THE COMPARISON OF ROOM IMPULSE RESPONSE MEASURING SYSTEMS

Ivo MATELJAN*, Kosta UGRINOVIĆ**

*Faculty of Electrical Engineering, University of Split,
R. Boskovica bb, 21000 Split, Croatia
**Faculty of Natural Sciences, Mathematics and Education, University of Split,
N.Tesle 12, 21000 Split, Croatia

ABSTRACT

This paper describes cross-correlation based systems for the estimation of the room impulse response. A various types of excitation signals are used: white noise, pink noise, swept-sine, MLS, and random phase multisine. Criteria to evaluate the quality of the impulse response estimation were: a) random and stationary noise rejection, b) rejection of nonlinearities introduced by measuring system, c) immunity to time-variant environmental disturbances and d) averaging time necessary to achieve reliable estimation.

The work shows that choice of an optimal excitation signal depends on measurement environment. In an environment with high noise, a random phase multisine gave the best estimation, otherwise excellent results were obtained with a log-sweep excitation. The robustness of the estimation can only be achieved with a random phase signals.

1. INTRODUCTION

The measurement of the impulse response is widely accepted as a basis for the estimation of acoustical parameters and characteristics of rooms and sound systems (frequency response, cumulative spectral decay, ETC, energy decay and reverberation characteristics, IACC, speech transmission index).

The purpose of this work is to compare the following cross-correlation based methods for the impulse response estimation: Fourier analyzer [2], MLS-based system [3] and a swept-sine system [4]. The main differences between these measuring systems are:

1. Excitation signals are periodic (MLS, random phase multisine) or transient (swept-sine, random noise). System that uses periodic excitation gives estimation of impulse response that is usually called periodic impulse response (PIR).
2. The signal acquisition system is implemented as a single channel or dual channel system. In dual channel systems, input and output of the system are measured. In

single channel systems, only the output is measured, while input is digitally generated discrete signal that is located in a computer memory.

3. The signal processing (correlation and averaging) is done in frequency domain (Fourier analyzer, swept-sine) or in time domain (MLS, swept-sine).

Other differences, such as the quality of A/D and D/A conversion, the type of antialiasing filters, use of dihtering and numerical accuracy are not characteristics of the measuring method. That is why, to make the objective comparison of measuring methods, all methods are implemented in the same measuring system – ARTA. It uses high quality PC soundcards (Terratec EWX 24/96) and software that gives user friendly access to setup of many acoustical measurements (see Fig. 1).

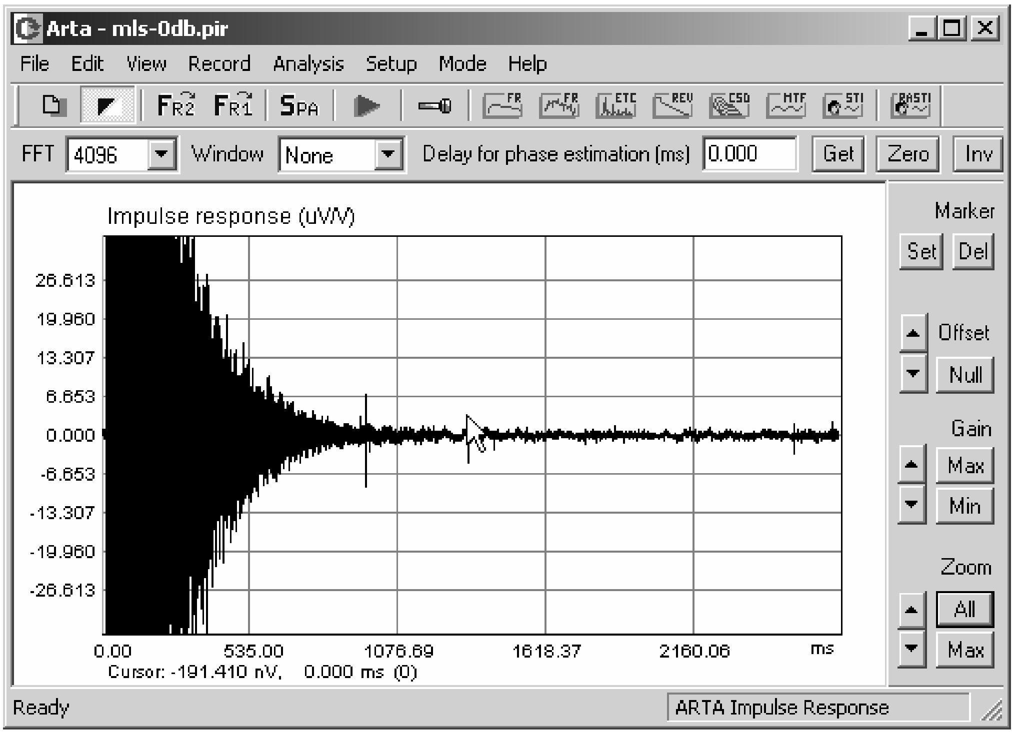

Figure 1. ARTA window for impulse response analysis (MLS-estimated PIR is shown)

Environmental and systematic errors are present in the estimated impulse response. The noise and time changes of the propagation medium are main sources of environmental errors, while the source of systematic errors is the measuring system (loudspeaker and microphone nonlinearity, dynamic range and bandwidth of the excitation signal, time-bandwidth resolution, the type of averaging and implementation dependant post processing algorithms).

This paper first presents measuring techniques to minimize these errors. Then, results of measurements in a real acoustical environment are presented to show:

- distribution of noise across the estimated impulse response,
- distribution of nonlinear distortions across the estimated impulse response, and
- reliability of the estimation in a time-variant environment.

The most popular system for measuring the impulse response is MLS-based system. Rife [3] postulated two properties of the MLS measuring system: energy conservation and phase randomization. The first property means that measuring S/N is preserved in the estimated PIR over all frequencies, while the second property means that the influence of a noise and distortions is uniformly distributed across the estimated PIR. The main problem with MLS-based systems is that they do not uniformly distribute errors due to nonlinear distortions across the estimated PIR, instead there is a transformation of distortions in a series of PIR spikes, which are impossible to distinguish from true reflections that can be present in the estimated impulse response (see Fig. 1). It is shown in [1] that better estimation can be obtained with low-crest excitation signals that belong to the class of random phase multisine signals (RPMS) [5]. A Fourier analyzer that uses the RPMS with white and pink spectrum is presented in [2]. In this work, the analyzer is extended with a swept-sine based impulse response estimator. The signal generation is based on work of Farina [4], who experimentally showed the distortion rejection property of logarithmic sweep.

2. PROPERTIES OF EXCITATION SIGNALS

Random Phase Multisine

The random phase multisine (RPMS) is a periodic signal whose spectrum has deterministic amplitude and random phase [5]. In ARTA, it is generated as wideband signal by using the inverse Fourier transform:

$$g[n] = \frac{1}{N}\sum_{k=0}^{N-1} A_k^{'} e^{j\varphi_k} e^{j2\pi kn/N} \ , \qquad (1)$$

where $A_k^{'} = \begin{cases} 0, & k = 0 \\ A_k, & k > 0 \end{cases}, \quad \varphi_{N-k} = \varphi_k \mid \text{random} \in [0, 2\pi], \ \varphi_{N/2} = 0$

If A_k = const., $g[n]$ is periodic white RPMS. In acoustical measurements the system noise is colored, and better choice to use the pink RPMS ($A_k = \sqrt{2} A_{2k}$). In ARTA, generation of pink RPMS sequences is implemented with a variable low frequency roll-off.

Amplitudes of RPMS have a normal distribution (proof in [1]), and crest factor is 12-13 dB. By further processing the crest factor can be lowered to values below 6dB [1]. In ARTA, the RPMS generation is implemented with a crest factor lower than 10dB.

The circular autocorrelation of RPMS (suppose A=1) is

$$AutoCorr[n] = \mathbf{IDFT}\left[GG^*\right] = \delta[n] - \frac{1}{N} \qquad (2)$$

where $\delta[n]$ is a periodic unit-sample sequence, with period N. Except the small dc offset (which is negligible for the large N) and periodicity implied by the circular correlation, this is the characteristics of the white process.

Maximum length sequence

The MLS signal is a binary sequence of length $N=2^k-1$. It has circular correlation defined with Eq. (2), with substitution of $N+1$ for N, so it has white spectrum. Only a limited number of MLS sequences can be generated [3]. Theoretically, the MLS sequence has crest factor 0dB, but in real implementations crest factor varies from 6 to 9 dB, due to the antialiasing filtering.

Swept-sine

The linear sweep has a white spectrum. It is defined by equations:

$$g(t) = \sin(2\pi\, f(t)), \qquad f(t) = f_1 t + (f_2 - f_1)t^2 / 2T \tag{3}$$

where T is the total sweep duration, f_1 is the start frequency, and f_2 is the stop frequency.

The logarithmic sweep has a pink spectrum. It is defined by the equation:

$$f(t) = \frac{f_1 T}{\ln(f_2 / f_1)} (e^{\frac{t}{T}\ln(f_2/f_1)} - 1)) \tag{4}$$

The swept-sine has crest factor 3dB. For the impulse response estimation, the sweep must be generated as a wideband signal with enough power spectral density level at all frequencies of the sampled system. In ARTA, the Hann window is applied to the beginning part and at the end part of the sweep. It smoothes the spectrum and gives the proper amount of spectrum level at all frequency bins.

3. THE MEASURING SYSTEM

Fig. 2 shows the measuring system. The computer generated signal g, after D/A filtering with transfer function D, is applied to the test system that has transfer function H. Note that H represent best linear fit of the possible nonlinear transfer function. The generator noise is n_g. The output from the test device, together with an additive system noise n_s, is acquired by the computer as a discrete signal sequence y. The acquisition process implies the use of an antialiasing filter that has transfer function A. In a dual channel system input to the test device is acquired by the computer, together with input channel noise n_x, as a discrete signal sequence x. In acoustical measurements, the influence of generator noise and noise of input channel can be neglected, as they are much smaller than system noise and distortions.

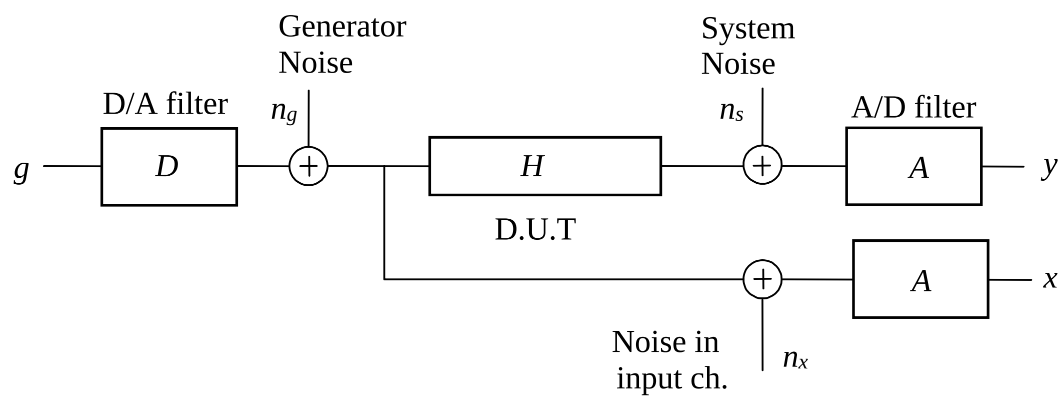

Figure 2. Block diagram of the measuring system

Dual channel system (Fourier analyzer)

In a dual channel system, we first estimate the frequency response function (FRF), and then we get the impulse response estimation by applying the inverse DFT to the FRF. In a classical Fourier analyzer the excitation is a random noise and a frequency response is estimated by dividing the averaged cross-spectrum X^*Y with averaged auto-spectrum X^*X of N input and output discrete signal sequences. We define the H_1 estimator as:

$$H_e(\omega) = \frac{\sum_{i=1}^{N} Y_i(\omega) X_i^*(\omega)}{\sum_{i=1}^{N} X_i(\omega) X_i^*(\omega)} \qquad (H_1 \; estimator) \tag{5}$$

where $H_e(\omega)$ denotes the estimated FRF and star (*) denotes the complex conjugate value. The effect of averaging can be expressed by the equation:

$$H_e(\omega) \cong H(\omega) + \frac{\sqrt{N}\left\langle N_s(\omega) A(\omega) X^*(\omega)\right\rangle}{N\left\langle X^*(\omega) X(\omega)\right\rangle} \cong H(\omega) + \frac{1}{\sqrt{N}} \frac{\left\langle N_s(\omega) G^*(\omega)\right\rangle}{\left\langle G(\omega) G^*(\omega)\right\rangle} \frac{D^*(\omega)}{\left|D(\omega)\right|^2} \tag{6}$$

where brackets <> denote the averaged value. Note that signal term is summed coherently, while the stochastic part of the noise is power summed. The conclusion is that averaging lowers the noise level proportionally with a square root of number of averages, thus improving the measurement S/N by 10log(N). If nonlinear distortions are present, then part of the system noise is coherent with a generated signal, and a better measure for the proportionality of the noise+distortion and a number of averages is $1/\gamma\sqrt{N}$, where γ is the input-output coherence function [6].

If the excitation is done with N different RPMS sequences, the estimator can be of the form:

$$H_e(\omega) = \frac{1}{N} \sum_{i=1}^{N} \frac{Y_i(\omega) X_i^*(\omega)}{X_i(\omega) X_i^*(\omega)} \tag{7}$$

This type of averaging is called the frequency domain asynchronous averaging. An important theoretical result [5] shows that this estimator has the same power in reduction of noise and distortion as H_1 estimator.

If the excitation is done with a single periodic sequence, repeated N times, the estimator can be of the form:

$$\bar{y}(t) = \sum_{i=1}^{N} y_i(t), \quad \bar{x}(t) = \sum_{i=1}^{N} x_i(t), \quad H_e(\omega) = \frac{\bar{Y}(\omega) \bar{X}^*(\omega)}{\bar{X}(\omega) \bar{X}^*(\omega)} \tag{8}$$

This type of averaging is called the time domain synchronous averaging. This estimator reduces the system random noise, but it cannot reduce nonlinear distortions and a system stationary noise that is periodic within the excitation period.

The quality of the frequency response estimation is noise dependent as follows:

1. The noise influence is high at frequencies near $f_s/2$, where we miss the excitation energy which is filtered with D/A smoothing filter (see Eq. 6). The problem is solved in program ARTA by digital filtering the estimated impulse response at frequency lower than converter's antialiasing filter cut-off frequency.
2. The noise influence is also high at extremely low frequencies, as system is AC coupled. This implies that soundcards with very low cut-off frequency have to be used.

3. When using a continuous noise generator, $X(\omega)$ does not have the constant spectrum at all frequency bins, as is the case for the RPMS excitation. This gives the frequency selective noise bias. It is high at frequencies where generator spectrum has notches. This resolution bias can be greatly reduced by increasing the number of averaging cycles. Results given in [2] suggests that it is necessary to average at least 8 spectrum averages, to get the result comparable with those obtained with a single period RPMS excitation.

The main problem of classical Fourier analyzer with continuous noise excitation is the large bias error when we measure frequency response in multipath environment. Signals that are delayed more than acquisition aperture time are always uncorrelated with excitation signal. This uncorrelated bias can be avoided by using the RPMS excitation under following conditions:

(i) random phase multisine signal must be periodic in the acquisition window,
(ii) start of the acquisition must be after a preaveraging cycle that is necessary to reach the steady state response,
(iii) when using frequency domain averaging, then for every generated sequence g_i (ii) must be fulfilled,
(iv) period of the RPMS must be greater than reverberation time (T_{60}) when we measure in an acoustical environment.

The following reasoning can approve the last requirement. The room acoustical response has the bandwidth of resonance peaks equal to $2.2/T_{60}$ [7]. If we choose that frequency difference between two multisine component is less than half of this value, to allow build up of all room resonances, we can conclude that period of the RPMS have to be equal or greater than reverberation time. Also, it follows that length of the preaveraging cycle must be greater or equal to the reverberation time.

Single channel system

In single channel systems, the computer generated signal $g(t)$ is used as the input signal. FRF estimators are defined by simply substituting $G(\omega)$ instead of $X(\omega)$ in Eq. (5), (6) and (8). Only estimator with a time domain synchronous averaging is applied in program ARTA, as Windows sound driver cannot control the delay between input and output channels (this control is necessary for frequency domain averaging).
The transfer function is estimated by the cross-spectrum calculation,

$$H_e(\omega) = \frac{\overline{Y}(\omega)\overline{G}^*(\omega)}{\overline{G}(\omega)\overline{G}^*(\omega)} \cong H(\omega)A(\omega)D(\omega) + \frac{1}{\sqrt{N}}\left\langle \frac{A(\omega)N_s(\omega)G^*(\omega)}{G(\omega)G^*(\omega)} \right\rangle \tag{9}$$

and can be considered as suboptimal FRF estimation, because H_e always gives transfer function H biased with transfer function of antialiasing filters.

The single channel system is a better choice than dual channel system if we use low-cost soundcards as they usually have restricted low frequency response and unmatched sensitivity of input channels.

MLS based system

The MLS excitation can only be applied in a single channel system. MLS has white spectrum ($G(\omega)G^*(\omega)$ = const.), and impulse response can be estimated by cross-correlation calculation in time domain (convolution of system output with time reversed input sequence). In program ARTA, a fast Hadamard transform is used for that calculation. It is faster and has less numerical errors than FFT transform.

Swept-sine system

The swept-sine can be applied as an excitation signal in single channel and dual channel systems. It must be treated as a transient excitation because sweep spectral components exist only a fraction of the sweep duration. In ARTA, the excitation signal is composed of two parts. First part of duration T_1 is the sweep defined with Eq. (3) or Eq. (4). The second part, that follows the sweep, has zero value and duration $T_2 \geq T_1$. The duration of this composite signal (T_1+T_2) is chosen to be equal to the DFT analysis time window. In acoustical measurements, the time interval T_2 must be greater than reverberation time. This way the impulse response estimation can be done with synchronous averaging, without time aliasing errors.

In ARTA both linear and logarithmic sweep are implemented. The logarithmic sweep has superior characteristics as it enables clear separation of linear from nonlinear response of the system with nonlinear harmonic distortions. In log-sweep, defined by Farina in Eq. 5 [4], N^{th} harmonic components are equally separated in time by value:

$$\Delta t = T \frac{\ln(N)}{\ln(f_2 / f_1)} \tag{10}$$

After the convolution of the system response with a time reversed sweep, these "ahead in time" components appear before the linear part of the impulse response. This is demonstrated in Fig. 3. For a sweep with f_1 = 10 Hz, f_2 = 24000 Hz, T = 2.73 s, both the Eq. (10) and a measurement give for the second harmonic response Δt = 242 ms, and for third harmonic response Δt = 385 ms.

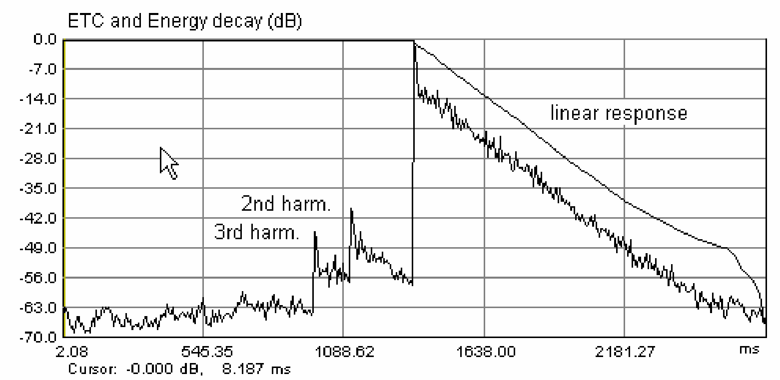

Figure 3. ETC and Schroeder Energy decay, measured with log-sweep

4. THE EXPERIMENTAL COMPARISON OF MEASURING SYSTEMS

We have made many measurements in various rooms (large and small music halls, airport terminals, churches and auditoria). Here we present only results that we judge to be important for general conclusions.

Table 1 shows rms and peak levels of impulse response tail noise measured with MLS, white RPMS, pink RPMS and log-sweep excitation signals. All signals were generated with a same power. Excitation levels were 0, -6, -12 and -18 dB ref. 81 dB(A) SPL (pink excitation) in a room with a volume 650 m^3 and T_{60} = 1.3 s. The microphone was placed 7m apart from the omnidirectional loudspeaker. The continuous noise, with speech spectrum shape, was held at constant level of 35 dB(A). The last row shows tail noise level obtained with 16 averages.

Table 1. RMS / peak level of impulse response tail noise (dB)

a) wideband tail noise level (rms/peak dB)

excitation level	MLS	White RPMS	Pink RPMS	Log-Sweep
0 dB	-70.9/-48.3	**-70/-57.43**	-66.7/-53.7	**-71.1/61.2**
-6 dB	-63.81/-47.6	-65.5/-53.7	**-68/-57**	-66.2/-58.1
-12 dB	-59.5/47.0	-59.8/-47.3	-67.2/-55	**-69/-56.5**
-18 dB	-57/-45	-55.1/-42	**-64.2/-52**	-61.8/-51.3
-18dB / 16avg	-66.7/-55.7	-62.18/-52.1	**-75/-62**	-72.5/-61.5

b) tail noise level in 250Hz octave band (rms/peak dB)

excitation level	MLS	White RPMS	Pink RPMS	Log-Sweep
0 dB	-57/-45	-58/-46	-58.6/-47.2	**-65.8/-53.7**
-6 dB	-51.9/-39.9	-51.8/-39.4	**-59.9/-48.5**	**-59.9/-48.5**
-12 dB	-46/-34	-46/-35.6	-52.61/-41	**-53.3/-42.4**
-18 dB	-41.4/-29.6	-41/-30	**-50.45/-38.7**	-49/-37
-18dB / 16avg	-51.8/-41	-52.7/-41.2	**-62.2/-50.51**	-59.2/-47.5

c) tail noise level in 1kHz octave band (rms/peak dB)

excitation level	MLS	White RPMS	Pink RPMS	Log-Sweep
0 dB	-67/-55.8	-63.8/-51.8	-59.2/-46.7	**-77/-61**
-6 dB	-58/-45	-65.2/-52.9	-61.2/-49.2	**-71/-59**
-12 dB	-54.5/-42.2	-56.2/45.2	-63.5/-51.2	**-66.2/-53.1**
-18 dB	-57.8/-45.2	-55.6/-46.5	**-62.2/-51.1**	-61/-49.2
-18dB / 16avg	-69/-56.3	-67/-57	**-72.7/-61.5**	-71.5/-60.5

These results show that under high S/N conditions the log-sweep gave best results, while at moderate and low S/N best results were achieved with a pink RPMS excitation. The averaging helps to reduce tail noise level in all cases. Bad results for MLS and pink RPMS excitation are at largest power level. MLS shows large PIR spikes, while pink RPMS pushes loudspeaker to large displacement at low frequencies, thereby producing large distortions. This is the reason why we implemented pink RPMS with variable low-frequency roll-off (20Hz-1000Hz).

Further experiments, with more uneven noise spectrum and temporal patterns (cracks, clicks, hand clapping), have revealed that swept-sine system unevenly distribute noise across the estimated impulse response (typical example shows Fig. 4), and in many cases the averaging do not increases the measurement S/N as expected. In such a cases random phase excitation gives better results.

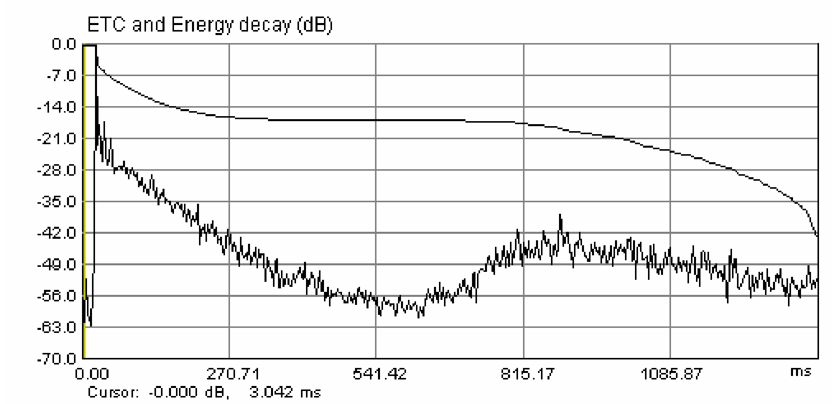

Figure 4. Typical time "colored" response of swept-sine system to impulsive noise

In a time-variant environment, the random phase excitation gives small S/N and distorted impulse response. Many experiments indicate that swept-sine system is less susceptible to time changes of environment medium, but the problem is that estimation is quite unpredictable, and that averaging usually gives worse result than a single measurement. Fig. 5 shows one case when sweep based system gave erroneous estimation. The measurement was made in a large hall with very low noise level (< 25 dB(A)), but with powered ventilation fan. A large low-frequency modulation of the impulse response is noticeable. This distortion was not present in the pink RPMS driven system.

In all measurements with a RPMS driven system the frequency domain averaging reduces the noise level, even in the time-variant environment. For example, we made measurements at the airport terminal (volume = 6000 m^3) where ventilation noise was 52 dB(A). With 16 averages tail noise level was -60 dB.

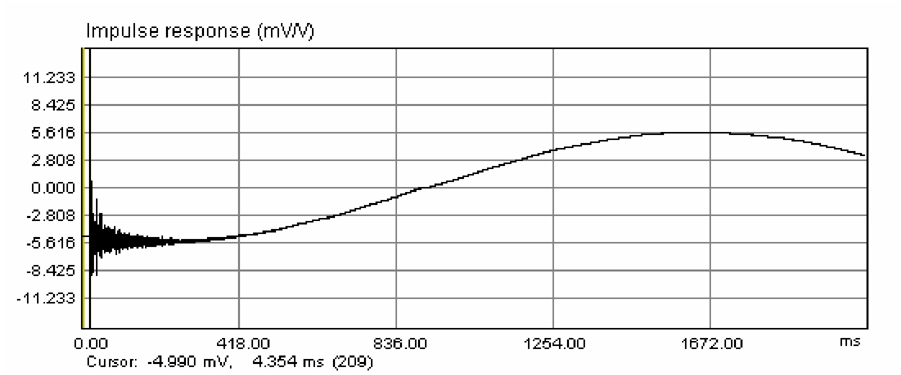

Figure 5. Impulse response of swept-sine system in room with low-speed ventilation

In a work [4] Farina concludes with a requirement for the swept-sine system: "there is no need to maintain tight synchronization between the sampling clock of the signal generator and of the digitizing unit employed for capturing the system response". We can't agree with this statement, as difference in I/O sampling frequencies in a single channel system generates time compressed sweep image at the beginning of the impulse response. Such impulse response has a low level of tail noise, and can be used for the estimation of reverberation time, but it cannot be used for correct estimation of the frequency response from the gated beginning part of the impulse response.

In a dual channel system there is no this type of distortion, as both input channels are always captured with a same sampling frequency.

5. CONCLUSIONS

Dual and single channel systems for the estimation of the impulse response are presented. Dual channel systems can only be realized with a high quality sound cards that have extended low-frequency response and matched sensitivity of input channels. Single channel systems can be realized with lower-quality sound cards.

This work shows that systems for measuring room impulse response with a pink RPMS and logarithmic sweep excitation outperform systems with white spectrum excitation in a real acoustical environment. The choice of an optimal excitation signal depends on measurement environment. In an environment with a high noise, the pink RPMS gives the best estimation, otherwise excellent result can be achieved with a logarithmic sweep excitation.

When using swept-sine system, the averaging does not always improve the measurement S/N as noise contribution is not uniformly distributed across the estimated impulse response. That is why it is better to use the long duration sweep, then to use the averaging. A high crest factor of the sweep and distortions redistribution property makes the swept-sine system ideal for a high-power loudspeaker testing in a low-noise environment.

None of tested systems is a reliable impulse response estimator in a time-variant environment, although the swept-sine system is more than other systems tolerant to the small perturbation of the propagating medium.

The most important conclusion is that robustness of the impulse response estimation can only be achieved with excitation signals that have random phase. Then, averaging always improve measurement S/N ratio. The frequency domain asynchronous averaging is the most effective type of averaging. It reduces the influence of random noise, stationary noise and nonlinear distortions.

6. REFERENCES

[1] Mateljan, I., "Signal Selection for the Room Acoustics Measurement", *Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,* New Paltz, New York, 1999.

[2] Mateljan, I., Ugrinović K. "The Fourier Analyzer for Acoustical Measurement", *Proc. ELMAR 2001*, Zadar, Croatia, 2001.

[3] Rife, D.D., Vanderkooy, J., "Transfer Function Measurement with Maximum-Length Sequences", *J. Audio Eng. Soc*., vol. 37, June 1989.

[4] Farina, A., "Simultaneous measurement of impulse response and distortion with a swept sine technique", *108 AES Convention*, Paris, 2000.

[5] Pinelton, R., Schoukens, J., "Measurement and Modeling of Linear Systems in the Presence of Non-Linear Distortions ", *Mechanical Systems and Signal Processing,* 16(5), 2002.

[6] Bendat, J. S., and Piersol, A. G., "Engineering applications of Correlation and Spectral Analysis", Wiley, New York, 1980.

[7] Pierce, A. D., "Acoustics – An Introduction to Its Physical Principles and Applications", McGraw-Hill, NewYork, 1981.