## 5. CONCLUSIONS

An adaptive filtering method for dynamical HRIR acquisition at any azimuth has been investigated and further improved. The key of the suggested approach is the use of *perfect sequences* (PSEQs). The PSEQ-LMS algorithm performs a rapid identification within  $N$  iterations. Furthermore, the PSEQ-LMS approach benefits from the occurrence of a *systematic* error with less disturbing fluctuations from one time-instant to the next. In applications with virtually moving sound sources, noise-like audible distortions can be avoided while maintaining convergence speed. Due to its high convergence rate and its simplicity, the PSEQ-LMS algorithm is superior, in particular, for rapid individualized acquisition of HRIRs.

Finally, we showed that the suggested instrumental quality measure correlates with the performed listening results and in that way represents an appropriate measure for the quality of continuous-azimuth HRIRs.

## 6. ACKNOWLEDGMENT

The authors would like to thank Michael Weinert and Christoph Höller, who performed numerous acoustics measurements and computer simulations within their student research projects.

## 7. REFERENCES

- [1] 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, USA, March 2008, pp. 393–396.
- [2] T. Ajdler, L. Sbaiz, and M. Vetterli, “Dynamic measurement of room impulse responses using a moving microphone,” *Journal of the Acoustical Society of America*, vol. 122, no. 3, pp. 1636–1645, September 2007.
- [3] 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.
- [4] V. Ipatov, “Ternary Sequences with Ideal Periodic Autocorrelation Properties,” *Radio Engineering Electronics and Physics*, vol. 24, pp. 75–79, 1979.
- [5] 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.