Audio Engineering Society
Convention Paper 7315
Presented at the 124th Convention
2008 May 17–20 Amsterdam, The Netherlands

Perceptual Study and Auditory Analysis on Digital Crossover Filters

Henri Korhola[1] and Matti Karjalainen[1]

[1] *Helsinki University of Technology, Department of Signal Processing and Acoustics, Espoo, 02150 TKK, Finland*

Correspondence should be addressed to Henri Korhola (hkorhola@gmail.com)

ABSTRACT

Digital crossover filters offer interesting possibilities for sound reproduction, but there does not exist many publications on how they behave perceptually. In this research, phase and magnitude errors in digital implementations of linear phase FIR as well as Linkwitz-Riley crossover filters are studied perceptually and by auditory analysis. In a headphone simulation listening experiment we explored the just noticeable level of degradation due to crossover filter artifacts. In a real loudspeaker experiment we explored rough guidelines for "safe" filter orders of linear-phase FIR crossover filters, which would not produce audible errors. Possibilities to predict the perceived errors were then explored using auditory analysis, including also third-octave magnitude spectrum and group delay as simple auditory correlates. Linear-phase FIR crossovers were found to produce different kind of phase errors than Linkwitz-Riley crossovers. The auditory analysis can qualitatively explain the perceptibility degradation.

1. INTRODUCTION

As digital technology has made its entrance into the audio world a couple of decades ago, crossover filters have mostly remained analog until very recent times. An interesting option for loudspeaker designers is a crossover filter with steep attenuation and constant group delay to approach an "ideal" crossover filter [1]. This may be possible to acquire with sufficiently high-order, linear-phase FIR crossovers [2].

Many publications exist on digital crossover filters [3, 4, 5, 6, 7, 8, 2, 9], but there are hardly any perceptual studies. This is interesting, because the perceived reality is not often coherent with signal processing theory. A lone exception is [10], which has been published after this study had been launched. The studies overlap slightly, though the approaches can be considered different.

This paper presents perceptual results from a study of linear-phase FIR crossover filters in different scenarios. Also digital implementations of well-known Linkwitz-Riley (LR) crossover filters [11] were included in the experiment. Filter order, crossover frequency, and input signal were varied, as well as off-axis listening with resulting misalignment of arriving signal components was included.

Listening experiments documented in Section 3 were performed over headphones and the crossover filters were simulated in Matlab. A real loudspeaker experiment with linear-phase FIR crossovers was conducted in a standardised (ITU-R BS.1116) listening room. The main goal of the loudspeaker experiment was to find rough guidelines for "safe" FIR filter orders that do not produce audible errors.

The off-axis ringing phenomenon due to lack of time alignment [7], occurring with high-order FIR crossover filters, was of special interest in the listening tests. This is why the listening angle was simulated in addition to filter properties.

The paper consists of this introduction, after which background knowledge of crossover filters is given with a short review of related psychoacoustics in Section 2. The listening experiments and their results are presented in Section 3. Analysis and discussion on the results are given in Section 4. Auditory analysis is presented in Section 5.

2. BACKGROUND

Crossover filters are special filters that divide the audio spectrum for loudspeaker drivers so that they can operate on their optimal frequency ranges. Crossover filters are a necessity in multi-way loudspeakers. The basic division of crossover filters can be made to passive, active, or digital crossovers.

In a two-driver case, the transfer function of a crossover filter consists of the low- and highpass outputs:

$$H(s) = H(s)_{\mathrm{L}} + H(s)_{\mathrm{H}} \tag{1}$$

where $H(s)_{\mathrm{L}}$ and $H(s)_{\mathrm{H}}$ are the transfer functions of the low- and highpass subfilters, respectively. $s = \alpha + j\omega$ and ω is $2\pi f$, f being frequency.

Crossover filters should protect the drivers from not being fed with unwanted signals, which means their attenuation capabilities should be adequate, but at the same time they should affect the reproduction as little as possible. The goals of crossover design (in decreasing order of importance) can be suppressed to the following, adapting Lipshitz and Vanderkooy [1] and Linkwitz [11]:

1. Flatness in magnitude response of the combined low- and highpass outputs on the main listening axis:

$$H(\omega) = |H(\omega)_{\mathrm{L}} + H(\omega)_{\mathrm{H}}| \equiv 1 \tag{2}$$

2. Steep enough attenuation rates of the subfilters to optimise the reproduced frequency range for each driver.

3. Acceptable polar response, i.e. phase difference between low- and highpass outputs is zero:

$$\phi(\omega)_L - \phi(\omega)_H = 0 \tag{3}$$

4. Acceptable phase response:

$$\phi(\omega) = \arg\left[H(j\omega)_{\mathrm{L}} + H(j\omega)_{\mathrm{H}}\right] \tag{4}$$

being linear at best:

$$\phi(\omega) = -k\omega \tag{5}$$

The group delay is a commonly used measure of phase distortion in crossover filter analysis telling how much a certain frequency component or frequency range of a signal is delayed (the first derivative of phase):

$$\tau_g(\omega) = \frac{-d\phi(\omega)}{d\omega} \tag{6}$$

It is called also the envelope delay as it tells how much the envelope curve of a complex signal that contains many frequencies is delayed. It is usually given in either samples or in milliseconds.

2.1. Crossover Filter Types

Passive crossovers have been the majority in the past, but active crossover filters have become increasingly common, especially in professional audio. Digital crossover filters are newcomers in the audio field, but some implementations already exist.

Passive crossovers filter the signal after it has been power-amplified. This means the crossover network is connected to drivers, which creates impedance problems. This pitfall does not exist with active crossovers, which filter the signal before it is power-amplified. Therefore one power amplifier for each channel is needed, as well as an external power source for the active filter. Digital crossover filters can be made with an audio signal processor, being thus adjustable with properly written software and offering interesting properties in signal processing, such as steep separation of frequency ranges and phase linearity.

2.1.1. Linkwitz-Riley Crossover Filter

Linkwitz-Riley crossover filter (L-R) is one of the best known active crossovers [11]. It consists of two cascaded Butterworth low- and highpass subfilters. The transfer function of second-order L-R crossover (Equation (9)) consists of the low- and highpass transfer functions summed:

$$H(s)_L = \frac{1}{(1+s_n)^2} \tag{7}$$

$$H(s)_H = \frac{s_n^2}{(1+s_n)^2} \tag{8}$$

$$\begin{aligned} H(s) = H(s)_L - H(s)_H &= \frac{(1-s_n)^2}{(1+s_n)^2} \\ &= \frac{(1-s_n)(1+s_n)}{(1+s_n)^2} = \frac{1-s_n}{1+s_n} \end{aligned} \tag{9}$$

where s_n is the frequency normalized to the crossover frequency:

$$s_n = \frac{s}{2\pi f} \tag{10}$$

The Linkwitz-Riley crossover has uniform magnitude response on the main listening axis and relatively good off-axis response. Commonly used 4th order L-R crossover has attenuation rate of 24 dB/oct, which is adequate. The combined phase response is non-linear, producing phase distortion to the output signal, and so the question would be: when is it audible? This controversial question has been under discussion for decades. Linkwitz himself concluded in the original paper from 1976 as well as on his webpages that at least with 4th order L-R filter, the phase errors are not audible. We try to find guidelines for the audibility of L-R phase errors in our headphone simulation listening experiment in Section 3. It is already known that the audibility of phase distortion is individual, and depends on the sound material used [12].

2.1.2. FIR Crossover Filter

FIR crossover filters can offer so called "brickwall" filtering, which means very steep separation of the frequency spectrum. Another interesting property is the possibility of linear-phase reproduction. Importantly, it exists only when the low- and highpass outputs are perfectly summed for the combined output, i.e. on the listening axis, when the time-delay between drivers is zero. The transfer function of an FIR filter is:

$$H(z) = \sum_{n=N_1}^{N_2} h(n) z^{-n} \tag{11}$$

where $N_2 - N_1$ is the length of the impulse response, $h(n)$ is the discrete-time impulse response, z is the complex variable $re^{j\omega}$, and $r = \left|H(e^{j\omega})\right|$.

2.2. Psychoacoustics

Psychoacoustics is a field that studies the perception of sound. While the general opinion says that the magnitude response plays a major role in sound perception, the phase response should not be neglected. Many publications regarding the perception of phase distortion have been published [13, 14, 15, 16, 17, 12, 18, 19, 20, 21], and the conclusions are that:

1. Phase distortion is audible with certain signals. Impulsive sounds are the most susceptible.
2. Perception of phase distortion is individual: there are clear differences between subjects.
3. Phase distortion is more audible with headphones than with loudspeakers.

4. A "rule of thumb" for audibility limit of group delay deviation would be 1.6 ms independent of the center frequency with all-pass filters, according to [21].

In our experiments, we wanted to find out what would be the rough guidelines for phase errors with different signals. We will talk about group delay error, which means the peak-to-peak variation of group delay.

3. LISTENING EXPERIMENT

3.1. Motivation and Goals

As mentioned, the two crossover filters under study were linear-phase FIR crossover and a digital implementation of the Linkwitz-Riley crossover. The more interesting part of the experiments was the linear-phase FIR crossover, because we wanted to find out how its brick-wall filtering with linearity in phase response affects perceptually the sound material with different parameters.

The goals of the study can be set by the following:

1. How much group delay deviations from uniform group delay are allowed for different signals, i.e., what is the Just Noticeable Difference (JND limit)? How the type of the group delay error affects the audibility?

2. How sensitive are digital crossover filters for off-axis response errors?

3. Is it possible to predict the results of the listening test by objective measures?

4. Rough "safety-limits" for FIR filter orders that would not produce audible errors?

3.2. Description of Listening Test

3.2.1. Test Procedure

The experiment was conducted as a comparison between the original, unprocessed sound sample, and the crossover-simulated sound sample. Simulation with headphones was chosen as the primary test method to reveal small degradations, but also a more brief listening test was conducted with a real loudspeaker to compare the results. Because the loudspeaker test was different from the simulation experiment, it is discussed in its own Section 3.4, starting from Page 8.

Impairment	**Grade**
Imperceptible	5.0
Perceptible,not annoying	4.0
Slightly annoying	3.0
Annoying	2.0
Very annoying	1.0

Table 1: The ITU small impairment scale. Grade 5 presents the imperceptible impairment, and 1 presents bad disturbance in the signal. The scale was used at intervals of 0.5 in the listening test.

In addition to the author, nine untrained subjects listened to the comparison samples in random order, using a graphical user interface (GUI) to respond, and they were supposed to give grades for the basic audio quality difference between the unprocessed and the crossover filtered samples. Audiometry tests were not performed, but they were considered having normal hearing. A screen shot of the Matlab [22] GUI is presented in Figure 21. Oral introduction to the topic, instructions, and guidance were given before the test.

Grading was based on Table 1 of ITU's small impairment scale [23], which is used for judging small distortions in audio signals. The grade of 5 presents the imperceptible difference in the sample, and the grade of 1 presents high disturbances in the audio quality. The scale is guided to be used at intervals of 0.1, but due to feedback received from preliminary tests, it was instead used at intervals of 0.5. Comprehensive training would have been necessary for using steps of 0.1. The only trained subject for this test was the author due to preliminary listening of different scenarios. As the samples were graded in randomized order using the GUI in Matlab, the author could participate to the test.

3.2.2. Test Material and Parameters

The test material consisted of three different sound samples:

1. 10 Hz square wave

2. The castanets
3. The tom-tom drum

The selection of test material was based on preliminary tests of the topic as well as general knowledge of psychoacoustics. The square wave is known to be highly revealing for distortions in the audio systems due to its waveform and its contents of odd harmonics. It has also been used in many perceptual studies, such as in [17, 12]. The castanets were chosen as a test sound because of their transient nature. The castanet percussions are concentrated in a short period of time, and so the distortions would affect the final waveform in an audible way. The third sample was the tom-tom drum. It has also a slight impulse nature in its waveform, but not as concentrated as the castanets. It was an interesting sample, because the drum is a very common musical instrument, and in contrast to the castanets, it has low frequency contents.

Defining the parameters for simulated crossover filters was planned as carefully as possible. The crossover frequency, at which the audio spectrum is divided into low- and highpass bands, was chosen to be at either 100/300 Hz, 1 kHz or 3 kHz to approximate the common divisions in multi-way loudspeakers. The orders of the crossover filters were chosen to be realistic for practical implementation, and not to exceed 32 for L-R or 2000 for FIR crossovers. Some exceptions were included. The delay between low- and highpass outputs was implemented to simulate the elevation of the listening angle, because most of the problems exist when the vertical angle is changed. With quite a normal separation of drivers of a two-way loudspeaker (0.25 m), the delay was limited to 0-0.5 ms to simulate far-field elevation angles between 0 and 45 degrees. The parameters have been gathered in Table 2 to conclude the diversity of possible scenarios. The large number of different scenarios forced to cut out inaudible (to the author at least) samples to make the test reasonable in size and duration. For the same reason, the parameters were not always the same for different signals.

3.2.3. Test Equipment

The listening test was carried out with a laptop computer, which was equipped with Matlab software [22]. To ensure the sound quality to be at an adequate level, an external sound card Fast Track Pro of M-Audio was used. The headphones used in the experiment were high quality Sennheiser HD-600 headphones.

A real loudspeaker was also used for the test to find out the differences between headphone simulation and real listening in a room, and also to conclude approximate "safety limits" for FIR crossover filters. In this case the crossover filter under study was a linear phase FIR crossover at the crossover frequency of 3 kHz. The orders of low- and highpass subfilters varied from 300 to 2400. The loudspeaker was a two-way loudspeaker with a bass-reflex enclosure. The enclosure's volume was 12 liters. The loudspeaker was equipped with bass-range driver L18RNX/P and tweeter 25TAFN/G of Seas [24]. The magnitude responses of the woofer and tweeter are illustrated in Figure 1. Probable crossover artifacts occuring in a very narrow band, equalization of the responses was considered unnecessary. The woofer was 78 cm, and the tweeter was 92 cm above the floor. The listening test took place in the listening room at the Acoustics Laboratory of Helsinki University of Technology.

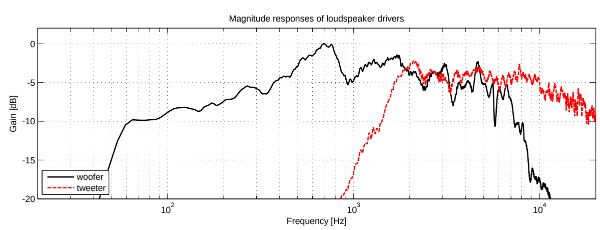

Fig. 1: Magnitude responses of drivers of the loudspeaker that was used in the real loudspeaker experiment.

3.3. Results of the simulated loudspeaker experiment

As there were so many varying factors, like Table 2 suggests, presenting and interpreting the results is not straightforward. The total count of different samples is 79, which are divided into 27 sub-groups. In each of these sub-groups some parameter is varied and its effect on the perception of distortions is studied. It must be emphasized that based on the preliminary tests, in many of the possible scenarios

Type	**L-R**	**FIR**
Signal	square/castanets/tom-tom	square/castanets/tom-tom
Order	4/8/12/16/20/24/32	500/600/700/800/1000/2000/10000
Crossover [Hz]	300/1000/3000	100/1000/3000
Delay [ms]	0/0.1/0.2	0/0.01 - 0.05/0.1 - 0.5

Table 2: Filter types and parameters used in the listening test. The number of possible combinations of different scenarios would have been too large, thus preliminary pruning was applied to the sample group.

distortion was considered inaudible, and they were left out of the test.

The chosen combinations of the results are presented for both FIR and L-R crossover filters. Different sub-groups have been plotted in the same plot in order to visualize the differences in the perception between signals. The results are plotted with Matlab's boxplot function. It plots the median as the red horizontal line inside the box. Lower and upper quartiles are plotted as the bottom and top line of the box. "Whiskers", the dashed lines, extend from the ends of the box to the adjacent values of the data, to the maximum of 1.5 times the interquartile range. The grades not in the range of 1.5 times the interquartile are plotted as outliers with an "x".

3.3.1. Results of FIR Crossover Filters

The results of 1000th to 2000th order FIR crossovers at 1 kHz are plotted in Figure 2. It can be seen that a 0.2 ms delay between drivers causes clearly perceivable distortion in the square wave. With a 2000th order FIR at 1 kHz, castanets are susceptible to distortions rather easily, and 0.3 ms delay between the low- and highpass bands produces a perceivable distortion. The tom-tom drum being the most insusceptible signal, it does not suffer from distortion even with 0.5 ms delay at 1 kHz.

Results for the two other crossover frequencies 100 Hz and 3 kHz are plotted in Figure 3. Excluding a couple of outlying grades, the results of FIR crossovers at 100 Hz tell that even very high orders do not produce audible distortions in any of the signals with the maximal delay between drivers (0.5 ms). This is illustrated in the top subplot of Figure 3.

On the contrary, differences become clearly audible at 3 kHz. The results of FIR crossover at 3 kHz with the maximum delay 0.5 ms are presented in the middle subplot of Figure 3. As the order of the crossover filter is varied, it seems that castanets begin to suffer from distortions somewhere between orders of 700 and 1000. The tom-tom presents clearly audible distortions at order of 700.

When testing the highest order of 2000 with FIR crossovers, interesting results occur, as the bottom subplot of Figure 3 shows. After the smallest delay of 0.01 ms, all the signals start to present audible distortions. The delay is really small, but still the medians of the results are clearly in the disturbance range (below or at the grade of 4). These results are interesting, because this type of FIR crossover should offer the "ideal" frequency domain characteristics of a crossover with very steep stopband attenuation and yet linearity in the phase response. However, as the audio spectrum is divided steeply in low- and highpass bands, ringing in the time domain occurs because of the Gibbs phenomenon [25]. This will be discussed later with impulse response demonstrations in Section 4.4.

3.3.2. Results of Linkwitz-Riley Crossover Filters

A digital implemenation of L-R crossover was simulated to study the audibility of phase distortion. The L-R crossover should offer a magnitude response of unity on the listening axis, but it suffers from phase distortion due to increasing group delay deviations as the order increases. Combined results of the Linkwitz-Riley implementations are presented in Figure 4. When the crossover frequency is at 300 Hz, the square wave presents more audible errors than the castanets or the tom-tom, like the top subplot of Figure 4 shows. Looking at the L-R crossover with an order of 16, the square wave results have received a median grade under 2, while castanets and

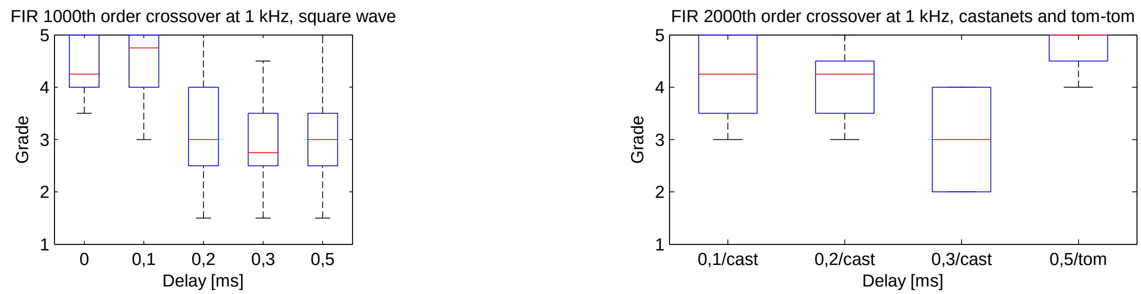

Fig. 2: Combined results of the listening test for FIR crossovers at 1 kHz. Square wave is susceptible to clearly audible degradation with 0.2 ms delay, castanets with 0.3 ms delay, while tom-tom receives high grade even with 0.5 ms delay.

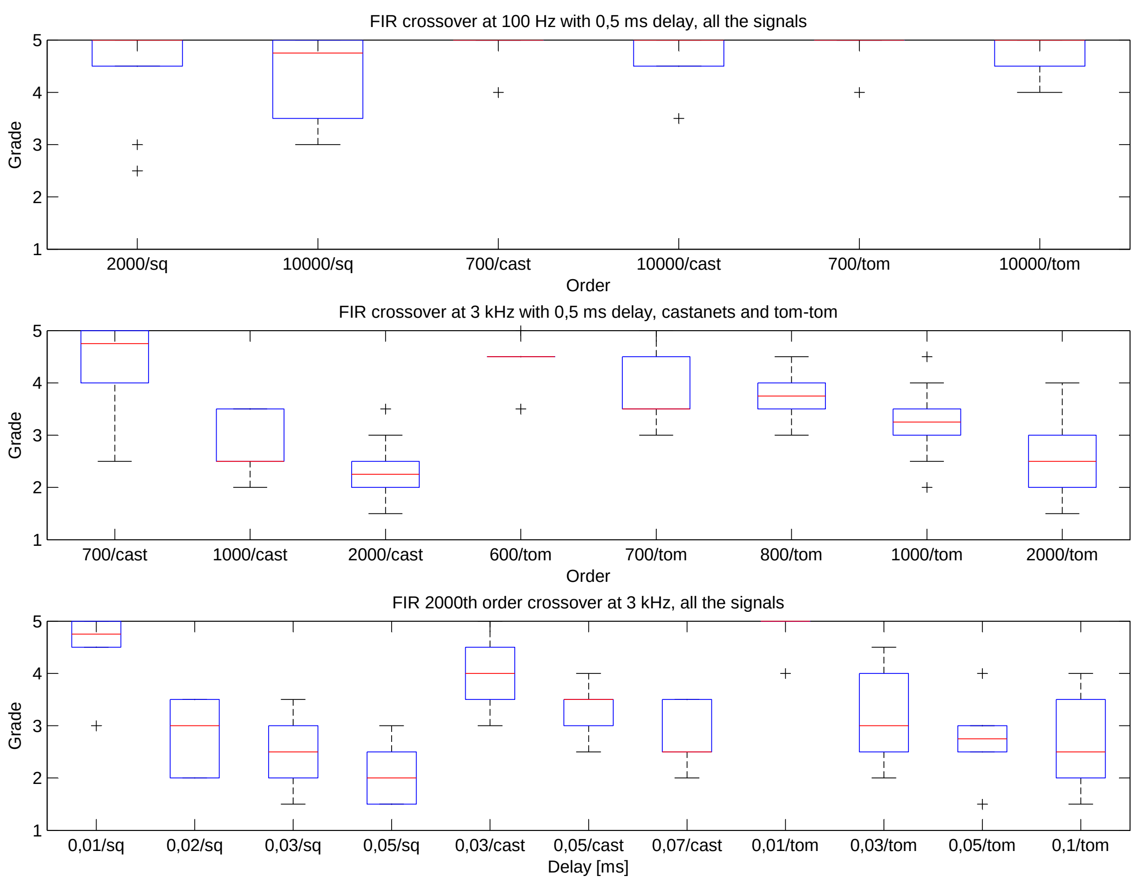

Fig. 3: Combined results of the listening test for FIR crossovers. Top: At 100 Hz no considerable degradations are perceived. Middle: At 3 kHz with long delay the grades seem to decrease after order or 600-700. Bottom: At 3 kHz with varying delay, even small delays seem to degradate the signals in an audible way.

tom-tom have median grades above 4.

At 1 kHz, the square wave once again suffers from distortion above the order of 8, whereas the castanets have received acceptable median grade up to order of 16 and the tom-tom does not seem to present considerable distortion up to order of 24. The results of the L-R listening test at 1 kHz for all the signals are plotted in the middle subplot of Figure 4.

With crossover frequency at 3 kHz, higher orders than with lower crossover frequencies seem to be acceptable in the light of the results. With the square wave, orders up to 20, and with castanets and tom-tom up to 32, the median grades are clearly above 4.

3.4. Results and Analysis of Loudspeaker Listening Experiment

A real loudspeaker was also used for testing linear phase FIR crossover filters. Comparing the results with headphone simulation was of interest. Five untrained but audio-related persons participated in the experiment. After seeing the results from headphone simulation, the test was decided to serve for finding the FIR filter orders that do not produce audible errors. The test signals used in the loudspeaker experiment were 10 Hz square wave and the castanets.

The listening positions in the listening room are depicted in Figure 5. The loudspeaker was placed as the left channel in stereo listening (for position 1). The experiment consisted of five different listening positions, in each of which the subject listened to the signals sitting on a chair as well as standing. Position 5 was directly in front of the loudspeaker so that the subject was directed toward the loudspeaker, different from positions 1-4. The room was a standardised (ITU-R BS.1116) listening room (see Section 3.2.3 on Page 5) with reverberation time of roughly 0.3 seconds at 500 Hz [26].

The threshold for audible errors were tried to find by playing pairs that consisted of the reference signal of low order (300) FIR crossover and the test signal with higher orders (600-2400) FIR crossover. Enough comparison pairs were played to find the order at which the test subject did not perceive errors.

The degradations were inaudible to all of the test subjects at the order of 600 for the 10 Hz square wave and at 900 for the castanets. Above these orders, most of the test subjects clearly perceived errors.

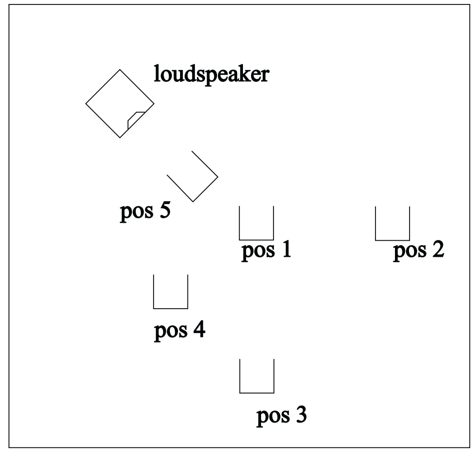

Fig. 5: Listening positions 1-5 of the real loudspeaker experiment. Loudspeaker was placed as the left channel in stereo listening (for position 1).

The "safe" orders that do not produce perceived errors seem to follow those of received from the headphone simulation for 10 Hz square wave (see Figure 10 on Page 11), while being larger for the castanets (see Figure 11 on Page 12). This is a natural occurrence especially for real life signals, as the reflections in a room make the perception of phase more difficult, compared to headphone listening without reflections.

The lowest orders that produced errors were recorded in front of the loudspeaker in position 5 (see Figure 5), when the subject was sitting on a chair. This is likely due to the dominance of direct sound over reflections. Qualitative inspections from the subjects suggested that the ringing of FIR crossovers was highly critical to the listening place. Slight changes in subject's head position could make the phenomenon either audible or inaudible. There were also differences between test subjects.

Remarkable is that the errors were clearly audible with a real signal and a real loudspeaker in a listening room for orders above 600 with the 10 Hz square wave. This suggests carefulness in designing and using digital crossover filters, especially linear phase

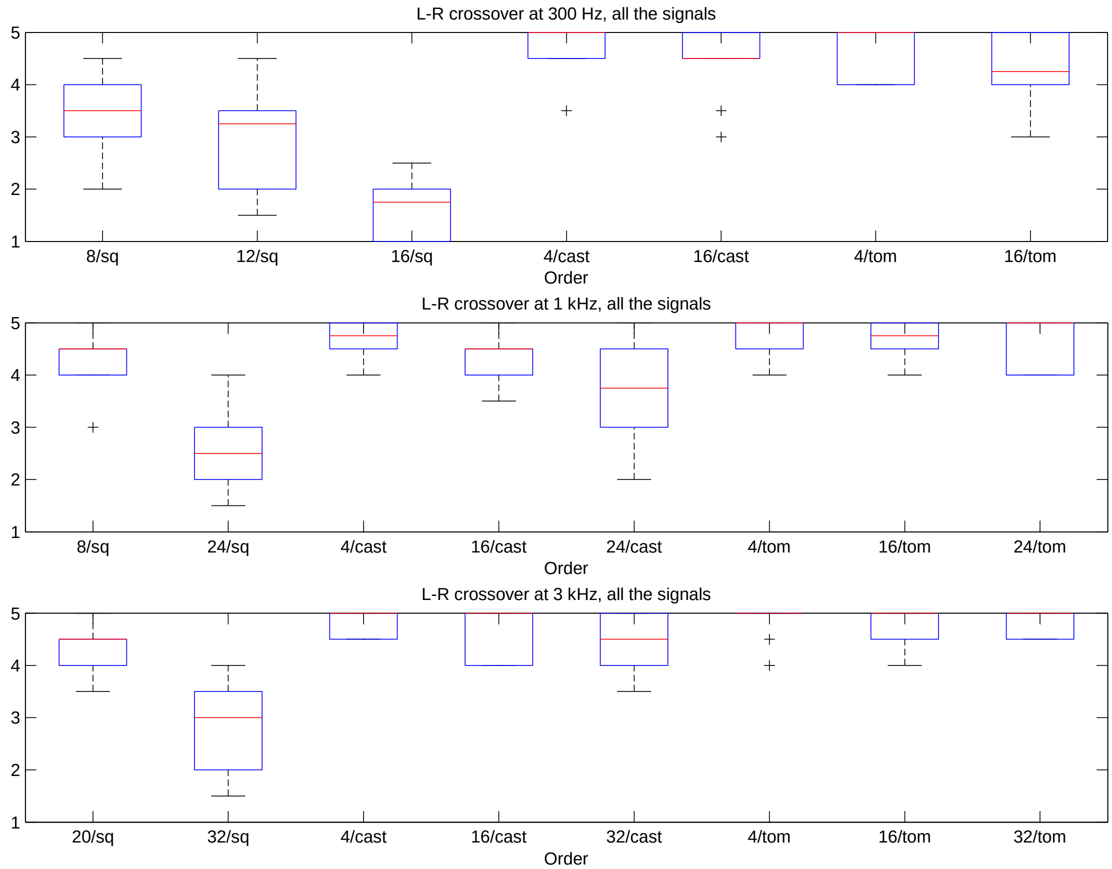

Fig. 4: Combined results of the listening test for L-R crossovers. Top: At 300 Hz, only the square wave seems to be degradated audibly, while the castanets and the tom-tom have quite high grades. Middle: At 1 kHz, the square wave suffers from degradation above the order of 8, the castanets above the order of 16 and the tom-tom does not present clear distortion even at the order of 24. Bottom: At 3 kHz, degradations become audible with the square wave above the order of 20, but the castanets and the tom-tom do not suffer from considerable distortions up to order of 32.

FIR crossovers with higher orders.

4. ANALYSIS AND DISCUSSION

The analysis concentrates on the computable measures of magnitude error and group delay error. In order to make the magnitude errors interpretable in terms of psychoacoustics, a smoothed third-octave spectrum was calculated in Matlab. It is a rough measure of resolution of the bandpass filterbank of hearing.

The reasons for audible errors can be of two natures. Regarding *magnitude errors*, a general rule of JND is about 1 dB deviation, if the sound pressure level is moderate and there is a comparable reference in the short term memory. In the crossover filter case, magnitude response should be uniform on-axis. In the experiment, L-R crossovers were tested both on- and off-axis. FIR crossovers were tested only off-axis, because the on-axis response should be nearly perfect independent of the crossover frequency and the filter order due to the "ideal" characteristics of a FIR type crossover filter.

4.1. Magnitude Errors of L-R Crossover Filters

For L-R crossovers having a good off-axis response and FIR crossovers approaching ideal cut-off with steep attenuation, smoothed third-octave spectra exhibit hardly any audible magnitude errors among the 79 test samples. The audibility limit is kept at 1 dB. The cases with a big magnitude error and a low average grade from the listening test are not so interesting. They are square wave signals filtered with 4th and 16th order L-R crossovers and having a 0.2 ms delay between drivers, and suffering from 8.8 dB and 3.9 dB peak magnitude errors. Regardless of that, the average grades are as good as 4.0 and 4.1, which suggests only a slightly perceptible change in the signal.

4.2. Group Delay Errors of L-R Crossover Filters

The JND limits of group delay distortions can be approximated from the results. They have been gathered in Table 3. Care must be taken in interpreting the results, but they offer guidelines for group delay limits. As a general conclusion, phase errors larger than roughly 1.5-2 ms due to crossover filter simulations may become audible. Furthermore, the absolute value of group delay error does not reveal all the problems.

Plotting the results from a different point of view may give a better insight into the problem. Grades from the listening test as a function of group delay are plotted in Figures 6, 7 and 8. The question is if something can be deduced from the plots. Looking at the figures, the grades and group delay errors seem to have a dependance. By these plots, the real world signals, the castanets and the tom-tom, would accept group delay errors of 3 ms before they become audibly distorted. No odd behaviour is recorded in the correlation between the grades and group delay errors, so guidelines can be drawn from the group delay error values that represent the phase distortion. Hence the point 1) of the experiment's goals is answered for L-R crossovers, though the rules of exact limits for different signals remain unclear.

4.3. Group Delay Errors of FIR Crossover Filters

The JND limits of group delay errors for FIR crossovers are not so straightforward, as we will find out. The grades from the listening test as a function of group delay error for the square wave are plotted in Figure 9. No strange behaviour exists and the slopes descend almost similarly to L-R crossovers in Figure 6.

Respectively, the grades from the listening test as a function of group delay errors for the castanets are plotted in Figure 10. This is the point where problems begin, regarding the correlation between group delay errors and grades. Remembering that the castanets and the tom-tom were quite insusceptible to

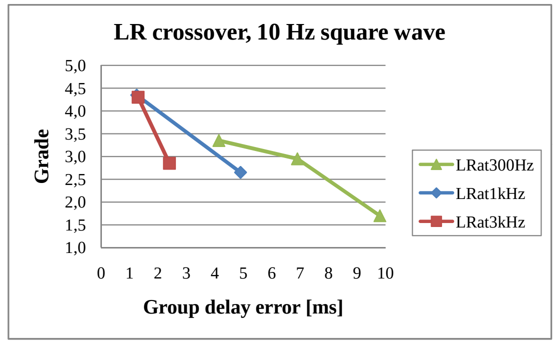

Fig. 6: Grade as a function of group delay error for 10 Hz square wave signal with L-R crossovers. There is no delay between drivers. The graphs behave regularly, descending as the group delay error increases.

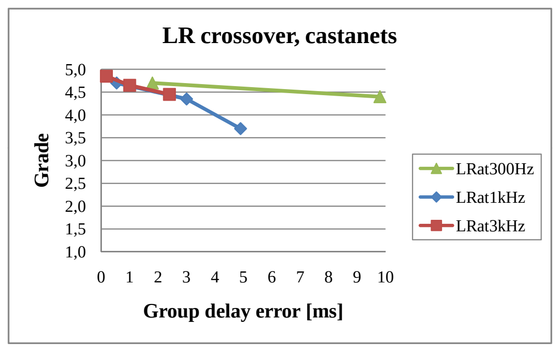

Fig. 7: Grade as a function of group delay error for castanets with L-R crossovers. There is no delay between drivers. As with square wave, the slopes are descending, though not as abruptly.

Signal	Crossover [Hz]	Group Delay limit [ms]
Square	300	1.8 - 4
	1000	1.25 - 4.9
	3000	1.3 - 2.4
Castanets	300	over 9.8
	1000	3 - 4.9
	3000	over 2.4
Tom-tom	300	over 9.8
	1000	over 4.9
	3000	over 2.4

Table 3: JND limits for audible group delay errors with Linkwitz-Riley crossovers. Due to sparsity of the data, the exact limits cannot be concluded.

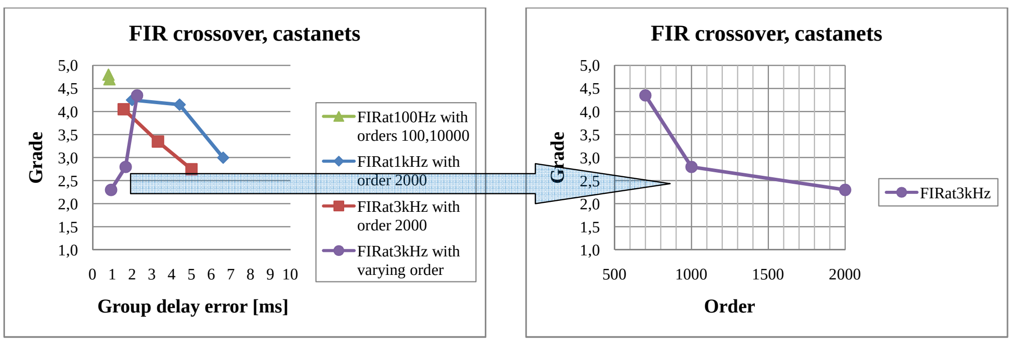

Fig. 10: Grade as a function of group delay error for the castanets with FIR crossovers. The last series in the left subplot is a FIR with varying order, and strange behaviour is noticed. It is extracted into the right subplot to show the effect of the filter order.

group delay errors, interesting results occur. The series of FIR crossovers with a varying order fits badly into the picture. The other series in Figure 10 show nice behaviour, but the last series in it has to be extracted to a plot of grade as a function of the filter order. The right subplot shows the extracted FIR case. It shows that straightly after an order of 700, the grade begins to decrease dramatically.

Finally, looking at Figure 11, we see the same kind of behaviour as with the castanets. The series of FIR crossovers with a varying order does not behave like the other series at all, which suggests that the value of group delay error is not explaining the perceived degradation. It is similarly extracted to the right subplot, like in Figure 10. The graph suggests that the tom-tom is even more susceptible to the errors with increasing order than the castanets. Immediately after the order of 500 problems begin, and the audio quality is not acceptable anymore.

The results imply that with very small magnitude errors the group delay deviations can explain the decrease in audio quality to some extent, but as Figures 10 and 11 clearly show, predicting audible errors by the values of group delay errors cannot always be done. Hence the point 1) of the experiment's goals remains unanswered for the FIR crossover case.

4.4. Ringing Phenomenon of FIR Crossover Filters

The explanation for the weird behaviour of grade vs. group delay plots is the ringing phenomenon,

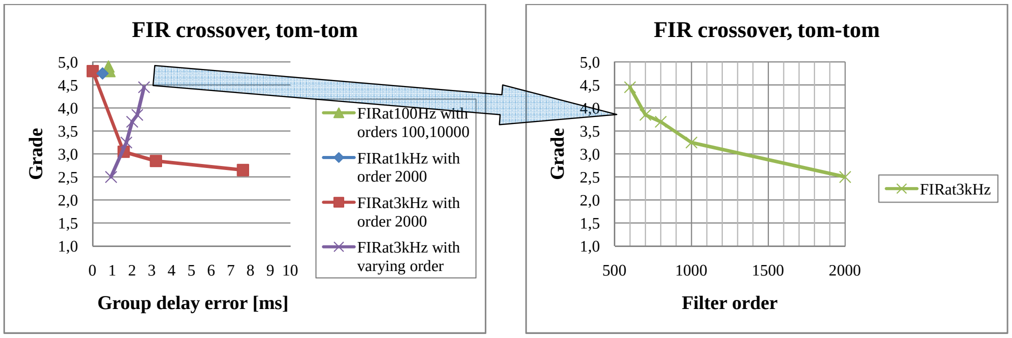

Fig. 11: Grade as a function of group delay error for the tom-tom with FIR crossovers. The last series in the left subplot is an FIR with varying order, and strange behaviour is again noticed. It is extracted into the right subplot to show the effect of the filter order.

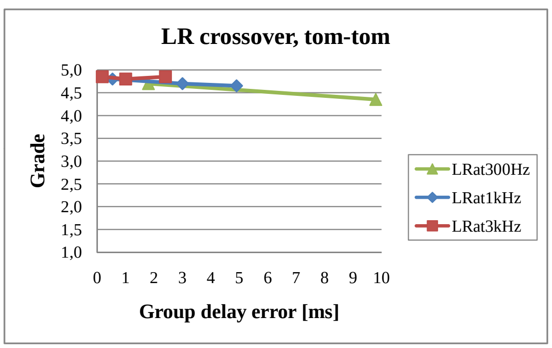

Fig. 8: Grade as a function of group delay error for tom-tom with L-R crossovers. There is no delay between drivers. As with the two other signals, the slopes are descending, but the rate is even smaller.

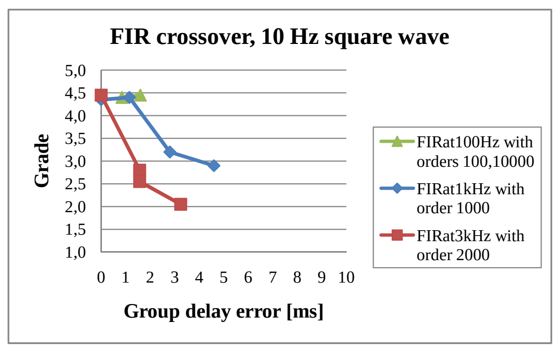

Fig. 9: Grade as a function of group delay error for the square wave with FIR crossovers. The graphs behave regularly, descending as the group delay error increases.

which occurs among FIR filters as the filter order increases. It happens due to the Gibbs phenomenon, which is the aftermath of the very steep attenuation of the low- and highpass filters [25]. On-axis it is no problem, because the low- and highpass impulse responses sum up nicely, and the crossover filter impulse response is just a delayed impulse. Off-axis the summing does not succeed and residual ringing will be left.

A comparison is made between two FIR samples with the castanets. Table 4 gathers the parameters. By the group delay error values, expecting a better grade for a 2000th order FIR crossover would be realistic. However, the average grade of the 700th order FIR is 4.4, while the 2000th order FIR has only received an average grade of 2.3. This seems counterintuitive according to the values of magnitude and group delay errors. The only factor which explains the difference is the much higher order of the crossover filter, apparently offering impressive properties for a crossover filter, but eventually substantially degradating the signal.

Regarding *ringing in the time domain*, zoomed plots of the impulse responses of 700th and 2000th order FIR crossover filters simulated at an off-axis position

Type	Order	Cross over [Hz]	Delay [ms]	Signal	GrpDel Err [ms]	Magn Err [dB]	Avg Grade
FIR	700	3000	0.5	cast	2.25	0.6	4.4
FIR	2000	3000	0.5	cast	0.93	0.2	2.3

Table 4: Comparison of samples to demonstrate the ringing phenomenon. The delay between drivers is 0.5 ms. Magnitude error is decreasing from 0.6 to 0.2 dB. Judging from the group delay and magnitude error values, the 2000th order FIR should receive much better grade, but the opposite is observed.

are presented in Figure 12. As the theory [25] dictates, the height of the ripple is unchanged when the filter order is increased, but the time span of ripples does change. This is clearly seen in Figure 12. The ringing lasts longer in the time domain at the both sides of the main response.

The time-domain masking effect in hearing seems not to prevent all the errors, because transient type sounds mask quite symmetrically in time [27]. This lays ground for the audibility of the error, especially when we notice that with linear phase FIRs there is always *pre- and post-ringing*, because of the symmetry of the impulse response. Often with real life signals, the signal itself masks the error. Because of the sharp rise and decay of the castanets and sharp rise the tom-tom signals, audible errors clearly remain. As seen in Figure 3 on Page 7, the castanets show susceptibility to errors at higher frequencies (1 and 3 kHz). With the tom-tom, the attack (rise) of the waveform is even steeper than in the castanet case, which could explain its susceptibility to ringing phenomenon errors. The decay of the tom-tom's waveform is slow compared to the percussion of the castanets, but pre-ringing could explain the errors.

Another demonstration of the ringing phenomenon is made between an L-R filtered and an FIR filtered tom-tom drum sample. Table 5 gathers the parameters of the samples, group delay error values, magnitude error values, and the average grades received.

As we can notice, group delay errors are quite small: only 2.4 ms for the L-R and 1.57 ms for the FIR crossover. Magnitude error is zero for the L-R because of listening on-axis and only 0.01 dB for the FIR according to the smoothed third-octave spectrum, which represents the bandpass filtering of hearing. So one could expect that the samples would receive approximately the same grades. Surprisingly, the average grades are far from each other, as the L-R crossover has received a grade of 4.9 and the FIR crossover only a grade of 3.1. The audible error that causes the low grade with FIR is a kind of "squeak" in the beginning of the drum hit, which is caused by pre-ringing phenomenon in the impulse response.

To study the problem more deeply, the impulse responses are first examined. They are plotted in Figure 4.4. There exists certain ringing in the L-R's impulse response, but it is practically decayed off after 5-6 ms. The duration of FIR's ringing in the impulse response is much longer. Worth noting is also the asymmetry of the L-R's impulse response. As the FIR crossover is realised with linear phase, the impulse response has to be symmetric [25]. This makes the pre-ringing phenomenon easily audible, as the signal itself masks the errors well forward with fast-rising, but slow-decaying sounds, such as the tom-tom. The same does not happen backward.

Figure 14 shows the behaviour of group delay of these crossover filters. The L-R crossover's group delay graph is on the left, and though it is not smooth, it does not have abnormalities in it. The FIR crossover's group delay graph is on the top right, and it has a very sudden change around the crossover frequency.

The essential question is: which one is audible, the minimal ripple in the magnitude response, or the strange behaviour of the group delay response? Both degradations are below the generally known and obtained audibility limits of group delay (see the results of L-R in Table 3 on Page 11) and of magnitude, which makes it tricky. As hearing performs analysis in both time- and frequency domains, the explanation for errors with FIR crossovers should be searched for with time-domain analysis.

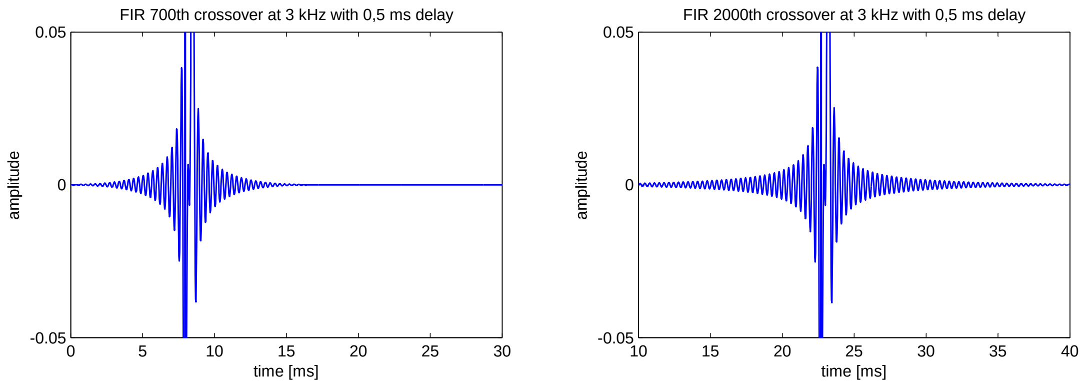

Fig. 12: Comparison of impulse responses of 700th order FIR (left) and 2000th order FIR (right) crossovers at 3 kHz in an off-axis position (the delay between drivers is 0.5 ms) to demonstrate the ringing phenomenon. Notice how the duration of ringing grows with increasing filter order.

Type	**Order**	**Cross over [Hz]**	**Delay [ms]**	**Signal**	**GrpDel Err [ms]**	**Magn Err [dB]**	**Avg Grade**
L-R	32	3000	0	tom-tom	2.4	0	4.9
FIR	2000	3000	0.03	tom-tom	1.57	0.01	3.1

Table 5: Comparison of samples to demonstrate the ringing phenomenon. The delay implemented to the FIR crossover is very small. Group delay errors are below the audibility limits received. Magnitude error of FIR is increasingly small, only 0.01 dB, which is well below the general perception level of 1 dB. Regardless, the average grades differ remarkably.

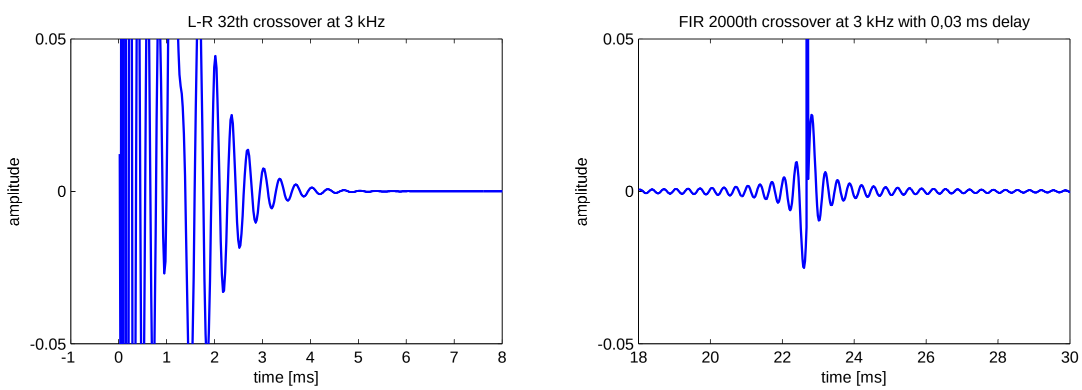

Fig. 13: Comparison of impulse responses of 32th order L-R (left) and 2000th order FIR (right) crossovers to demonstrate the ringing phenomenon. Notice the asymmetry in ringing of L-R's vs. the symmetry of FIR's. The impulse responses are closely zoomed to illustrate the phenomenon better.

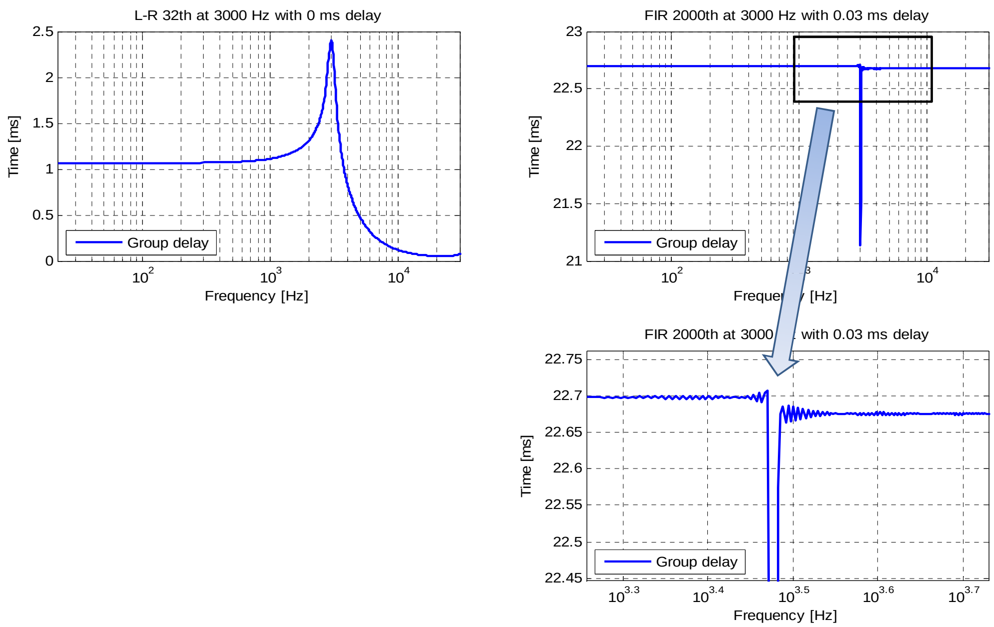

Fig. 14: Group delay graphs of L-R (left) and FIR (right) crossovers. Notice the abrupt changes of FIR crossover's group delay around the crossover frequency in the zoomed subplot (bottom).

However, the point 2) of the experiment's goals on Page 4 can be answered after these two inspections: High order FIR crossovers seem to be highly susceptible to off-axis errors. The delay of 0.03 ms corresponds only to a flight time difference of 1 cm between the loudspeaker's drivers. This equals an off-axis position of only 2-3 degrees when the loudspeaker drivers are separated by 0.25 m.

4.5. Conclusions from the listening Experiments

1. For Linkwitz-Riley crossovers, JND limits for group delay deviations vary between signals, with the 10 Hz square wave being the most susceptible, the castanets being less susceptible and the tom-tom being the least susceptible to phase distortions. The guidelines for JND limits can be read from Table 3 on Page 11.

2. For FIR crossovers, JND limits for group delay errors cannot be obtained, as no systematic correlation between group delay error values and grades exists.

3. The shape of the group delay graph might be crucial for perceived errors. FIR crossovers' group delay curves show irregularities compared to L-R crossover's group delay curves.

4. FIR crossovers seem to be highly susceptible to off-axis errors with higher filter orders. The flight time difference of only 0.02-0.03 ms between low- and highpass bands at 3 kHz was found to produce audible ringing with high FIR orders of the scale 1000-2000. Rough safety limits would be to keep the order of a linear-phase FIR filter at/under 600 at 3 kHz according to both the headphone simulation and the real loudspeaker experiment.

5. Predicting the results can be to some extent made for L-R crossovers by the group delay error values. Prediction with FIR crossovers seem

to be much more complicated, as the ringing caused by the Gibbs phenomenon causes different behaviour with higher filter orders.

6. The apparent characteristics of an "ideal" crossover filter turn out not to be pursued at any cost, because "brick-wall" attenuation with a linear phase response demands a relatively high-order FIR filter, which is exposed to the ringing phenomenon.

7. It seems obvious that the auditory effect of ringing has to be studied in the time domain.

5. AUDITORY ANALYSIS

5.1. Structure of Filterbank Model

A simple filterbank auditory modeling analysis was used to interpret the results of the listening test. Because the degradations in the signals due to the crossover filters occurred in a narrow band, the external and middle ear modeling were left out assuming that the auditory response is relatively flat in the crossover region. The auditory analysis was implemented in Matlab and consisted of the following steps:

1. Zero-phase (forward and backward) bandpass filtering using a bandpass Butterworth filter with a bandwidth corresponding to Bark scale [28]:

$$\Delta f_{\mathrm{cb}}[\mathrm{Hz}] = 25 + 75[1 + 1.4(f_{\mathrm{c}}[\mathrm{kHz}])^2]^{0.69} \quad (12)$$

   where f_c is the characteristic frequency of the auditory channel.

2. Full-wave rectification instead of half-wave hair-cell rectification to smooth the response by taking the absolute value of the signal.

3. Monaural time resolution by 3rd order lowpass filtering at 300 Hz for simulating auditory nerve synchrony. No adaptation or temporal integration was used.

Filtering with Butterworth type bandpass filter was done with Matlab's "filtfilt"-function. A second order bandpass filter was applied three times both in forward and reverse directions, and zero-phase filter was thus obtained with smooth, symmetrical responses. Using a single higher-order Butterworth type bandpass filter without reverse filtering (i.e. non-linear phase) was found to produce more ripple in the auditory channel envelope response, which made the interpretation more difficult, so it was omitted. Time-alignment of original and crossover-simulated signals was also easier with a zero-phase bandpass filter.

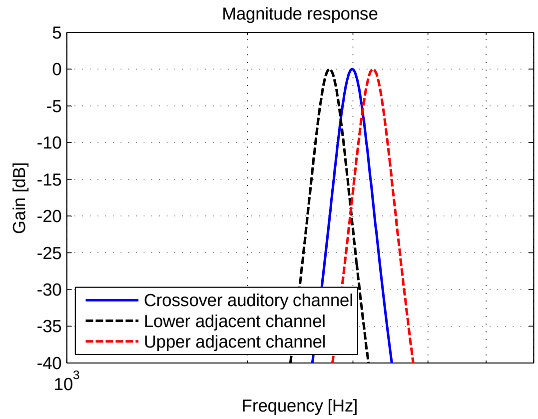

Fig. 15: Magnitude responses of used filterbank's three channels. The main channel's center frequency corresponds to the crossover frequency (here 3000 Hz). The lower (center frequency 2750 Hz) and upper (center frequency 3250 Hz) adjacent channels are also plotted. The -6 dB bandwidth is roughly 260 Hz.

The magnitude responses of the used filterbank in three auditory channels are illustrated in Figure 15. The main auditory channel's center frequency corresponds to the crossover frequency that in this case is 3000 Hz. Lower (center frequency 2750 Hz) and upper (center frequency 3250 Hz) adjacent channels are also plotted in the figure. The -6 dB bandwidth is approximately 260 Hz.

5.2. Auditory Response to 10 Hz Square Wave

5.2.1. FIR Crossover

First is studied how the model responses to the 10 Hz square wave signal in the time domain. As hear-

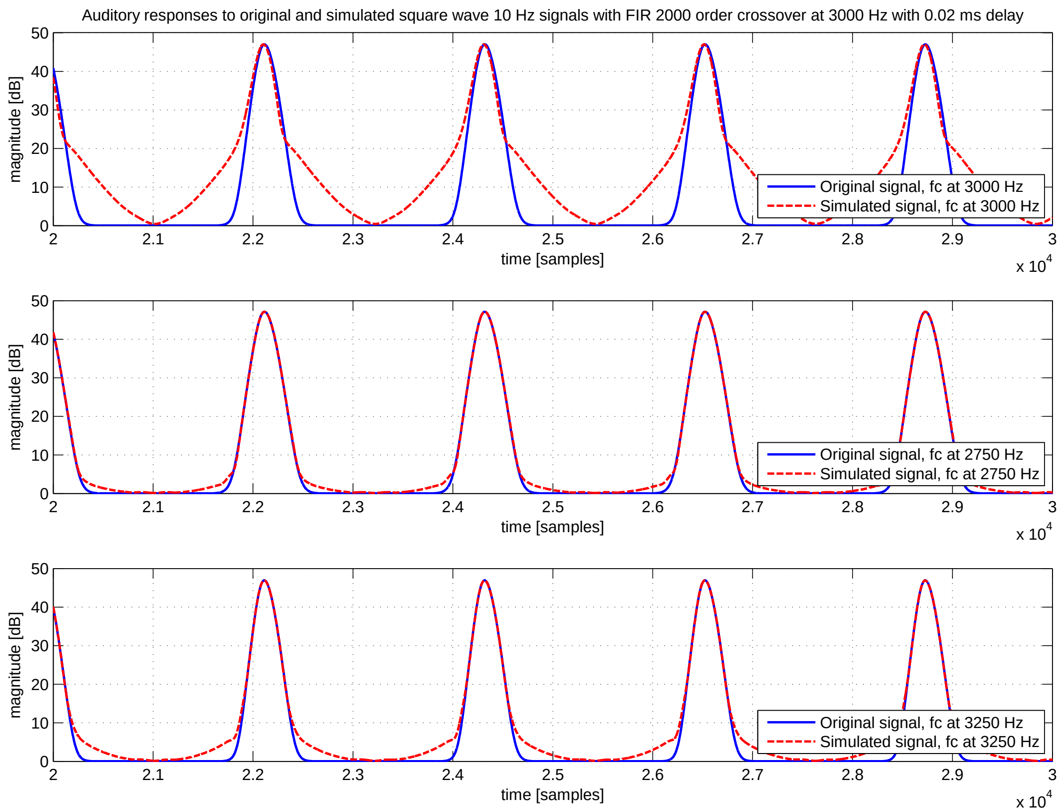

Fig. 16: Auditory analysis responses to 10 Hz square wave signal with 2000th order FIR crossover filter on three adjacent auditory channels: 3000 Hz (crossover channel, top subplot), 2750 Hz (lower channel, middle subplot), and 3250 Hz (upper channel, bottom subplot). The delay between drivers is now 0.02 ms. The crossover-simulated signal response is clearly spread in time on the auditory channel corresponding to the crossover frequency, which is heard as audible ringing in the sample. The average grade was 2.8.

ing perceives changes rather than the steady state in a stimulus, either the rising or falling edge of the square wave signal causes an auditory response. These auditory responses to the 10 Hz square wave signal on three channels (crossover channel at 3000 Hz, lower adjacent at 2750 Hz, upper adjacent at 3250 Hz) for a 2000th order FIR crossover filter are shown in Figure 16. The auditory responses to original signals are symmetrical and emerging at constant intervals of 50 ms, as expected from the square waves' waveform. Figure 16 shows also the spreading of the time domain response for the same FIR 2000th order crossover at 3 kHz when the delay between the loudspeaker drivers is 0.02 ms. Though behaving quite nicely on the lower and upper auditory channels, the crossover-simulated signal response is vastly spread on the auditory channel of 3000 Hz, which corresponds to the crossover frequency. In spite of the group delay error being 1.57 ms, very clearly audible ringing is heard when listened in the test. The case of FIR 2000th order crossover at 3 kHz with 0.02 ms delay has received an average grade of 2.8, which implies a clear degradation in audio quality.

5.2.2. Linkwitz-Riley Crossover

When studying the on-axis case with the Linkwitz-Riley crossover, it can be seen that despite of the group delay error of 2.4 ms, which is clearly audible according to the average grade of 2.9 from the listening test, the L-R crossover filter's time domain auditory response shows only slight difference between the original and crossover simulated samples, as Figure 17 illustrates. This is an example of different kind of a phase error than with FIR crossover's ringing that causes spreading in time.

Studying the off-axis case, the auditory response to the 10 Hz square wave signal with L-R 32th order crossover off the main axis is shown in Figure 18. The delay between drivers is now 0.2 ms. Now the auditory response to the square wave shows a clearer difference between the original and crossover simulated signals. The response curve has a bump on the left side on the crossover channel at 3 kHz, while the adjacent channels do not seem to have any abnormalities. It must be remembered that this sample has a magnitude error larger than 2 dB, which is above the JND limit.

5.3. Auditory Response to Castanets

Another illustration of the ringing phenomenon is presented in Figures 19 and 20. Table 4 on Page 13 gathers the parameters of the samples. Figure 19 shows the auditory responses to the castanet signal with a 700th order FIR crossover, suffering from weird behaviour at the top of the graph, but still receiving an average grade of 4.4, which means no audible distortion.

On the contrary, Figure 20 illustrates the auditory responses to the castanets with a 2000th order FIR crossover, and clear spreading is noticed. This means audible ringing in the sample, just as with the 10 Hz square wave signal. The smoothed third-octave spectra present only 0.6 and 0.2 dB magnitude errors and 2.25 and 0.93 ms group delay errors for 700th and 2000th order FIR crossovers, respectively. So both the group delay error and the magnitude error decrease as the order increases, but the average grade decreases from 4.4 for 700th order FIR crossover to 2.3 for 2000th order FIR crossover.

Analysing and noticing audible errors of FIR crossover filters may thus demand both analysing the group delay and magnitude error values as well as time domain analysis of the auditory responses.

5.4. Conclusions from Auditory Analysis

1. When phase errors are studied, L-R crossovers show completely different auditory response (differences in channels' group delays, see Figure 17 on Page 19) than high-order FIR crossovers (spreading in time, see Figure 16 on Page 17). Off the main axis, the auditory responses of L-R crossovers show clearer abnormalities (see Figure 18 on Page 20).

2. The ringing caused by the Gibbs phenomenon in high-order FIR crossovers seems to be visible in the auditory responses to 10 Hz square wave and castanet signals, which helps to predict the degradations, when the group delay error values do not imply any degradation. The tom-tom drum has so complex waveform that auditory analysis could not be interpreted in practice.

3. Quantitative measures of audio quality are challenging to find because of parallel time and frequency domain analysis of hearing. From points 1) and 2) it can be concluded that either group delay error values, magnitude error values, or temporal auditory analysis by inspection reveals the degradation in a signal. The L-R crossovers seem to correlate better with the group delay error values, while high-order FIR crossovers may show qualitatively clearer degradations in the auditory channel envelope responses.

6. ACKNOWLEDGEMENTS

Part of the work of Henri Korhola has been funded by the Academy of Finland, project No. 106956 (Virtual Acoustics and Audio 2015).

7. REFERENCES

[1] S. P. Lipshitz and J. Vanderkooy. A family of linear-phase crossover networks of high slope derived by time delay. *J. Audio Eng. Soc.*, 31:2–20, Jan/Feb 1983.

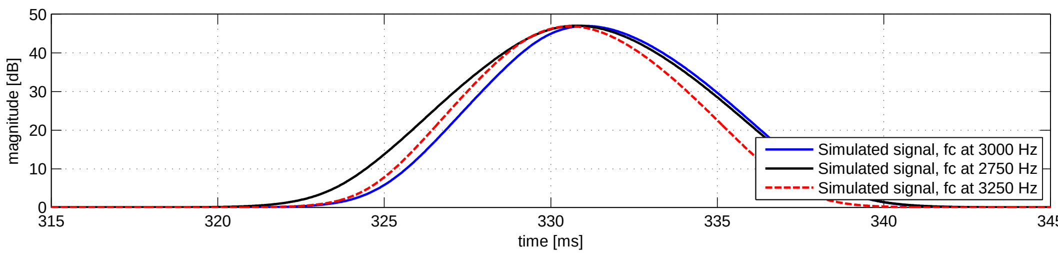

Fig. 17: Auditory analysis responses to 10 Hz square wave signal with 32th order L-R crossover filter on three auditory channels without the original signal: 3000 Hz (crossover channel), 2750 Hz (lower channel), and 3250 Hz (upper channel). The delay between drivers is 0 ms. Notice how the responses differ in time due to changing group delay of the L-R crossover, producing audible phase errors.

[2] J. Baird and D. McGrath. Practical application of linear phase crossovers with transition bands approaching a brick wall response for optimal loudspeaker frequency, impulse and polar response. In *The 155th Convention of the AES*, New York, October 2003.

[3] R. M. Aarts and A. J. M Kaizer. Simulation of loudspeaker crossover filters with a digital signal processor. In *The 82nd Convention of the Audio Engineering Society, London*, London, March 1987.

[4] R. Wilson, G. Adams, and J. Scott. Application of digital filters to loudspeaker crossover networks. *J. Audio Eng. Soc.*, 37(6):455–464, June 1989.

[5] P. L. Schuck and G. Klowak. Digital FIR filters for loudspeaker crossover networks. In *The 85th Convention of the Audio Engineering Society*, Los Angeles, November 1988.

[6] P. L. Schuck. Digital FIR filters for loudspeaker crossover networks ii: Implementation example. In *The 7th International Conference of the Audio Engineering Society*, Toronto, May 1989.

[7] R. Greenfield. Polar response errors in digital crossover alignments. *J. Audio Eng. Soc.*, May 1996.

[8] P. Reviriego, J. Parera, and R. Garcia. Linear-phase crossover design using digital iir filters. *J. Audio Eng. Soc.*, 46(5):406–411, May 1998.

[9] D. Ponteggia M.Di Cola, M. T. Hadelich and D. Saronni. Linear phase crossover filters advantages in concert sound reinforcement systems: a practical approach. In *The 121th Convention of the Audio Engineering Society*, San Francisco, October 2006.

[10] D. Z. Todorovic I. M. Salom and L. D. Milic. The influence of impulse response length and transition bandwidth of magnitude complementary crossovers on perceived sound quality. *J. Audio Eng. Soc.*, 56(11):941–954, November 2007.

[11] S. Linkwitz. Active crossover networks for noncoincident drivers. *J. Audio Eng. Soc.*, 24(1):2–8, Jan/Feb 1976.

[12] S. P. Lipshitz, M. Pocock, and J. Vanderkooy. On the audibility of midrange phase distortion in audio systems. In *The 67th Convention of the Audio Engineering Society*, New York, October 1980.

[13] R. Plomp and H. J. M. Steeneken. Effect of phase on the timbre of complex tones. *Journal of the Acoustical Society of America*, 46:409–421, March 1969.

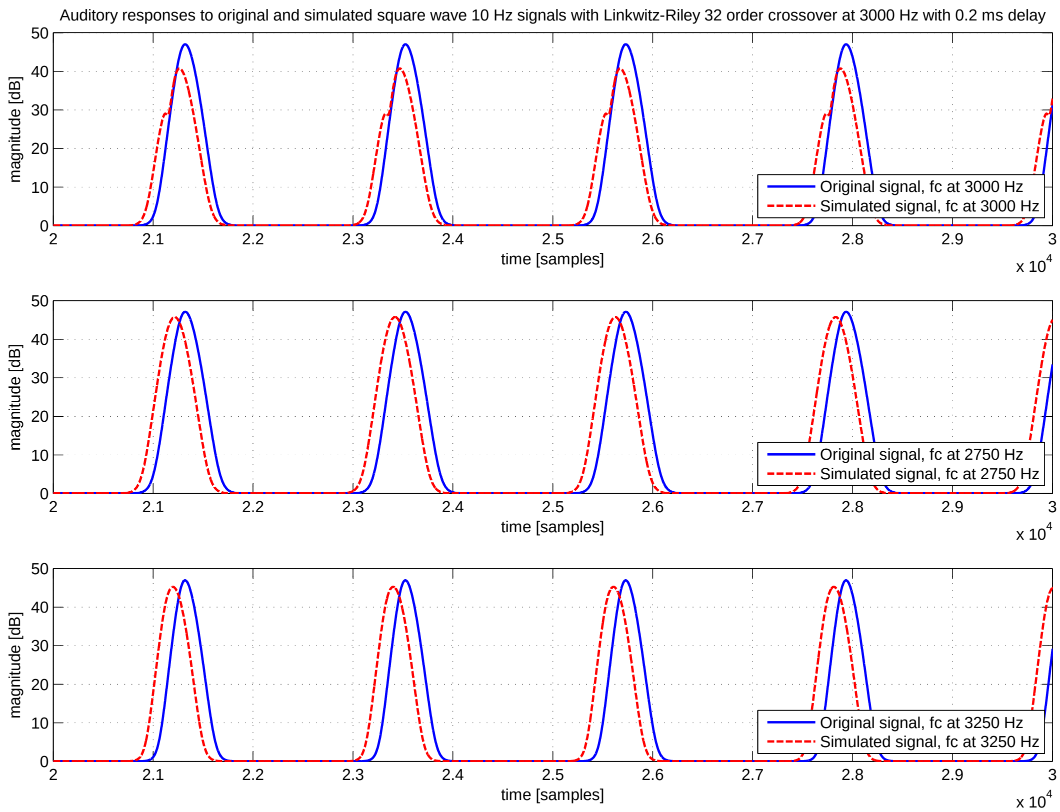

Fig. 18: Auditory analysis responses to 10 Hz square wave signal with 32th order L-R crossover filter off-axis on three adjacent auditory channels: 3000 Hz (crossover channel, top subplot), 2750 Hz (lower channel, middle subplot), and 3250 Hz (upper channel, bottom subplot). The delay between drivers is 0.2 ms. On the auditory channel corresponding to the crossover frequency, crossover-simulated signal response shows a bump on the left side. The average grade was 3.1.

[14] V. Hansen and E. R. Madsen. Threshold of phase detection by hearing. In *The 44th Convention of the Audio Engineering Society*, Rotterdam, February 1973.

[15] E. B. Jensen and H. Moller. On the audibility of phase distortion in audio systems. In *The 47th Convention of the Audio Engineering Society*, Copenhagen, March 1974.

[16] J. Blauert and P. Laws. Group delay distortions in electroacoustical systems. *J. Audio Eng. Soc.*, (63(5)), May 1978.

[17] H. Suzuki, S. Morita, and T. Shindo. On the perception of phase distortion. *J. Audio Eng. Soc.*, 28(9):570–574, February 1980.

[18] L.R. Fincham. The subjective importance of uniform group delay at low frequencies. In *The 74th Convention of the Audio Engineering Society*, New York, October 1983.

[19] J. A. Deer and P. J. Bloom. Perception of phase distortion in all-pass filters. In *The 77th Convention of the Audio Engineering Society*, Hamburg, March 1985.

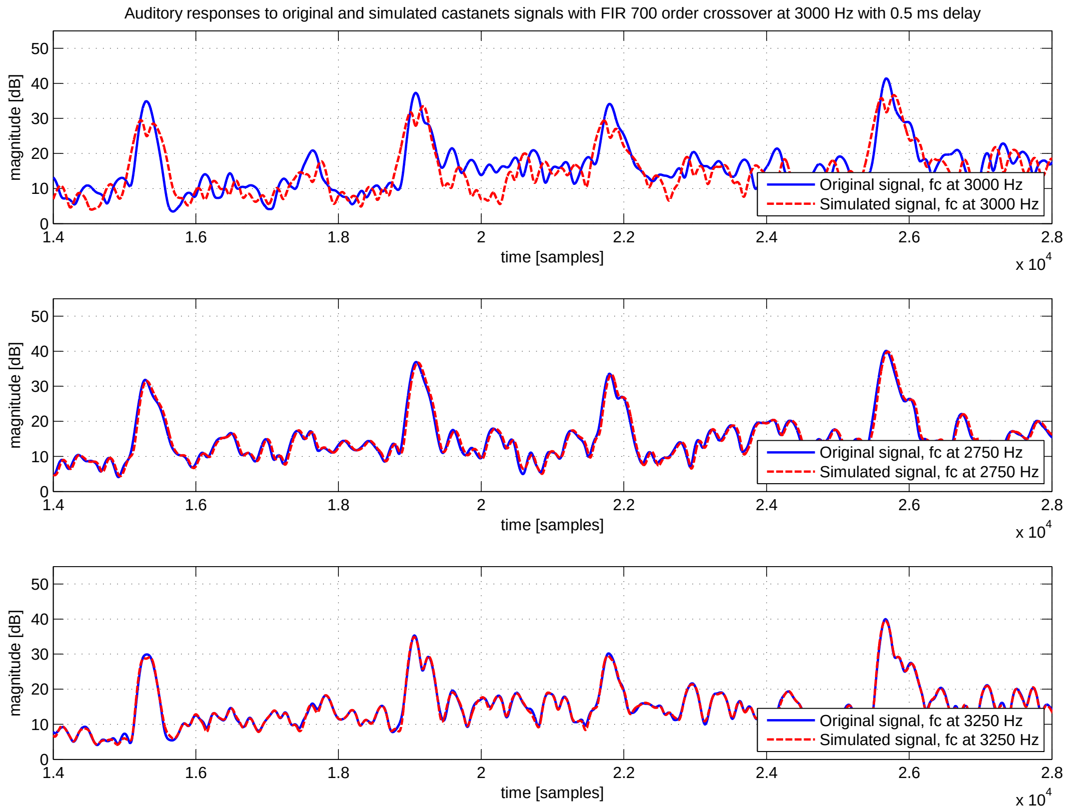

Fig. 19: Auditory analysis responses to castanet signal with 700th order FIR crossover filter on three adjacent auditory channels: 3000 Hz (crossover channel, top subplot), 2750 Hz (lower channel, middle subplot), and 3250 Hz (upper channel, bottom subplot). The delay between drivers is 0.5 ms. The center channel has a center frequency corresponding to the crossover frequency. On that auditory channel, the crossover simulated signal response is flattened from the top, but not much spread. The average grade was 4.4.

[20] G. J. Krauss. On the audibility of group distortion at low frequencies. In *The 88th Convention of the Audio Engineering Society*, Montreaux, March 1990.

[21] H. Moller, P. Minnaar, S. K. Olesen, F. Christensen, and J. Plogsties. On the audibility of all-pass phase in electroacoustical transfer functions. *J. Audio Eng. Soc.*, 55(3):115–134, May 2007.

[22] Mathworks Inc. www.mathworks.com. World Wide Web, December 2007.

[23] International Telecommunication Union (ITU). Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. ITU-R BS.1116-1, 1997.

[24] SEAS. www.seas.no. World Wide Web, January 2008.

[25] S. K. Mitra. *Digital Signal Processing: A Computer-Based Approach*. McGraw-Hill, Singapore, 1998. 864 pages.

[26] A. Jarvinen. Kuunteluhuoneen suunnittelu ja mallinnus. Master's Thesis, Helsinki University

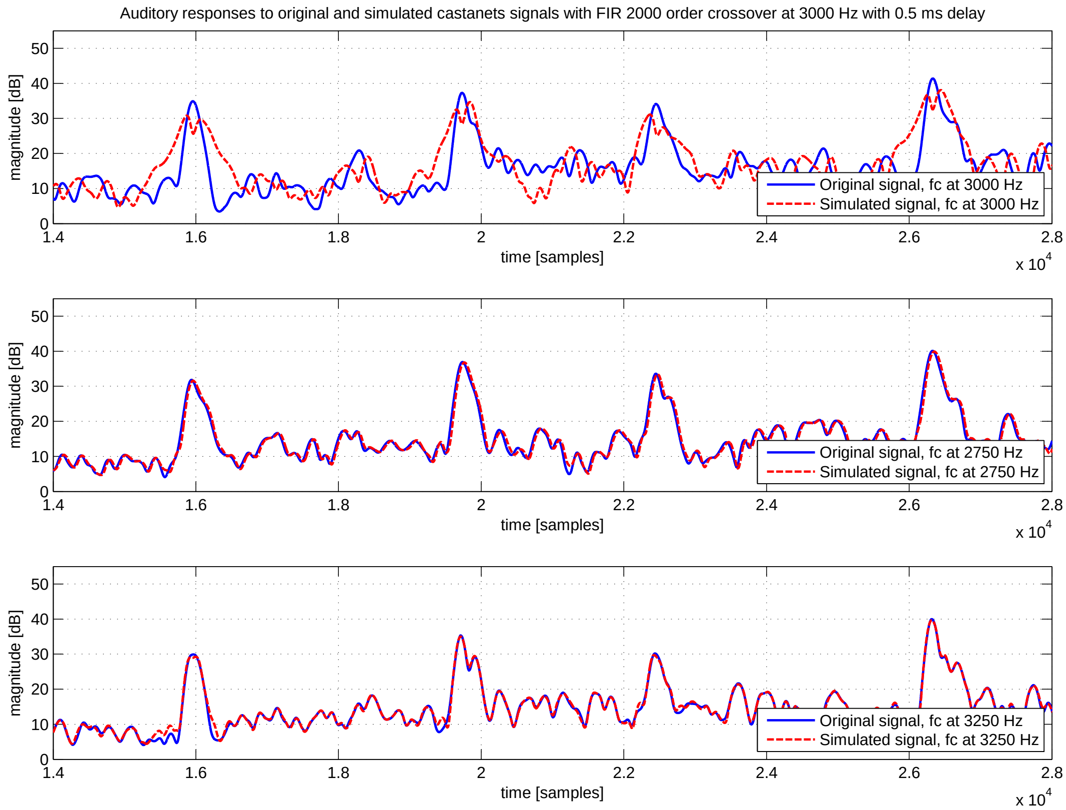

Fig. 20: Auditory analysis responses to castanet signal with 2000th order FIR crossover filter on three adjacent auditory channels: 3000 Hz (crossover channel, top subplot), 2750 Hz (lower channel, middle subplot), and 3250 Hz (upper channel, bottom subplot). The delay between drivers is 0.5 ms. On the auditory channel corresponding to the crossover frequency, crossover simulated signal response is flattened from the top, and also clearly spread in time. The average grade was 2.3.

of Technology, 1999.

[27] M. Karjalainen. *Communication Acoustics (in Finnish)*. Otamedia Oy, Espoo, revised 1st edition, 2000.

[28] E. Zwicker and H. Fastl. *Psychoacoustics: Facts and Models*. Springer, 3rd edition edition, 2006.

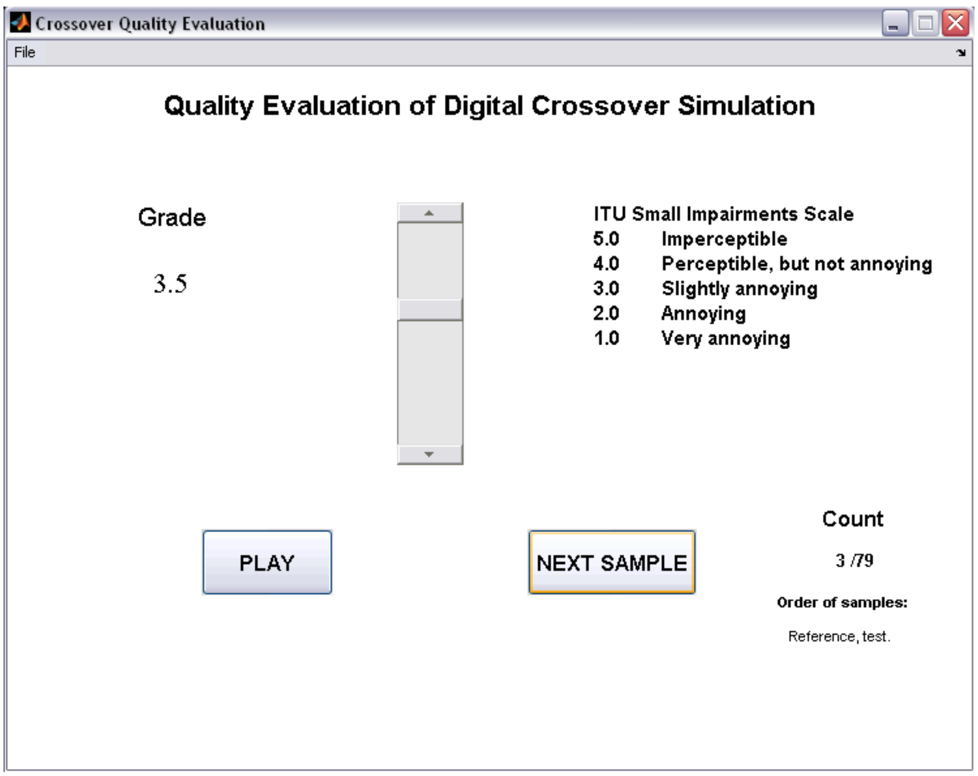

Fig. 21: Graphical User Interface of the Listening Test