Concert Hall Impulse Responses — Pori, Finland: Reference

Juha Merimaa[1], Timo Peltonen[2], and Tapio Lokki[3]
juha.merimaa@hut.fi, timo.peltonen@akukon.fi, tapio.lokki@hut.fi

[1]Laboratory of Acoustics and Audio Signal Processing
Helsinki University of Technology
P.O.Box 3000, FI-02015 TKK, Finland

[2]Akukon Oy Consulting Engineers
Kornetintie 4 A, FI-00380 Helsinki, Finland

[3]Telecommunications Software and Multimedia Laboratory
Helsinki University of Technology
P.O.Box 5400, FI-02015 TKK, Finland

May 6, 2005

Abstract

This document describes a set of impulse response measurements made in the Promenadikeskus concert hall in Pori, Finland. Altogether three source points and seven receiver points were utilized. The responses were measured using an omnidirectional sound source and a subwoofer, as well as a selection of microphones. The published database includes measurements with a pair of omnidirectional microphones, a pair of sideways pointing cardioid microphones, a dummy head, and a SoundField MKV microphone system. Additionally, diffuse field equalized dummy head responses and SIRR processed 5.0 responses are provided. The document gives a detailed description of the measurement and post processing procedures. Results of a comprehensive analysis of the responses are presented in an accompanying document (Merimaa et al., 2005).

1 Introduction

During years 2000–2002, a series of acoustical measurements were conducted in major Finnish concert halls within the framework of the TAKU/VÄRE technology project. The goal of the measurements was not only to gain more information of the measured halls, but also to acquire impulse responses with high enough quality for auralization and subsequent perceptual studies. Finding and partly developing suitable methods for the measurements, post processing, and analysis of the responses was a result of considerable work. This document describes the different steps related to measurements in the Pori Promenadikeskus concert hall on November 29th, 2002. The document also serves as a reference for a selection of the acquired responses, released free for noncommercial use (see copyright, Appendix B).

The measurements were performed using an omnidirectional loudspeaker and a subwoofer, both excited with sinusoidal sweep signals. The published responses include measurements with a pair of omnidirectional microphones, a pair of sideways pointing cardioid microphones, a dummy head, and a SoundField microphone system. The nonideal power responses of the sound sources were carefully compensated, and for each pair of source-receiver positions, the responses to both sound sources were combined with a digital cross-over. Furthermore, the different sensitivities of the receivers were compensated, and the levels of the responses are aligned relative to the power of the sound sources as measured in an anechoic chamber. For auralization purposes, the responses were denoised, and multichannel responses for loudspeaker reproduction with a standard 5.0 setup were computed with the Spatial Impulse Response Rendering (SIRR, Merimaa and Pulkki, 2004, 2005) method. In order to help interpreting the results and selecting appropriate responses, the data corresponding to each source-receiver pair were also analyzed using standard room acoustical parameters and graphical methods, as presented in the accompanying document (Merimaa et al., 2005).

The scope of this document is two-fold: As a reference documentation it describes the measurement and postprocessing procedures in considerable detail. However, where appropriate some low-level background information is included to help potential users that are less familiar with microphone and reproduction techniques, and/or concert hall acoustics. The document is organized as follows. The hall and the measurement positions are introduced in Section 2. The measurement system and the utilized acoustical transducers are presented in Section 3. Section 4 includes a reference description of the processing involved in extraction of the impulse responses and in the system compensation. Section 5 describes the applied (partly optional) post processing related to auralization of the responses, including some background on the reproduction techniques. Finally, Section 6 summarizes the document. Appendix A gives a reference listing of the provided sound files and their channel specifications, and Appendix B includes the copyright statement. The responses as well as this and the accompanying document (Merimaa et al., 2005) presenting analysis results of the responses can be downloaded from `http://www.acoustics.hut.fi/projects/poririrs/`.

2 Hall and measurement positions

The measurements were performed in the concert hall of Promenadikeskus located in Pori, Finland. The 700 seat hall was designed by Architect Company Güttner and built in 1999. The acoustic consulting was provided by Akukon Oy Consulting Engineers. The hall is illustrated with architectural drawings in Fig. 1, photographs in Figs. 2–4 and average room acoustical parameters in Table 1. For a description of the parameters[1] and the computational methods, as well as further analysis of the individual source-receiver combinations, see the accompanying document (Merimaa et al., 2005).

The hall is roughly shoebox shaped with dimensions of 33 x 23 x 15 m (length x width x height) yielding a total volume of approximately 9300 m^2. There are balconies on both sides, and the floor rises towards the rear part of the hall. A variety of diffusers has been installed on the walls, and the hall has a set of hanging reflectors above the stage (see Figs. 3 and 4) projecting more early sound both to the audience as well as to the musicians. The upholstered seats have a folding seat cushion with perforation on the bottom plate, yielding a relatively small difference between the absorption of an occupied and an unoccupied seat. The first three rows of seats can be removed but they were present during the measurements.

The measurements were conducted in an empty hall with three source positions on the stage, four receiver positions in the audience area, and three receiver positions on the stage. The positions were chosen according to recommendations of Gade (1989), and they are illustrated in the floorplan in Fig. 5 and listed in Table 2. Each stage position is specified with the distance to the foremost point on the stage (on the midline, denoted front of the stage) in the direction of the midline, as well as with the distance of the position from the midline. The receiver positions in the audience area are specified with row and seat numbers, where row 1 is the first (removable) row in front of the stage and seat numbers are given relative to the midline of the hall. The distance of the

[1] The symbols of the parameters should be familiar to concert hall acousticians except for LF_P and LF_{SF}, where the subscripts signify measurements with two different microphones.

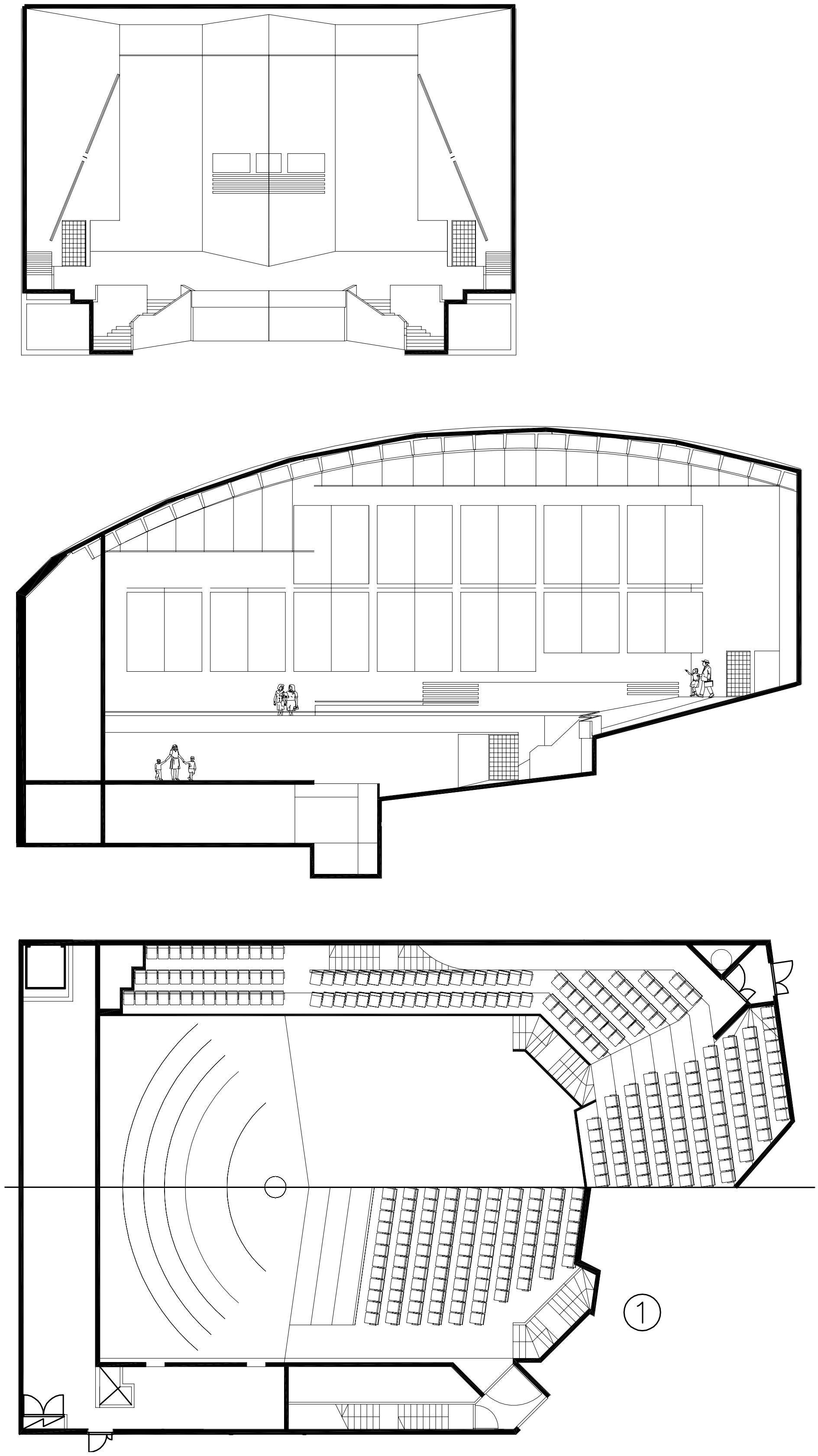

Figure 1: Architectural drawings of the Pori Promenadikeskus concert hall as seen from the front (top panel), from the left side (middle panel), and from above (bottom panel). The upper and lower halfs of the bottom panel illustrate the (symmetrical) organization of the seats on two different levels.

Figure 2: Pori Promenadikeskus concert hall as seen from the right side of the stage.

Figure 3: Pori Promenadikeskus concert hall as seen from the back of the (upper) floor.

Figure 4: A view on the side of the hall (left panel) and the dummy head and torso placed on a seat for measurements (right panel).

	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
T_{30} (s)	2.7	2.5	2.4	2.4	2.1	1.7	1.2
EDT (s)	2.5	2.4	2.4	2.3	2.0	1.6	0.9
G (dB)	8.9	8.7	8.9	9.6	9.9	8.2	4.3
C_{80} (dB)	-0.3	-3.6	-3.1	-1.8	-0.6	1.6	6.6
$1-\mathrm{IACC}_E$	0.10	0.26	0.74	0.74	0.74	0.73	-
$1-\mathrm{IACC}_L$	0.12	0.31	0.87	0.89	0.94	0.94	-
LF_{P}	0.24	0.24	0.28	0.37	0.34	0.41	-
$\mathrm{LF}_{\mathrm{SF}}$	0.37	0.33	0.34	0.37	0.48	0.59	-
SNR (dB)	66.1	65.7	67.3	72.0	74.3	74.2	70.5

Table 1: Room acoustical parameters analyzed at octave bands and averaged over all responses with receiver positions in the audience area.

receivers from the stage increases with their numbering. R1, R2, and R3 were located on the main floor, wheras R4 was on the elevated rear floor. For all positions, left and right are defined as seen from the floor when looking towards the stage. The average room acoustical parameters presented in Table 1 were calculated over all source positions, and receiver positions R1–R4.

During the measurements, the sound sources were consequently placed at each marked source position while the receivers were left untouched at their current positions. After the responses to both sources at all source points were measured, the receivers were moved to their next positions and the same procedure was repeated.

3 Measurement system

The responses were acquired using the IRMA measurement system described earlier by Peltonen (2000) and Peltonen et al. (2001). The system hardware consists of a rack mounted PC, a multi-channel sound card (Korg 1212 I/O), and external A/D converters (Korg 880). The measurements were performed using custom software running in the Matlab environment. For all measurements, 48 kHz sampling rate and 16-bit A/D and D/A conversion were used.

3.1 Sound sources

Two sound sources were utilized in the measurements. High frequencies were excited with an omnidirectional speaker consisting of 12 Vifa MG10SD09 4” loudspeaker elements mounted on a 300 mm dodecahedron cabinet constructed of 18 mm plywood. The frequency range of the measurements was extended with an EAW SB48e subwoofer unit consisting of two 8” loudspeaker elements. Both sources were individually fed using different channels of a Lab.gruppen LAB 1000 power amplifier. During the measurements, the omnidirectional speaker was mounted on a tripod at a height of 100 cm from the stage floor (Gade, 1989), wheras the subwoofer was set on top of a 10 cm high wooden base.

3.2 Microphones

The selection of microphones consisted of a pair of omnidirectional measurement microphones (DPA Type 4006), a stereo cardioid microphone (Pearl TL-4), a dummy head (Brüel and Kjær HATS custom fitted with DPA Type 4053 microphones), and a SoundField MKV microphone system (multi-capsule microphone and a related preamplifier and processor). Signals from the omnidirectional, cardioid and binaural microphones were amplified using a Sonolab custom 2 channel microphone amplifier and the microphone front end of a Tascam DA-P1 portable DAT recorder.

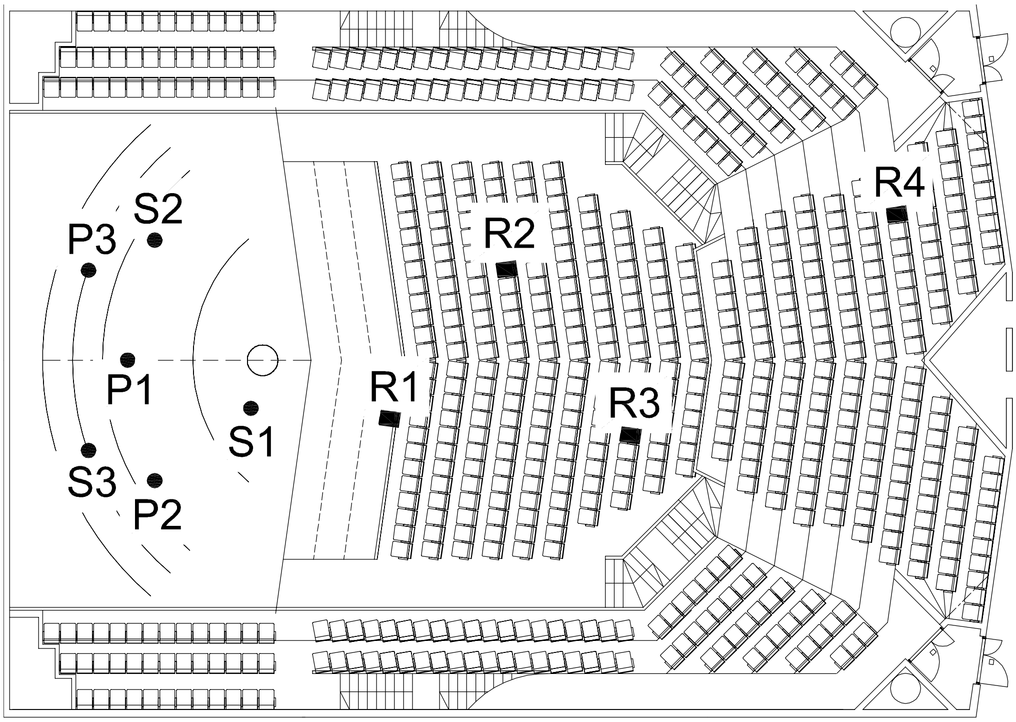

Figure 5: Floorplan of the Pori Promenadikeskus concert hall. See text and Table 2 for a description of the indicated measurement positions.

Source positions

Source	Distance on the stage from front (m)	from midline (m)
S1	2.0	1.6 left
S2	5.2	4.0 right
S3	7.4	3.0 left

Receiver positions on the stage

Receiver	Distance on the stage from front (m)	from midline (m)
P1	6.1	0.0
P2	5.2	4.0 left
P3	7.4	3.0 right

Receiver positions in the audience area

Receiver	Row	Seat from midline
R1	3	4^{th} left
R2	7	6^{th} right
R3	11	5^{th} left
R4	20	8^{th} right

Table 2: Source and receiver positions.

The rightmost DPA 4006 was also used extensively in the reference measurements described in Section 4.

The front direction for each receiver position in the audience area was defined as the direction towards the stage perpendicular to the backs of the seats at the corresponding location. At each measurement position, the dummy head was placed approximately where the head of a human listener sitting on the same seat would be located. Due to practical reasons, the other microphones were mounted on a stand slightly forward from the center of the dummy head, at a height of 115 cm corresponding approximately to the height of the ears of the dummy head. The positioning of the microphones was not precise enough to assume phase synchrony between different microphones. Both of the omnidirectional microphones were facing the front direction and separated by 10 cm in the left-right direction. The cardioids of the Pearl microphone were pointing to the left-right directions with the main purpose of these measurements being computation of the lateral energy fraction parameter (see Merimaa et al., 2005, Section 2.4). The SoundField microphone was aligned with the positive X direction towards the front.

For the receiver positions on the stage, the front direction was defined as the direction perpendicular to and towards the front wall. Only the DPA 4006 and SoundField microphones were used on the stage

Unfortunately, when starting the measurements, a contact failure was found in one of the figure-of-eight outputs of the SoundField processor. In order to overcome this problem, the microphone was rotated and the channels were switched such that the operational channels were aligned at the front-back (X) and the left-right (Y) directions. The steerable output of the preamplifier was adjusted to yield the missing up-down (Z) signal.

3.3 Excitation

The impulse responses were measured using logarithmic sinusoidal sweep excitation. Such sweeps have been earlier shown to yield superior results compared to other common excitation signals, when using nonideal transducers (Farina, 2000; Müller and Massarani, 2001; Stan et al., 2002). The special property of the sweeps is that they allow separation of the harmonic distortion products from the linear response in the deconvolution phase. This does not only directly decrease the measurement artifacts in the recovered impulse responses, but also enables measurements with considerably higher sound pressure levels (SPL) than is optimal, for instance, when using maximum length sequence (MLS, Schroeder, 1979) excitation. Furthermore, possible slight time variances due to temperature drifts and air movement in the hall do not corrupt the results when measuring with a single sweep.

For our measurements, the sweeps were synthesized in the time domain (Farina, 2000; Müller and Massarani, 2001). The sweep for the subwoofer extended from $35 - 400$ Hz lasting 2 s, and the sweep for the omnidirectional loudspeaker spanned the frequency range between $60 - 20000$ Hz lasting 4 s. In both cases the acquisition continued after the end of the excitation signal, with the full acquisition time being approximately 5.46 s (262144 samples) for both sweeps. Excitation levels were set experimentally to yield an optimal SNR while still allowing the separation of the distortion products from the linear responses. The sweep signals are illustrated in Fig. 6.

Note that logarithmic sweeps are characterized with a lowpass magnitude spectra with a decay of 3 dB/octave. This corresponds roughly to the spectral shape of background noise in many typical measurement situations, which means that the concentration of more energy to low frequencies is also a desired feature.

4 Extraction of impulse responses

After recording the responses to the sweep excitation signals, further processing is needed to recover the corresponding impulse response. For our purposes, all required processing was combined into two filters, one for the measurements with the subwoofer and one for the measurements with

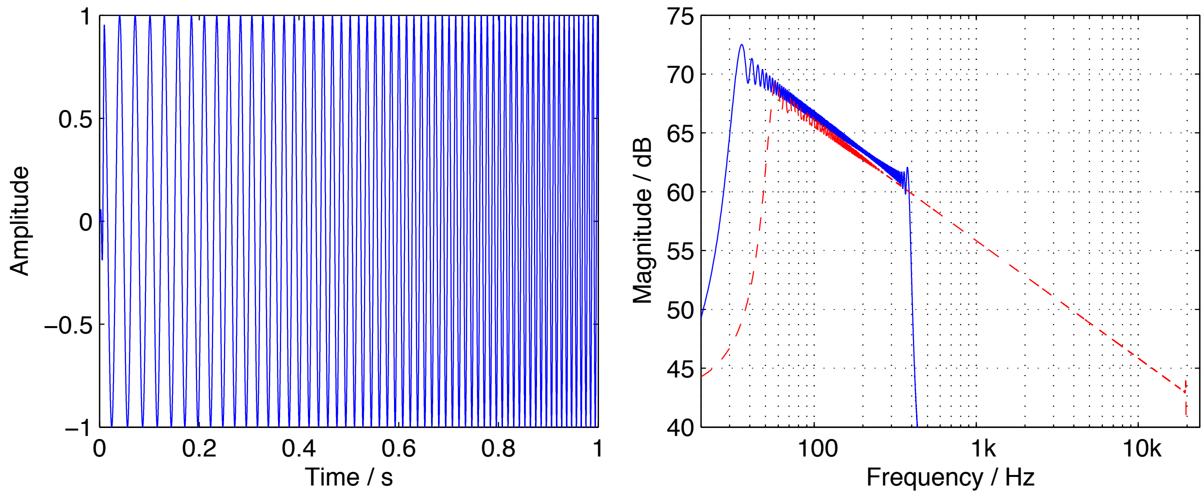

Figure 6: Left panel: Time-domain plot of the first second of the subwoofer sweep. Right panel: Magnitude spectra of the exitation signals for the subwoofer (solid line) and the omnidirectional speaker (dashed line).

the omnidirectional loudspeaker. These filters, applied in frequency domain, realize deconvolution of the sweep excitations, compensation of the magnitude responses of the subwoofer and the omnidirectional loudspeaker, as well as the cross-over between them. Finally, all that is needed is summation of the filtered responses to both sound sources. In the following, the design of each processing step is described in more detail. All filter design and filtering was performed in Matlab using double precision floating point arithmetics.

4.1 Deconvolution and cross-over

The sudden onsets and offsets created by the time domain sweep synthesis method result in some ripple in the excitation spectra, as can be seen in Fig. 6. In order to properly compensate this ripple, the deconvolution was designed in the frequency domain (Müller and Massarani, 2001). The spectra of the sweeps were first inverted, and the inverse spectra were zero-phase bandpass filtered in order not to boost the measurement noise outside the frequency range of the excitation. Since the omnidirectional loudspeaker and the subwoofer were located at slightly different heights during the measurements, a steep zero-phase crossover filter was designed to alleviate phase cancellation problems. 150 Hz was chosen as the cross-over frequency optimizing the overall signal-to-noise ratio (SNR).

The results of applying the deconvolution and cross-over to the excitation signals are illustrated in Fig. 7. Summing the responses yields a zero phase bandpass filtered impulse with a flat passband (± 0.1 dB) and -3 dB points at 35 Hz and 19 kHz. The bandpass corresponds roughly to the roll-off of the magnitude spectra at the extrimities of the excitation signals. Note that the use of non-causal zero phase filters (which is possible in the frequency domain) spreads the impulse evenly to both directions in the time domain.

4.2 Source compensation

In order to compensate for the effect of the sound sources, their impulse responses were measured with the rightmost of the DPA 4006 microphones (denoted reference DPA) in the large anechoic chamber of the Acoustics laboratory of Helsinki University of Technology. At high frequencies, the directivity pattern of the omnidirectional loudspeaker splits into several beams. Thus, for compensation of the average response, the speaker was measured from altogether 108 directions.

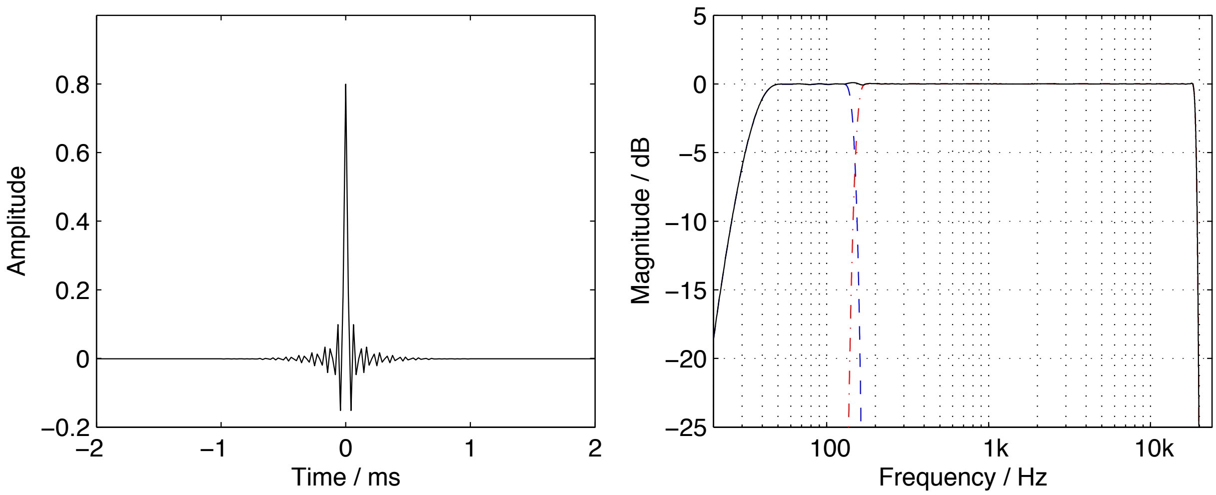

Figure 7: Left panel: Time domain impulse response of the excitation signals after deconvolution, cross-over and summation. Right panel: Magnitude spectra of the individual excitation signals after deconvolution and cross-over (subwoofer = dashed line, omnidirectional speaker = dash-dotted line), as well as the sum of the two previous responses (solid line).

The power spectra of the measurement results were averaged to approximate the diffuse field response of the speaker. The resulting power spectrum was smoothed with 1/3 octave resolution, and a minimum phase FIR filter was designed to compensate the smoothed response. A similar procedure, although with fewer measurement points, was applied to design the compensation filter for the subwoofer. The phase response of the omnidirectional loudspeaker was also studied, but the directional dependence of the phase especially at high frequencies made exact compensation impossible. Furthermore, the minimum phase filter appeared to yield good compensation results.

The inverse (non-smoothed) power responses and the designed compensation filters are illustrated in Fig. 8. For both sound sources, the frequency range of compensation was limited at the other end close to the 150 Hz cross-over frequency. The power response of the subwoofer can be seen to be almost flat between $50 - 300$ Hz. A limited boost was applied to extend the frequency range to 35 Hz. The power response of the omnidirectional loudspeaker is far less ideal, although typical for dodecahedron sources with a small cabinet. Nevertheless, the compensation filter flattens the response to ± 1 dB within a range of $100 - 17000$ Hz.

Based on a further measurement of the sound sources in the anechoic chamber, the gain of the compensation filters was adjusted to yield the same power for both sources at a frequency range of $100 - 200$ Hz. The measurement was performed with the reference DPA at a distance of 2 m from the sound sources. Furthermore, the measured and system compensated responses were combined with the cross-over described in the previous section, and the energy of the combined response was calculated. Since the same measurement system and settings were used in the concert hall, this serves as a reference for the level calibration discussed in the next section. Note that Fig. 8 does not indicate the final levels of the compensation filters.

The choice of the low and high frequency limits of the overall system compensation was made such that the level of exitation is still high enough above the background noise to yield meaningful measurement results. The gain applied to the extremities of the frequency range is still fairly aggressive, and it boosts the background noise somewhat. Furthermore, the directivity of the omnidirectional source deviates considerably from omnidirectional at the highest frequencies. However, for auralization purposes it was considered important to extend the frequency range as far as possible. The noise problems are also dealt with in a separate denoising process described in Section 5.1.

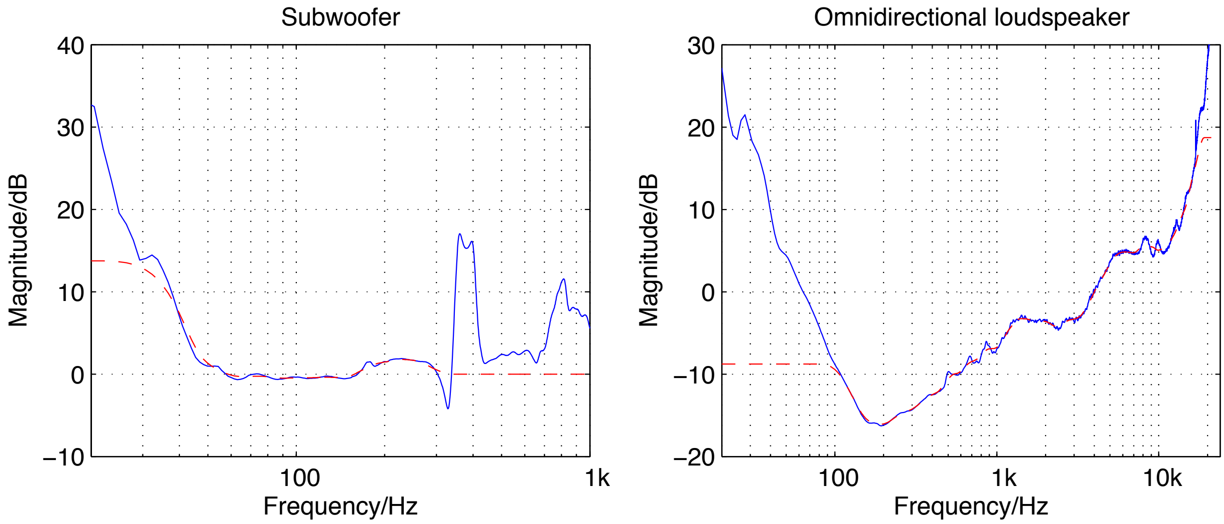

Figure 8: Inverse power responses of the sound sources (solid lines) and the magnitude responses of the compensation filters (dashed lines).

4.3 Microphone level calibration

For the computation of the room acoustical parameters (see Merimaa et al., 2005) it is preferable to normalize levels relative to the energy of an impulse response of the sound source measured at a distance of 10 m in free field. This level was arbitrarily set to -10 dB relative to the maximum of the wav files used as the final format of the impulse response database. Since we only had free-field measurements of the sound sources at a distance of 2 m, a scaling procedure was applied to find the normalization gains. Assuming an ideal point source, the sound energy decays relative to the inverse of the square of the distance to the source. Thus, according to the chosen level alignment the energy of an impulse response measured at a distance of 2 m in free field should be $-10 + 10\log_{10}(10^2/2^2) \approx 4$ dB. Based on the reference measurements of the sound sources (see the previous section), a normalization gain was determined for the reference DPA. The level setting was still verified using measurements of the omnidirectional loudspeaker at a distance of one meter in the hall. The energy of the responses pruned to include only the direct sound was within 1 dB of the expected $-10 + 10\log_{10}(10^2/1^2) = 10$ dB.

The relative sensitivities of the different microphones were determined from impulse response measurements of a Genelec 1032 loudspeaker at the same distance in the anechoic chamber. The level alignment relative to the reference DPA was performed by equalizing the A-weighted energy of each response. The reference for the dummy head was determined with a sound source in front of the head, and for the SoundField microphone, a single gain factor was calculated for the omnidirectional W signal and applied to both the omnidirectional and figure-of-eight signals. Furthermore, a slight spectral tilt in the order of 3 dB over 50 Hz – 20 kHz exhibited by the leftmost DPA microphone was equalized so that the frequency response corresponded to the reference DPA which was considered more reliable.

5 Post processing related to auralization

The previous section described the processing steps that were necessary to recover the impulse responses and to compensate some non-idealities of the measurement system. On the contrary, this section discusses processing methods that are optional in the sense that they do not make the responses more ideal or increase their information content. However, depending on application, the methods can still be very useful. Out of the techniques described in the following subsections, denoising was applied to all responses in the database, wheras diffuse field equalization and Spatial Impulse Response Rendering (SIRR) were used to produce additional responses for binaural headphone reproduction and multichannel loudspeaker reproduction.

5.1 Denoising

As discussed earlier, the inherent background noise of the hall and the measurement devices limits the achievable signal-to-noise ratio (SNR). Outside the extremities of the frequency range the measurements resulted in SNRs in the order of $60-80$ dB depending on the source-receiver positions and frequency. Exact SNRs for the omnidirectional measurements are listed in the accompanying document (Merimaa et al., 2005). The achieved SNRs are more than enough for characterization of the hall using standard room acoustical parameters (ISO 3382, 1997). However, the noise can still be audible when listening to the responses. One solution would be to cut the responses before the decay reaches the level of the background noise. However, a perceptually better method is to extrapolate the responses with random noise having the same time-frequency decay characteristics as the beginning of the response (Jot et al., 1997).

The analysis and denoising algorithm of Jot et al. (1997) operates in the frequency domain using short-time Fourier transform (STFT). The algorithm is based on a model of reverberation assuming a frequency-dependent exponential decay of the room responses. The decay at each frequency band is estimated by fitting a line to the logarithm of the backward integrated echogram (Schroeder, 1965, see also computation of the reverberation time in Section 2.1 of the accompanying document) before the response reaches the level of the background noise. The parameters of the line, as well as the transition time from measurement results to background noise, are found using an iterative procedure. Once these data are known, the responses are processed such that at each frequency band, the tail of the responses corrupted by background noise is replaced by random Gaussian noise decaying with the estimated rate.

The responses in the published database were denoised using 1024 sample STFT with 75 % overlapping time frames. The time frames were windowed with a 4th order Blackmann-Harris window in the analysis phase, and with Hann windows in the actual denoising procedure (Jot et al., 1997). Due to the limited frequency range of the omnidirectional loudspeaker, the analysis was limited to $0-17$ kHz and the frequencies above 17 kHz were replaced with Gaussian noise starting at the same time with the direct sound at lower frequencies and decaying at the rate of the measurement data averaged over $15.5-17$ kHz. Furthermore, the beginning of the responses was set to zero with a 2 ms fade-in using a raised cosine onset ramp just before the arrival of the direct sound. Several denoised responses are depicted in the figures in (Merimaa et al., 2005). In each case, also the transition time from the measured data to the extrapolated noise is indicated.

Note that the extrapolation of the high frequencies creates an unrealistic early response. This will be visible, for instance, in wide-band decay plots of the responses where too much sound energy can be seen between the early reflections. For pure analysis purposes it is advisable to lowpass filter the responses.

5.2 Diffuse field equalization of binaural responses

Diffuse field equalization relates to reproduction of the binaural responses. The internal microphones of the utilized dummy head are placed approximately where the eardrums of a human listener would be located. The measurement data already include the effect of the head, the pinnae and the ear canals on the sound, and are thus ideal for analysis of perception using, e.g., auditory models. However, in auralization, the sound should be delivered unmodified to the eardrums of the listener, which is not directly possible with typical reproduction systems. Undesired coloration may be produced either by the transmission path, or due to design features of the reproduction system.

Consider first playback of the responses through a pair of loudspeakers. Although this does not

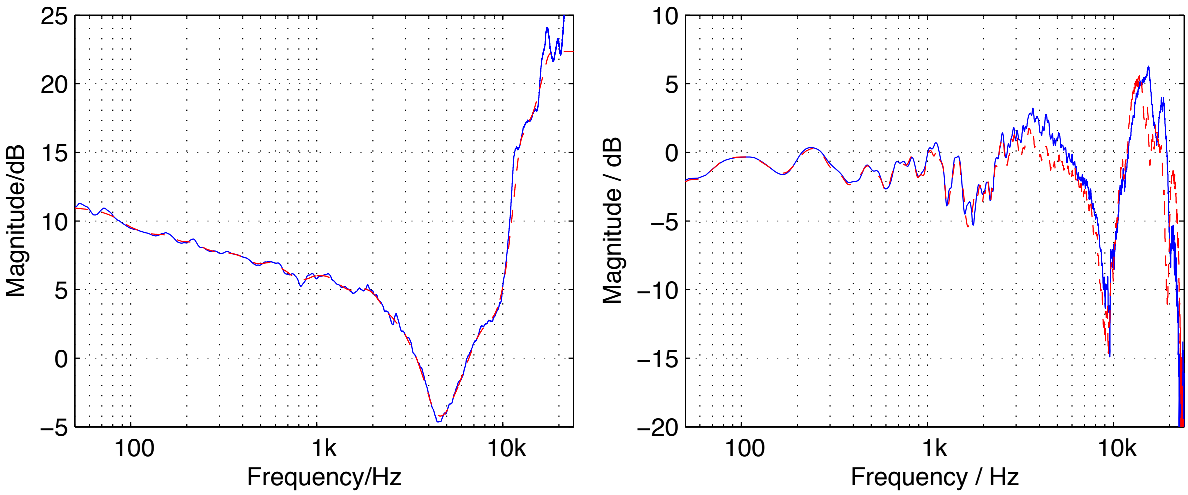

Figure 9: Left panel: Inverse diffuse field response of the dummy head (solid line) and the designed diffuse field equalization filter (dashed line). Right panel: Left (solid line) and right (dashed line) frontal HRTFs of the dummy head after diffuse field equalization.

give correct spatial reproduction, the reason for the coloration is obvious. The sound that already includes the effect of the auditory system up to the ear drum, travels a second time through a similar path in the peripheral auditory system of the listener. With headphones, the transmission path is different but a similar coloration may still be created due to typical design of headphones. Since commercial recordings are typically designed for loudspeaker reproduction, headphones are usually equalized to produce similar frequency characteristics to the ear drums of the listener as a pair of loudspeakers in a room. This equalization needs to be compensated. The best solution would be to design compensation filters for the exact frequency response of a chosen headphone model. However, a general compensation can be found by looking into the design criteria of headphones. A common reference for the design is to create the frequency characteristics of a sound source in an ideally diffuse sound field (diffuse field calibration, e.g. Theile, 1986). Now, if the binaural responses are equalized to yield a flat frequency response in a diffuse field (diffuse field equalization), the diffuse field frequency response is left to be created by the headphones, or in another way of thinking the "ideal" response of the headphones is compensated.

In order to design the diffuse field equalization, the head-related transfer functions (HRTFs) of the utilized Brüel and Kjær HATS dummy head were measured in an anechoic chamber. Measurements were performed at 7 different elevation angles at 10° steps in the azimuth direction, as well as from directly above the dummy head. A weighted root-mean-square (RMS) magnitude response of the resulting 253 transfer functions (assuming symmetry of the head) was calculated to yield an estimate of the diffuse field response of the dummy head. The inverse of the resulting response was smoothed with 1/3 octave resolution, and a minimum phase FIR equalization filter was designed. The magnitude spectra of the inverse (non-smoothed) diffuse field response and of the designed filter are shown in Fig. 9. The lower cutoff frequency of the measurement system was limited to 50 Hz, so below this frequency the filter is set to have a flat frequency response. Fig. 9 also illustrates the magnitude spectra of the frontal HRTFs after the diffuse field equalization. As can be seen, the equalization actually flattens the spectra. The level calibration followed the convention described in Section 4.3 such that the A-weighted power of the frontal HRTFs is preserved. Note that this will still in many cases result in different subjective loudness when listening to the equalized and non-equalized responses.

The database includes both diffuse field equalized and non-equalized binaural responses. As discussed, the equalized versions are recommended for general headphone listening, wheras the non-equalized responses are suitable for analysis purposes or for listening with separate headphone equalization.

5.3 Spatial Impulse Response Rendering

From the B-format SoundField microphone measurements it is possible to form responses for multichannel loudspeaker reproduction using several different techniques. Ambisonics (Gerzon, 1973) is one well-known systematic technique for this purpose. Several commercial and freeware Ambisonics decoders are available and an interested user is pointed to `http://www.ambisonic.net/`. Instead of Ambisonics, we provide loudspeaker responses created with a recent method called Spatial Impulse Response Rendering (SIRR) (Merimaa and Pulkki, 2003, 2004, 2005; Pulkki et al., 2004a,b; Pulkki and Merimaa, 2005), which has been shown to yield superior results in reproduction of room responses.

SIRR is an analysis-synthesis approach, where the direction of arrival and diffuseness of sound are analyzed as a function of time and frequency. Based on the analysis data and an omnidirectional measurement, responses for each loudspeaker in the reproduction system are synthesized. The synthesis utilizes two different techniques: nondiffuse part of the energy of each time-frequency component is positioned as sharply as possible in the correct direction with amplitude panning, wheras diffuse part is applied to all loudspeakers in a decorrelated form. This gives the advantage that discrete sharply localizable reflections can be synthesized with a precision that is only limited by the loudspeaker setup instead of microphone technology. Furthermore, the diffuse synthesis preserves the low interaural cross-correlation of the late reverberation and avoids comb-filtering created by several loudspeakers emitting the same signal in phase due to cross-talk.

SIRR is based on energetic analysis of a sound field, and the analysis data are very similar to the directional plots shown in the accompanying document (Merimaa et al., 2005). The responses included in the database were computed for a standard 5.0 loudspeaker setup (ITU-R BS.775-1, 1994) as described in (Merimaa and Pulkki, 2005; Pulkki and Merimaa, 2005). The processing was implemented using short-time Fourier transform (STFT) with 256 sample Hann windowed time frames zero-padded with another 256 samples. Directional synthesis of sharply localizable time-frequency components, as well as diffuse synthesis at frequencies below 1000 Hz was performed using the Vector Base Amplitude Panning (VBAP) algorithm (Pulkki, 1997), wheras diffuse synthesis above 1000 Hz was realized with the phase randomization method.

6 Summary

The measurement and processing of a set of room impulse responses from the 700-seat Pori Promenadikeskus concert hall was described. The hall was excited with an omnidirectional loudspeaker and a subwoofer using a sinusoidal sweep signal. The choice of microphones in the published database includes a pair of omnidirectional microphones, a stereo cardioid microphone, a dummy head, and a SoundField MKV microphone system. The responses were carefully system compensated and denoised. Furthermore, diffuse field equalized dummy head responses and SIRR-processed 5.0 loudspeaker responses are provided. More detailed analysis of each source-receiver combination is presented in the accompanying document Merimaa et al. (2005).

Acknowledgements

The responses were measured as part of the TAKU project (Control of Closed Space Acoustics) in the VÄRE technology program funded by the Finnish National Technology Agency Tekes. The authors would like thank Nokia Research Center and Nick Zacharov for lending us the SoundField microphone. A large part of the processing of the responses was done while Juha Merimaa was working for the Hearing Organisation And Recognition of Speech in Europe (HOARSE, HPRN-CP-2002-00276) project at Institute of Communication Acoustics, Ruhr-Universität Bochum, Germany.

References

A. Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *AES 108th Convention*, Paris, France, 2000. Preprint 5093.

A. C. Gade. Acoustical survey of eleven European concert halls — A basis for discussion of halls in Denmark. Technical Report 44, The Acoustics Laboratory, Technical University of Denmark, 1989.

M. A. Gerzon. Periphony: With-height sound reproduction. *J. Audio Eng. Soc.*, 21(1):2–10, 1973.

ISO 3382. Acoustics — Measurement of the reverberation time of rooms with reference to other acoustical parameters. International Standards Organization, 1997.

ITU-R BS.775-1. Multichannel stereophonic sound system with and without accompanying picture. International Telecommunication Union Radiocomminication Assembly, 1994.

J.-M. Jot, L. Cerveau, and O. Warusfel. Analysis and synthesis of room reverberation based on a statistical time-frequency model. In *AES 103rd Convention*, New York, NY, USA, 1997. Preprint 4629.

J. Merimaa, T. Peltonen, and T. Lokki. Concert hall impulse responses — Pori, Finland: Analysis results, 2005. Available at `http://www.acoustics.hut.fi/projects/poririrs/`.

J. Merimaa and V. Pulkki. Perceptually-based processing of directional room responses for multichannel loudspeaker reproduction. In *Proc. IEEE Workshop on Appl. of Sig. Proc. to Audio and Acoust.*, pages 51–54, New Paltz, NY, USA, 2003.

J. Merimaa and V. Pulkki. Spatial Impulse Response Rendering. In *Proc. 7th International Conference on Digital Audio Effects*, pages 139–144, Naples, Italy, 2004.

J. Merimaa and V. Pulkki. Spatial Impulse Response Rendering I: Analysis and synthesis. *J. Audio Eng. Soc.*, 2005. Submitted in February.

S. Müller and P. Massarani. Transfer-function measurement with sweeps. *J. Audio Eng. Soc.*, 49 (6):443–471, 2001.

T. Peltonen. A multichannel measurement system for room acoustics analysis. Master's thesis, Helsinki University of Technology, Department of Electrical and Communications Engineering, Laboratory of Acoustics and Audio Signal Processing, 2000.

T. Peltonen, T. Lokki, B. Goutarbès, J. Merimaa, and M. Karjalainen. A system for multi-channel and binaural room response measurements. In *AES 110th Convention*, Amsterdam, The Netherlands, 2001. Preprint 5289.

V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, 1997.

V. Pulkki and J. Merimaa. Spatial Impulse Response Rendering: Listening tests and applications to continuous sound. In *AES 118th Convention*, Barcelona, Spain, 2005.

V. Pulkki, J. Merimaa, and T. Lokki. Multi-channel reproduction of measured room responses. In *Proc. 18th International Congress on Acoustics*, pages 1273–1276, Kyoto, Japan, 2004a.

V. Pulkki, J. Merimaa, and T. Lokki. Reproduction of reverberation with Spatial Impulse Response Rendering. In *AES 116th Convention*, Berlin, Germany, 2004b. Preprint 270.

M. R. Schroeder. New method for measuring reverberation time. *J. Acoust. Soc. Am.*, 37(3): 409–412, 1965.

M. R. Schroeder. Integrated-impulse method measuring sound decay without using impulses. *J. Acoust. Soc. Am.*, 66(2):497–500, 1979.

G.-B. Stan, J.-J. Embrechts, and D. Archambeau. Comparison of different impulse response measurement techniques. *J. Audio Eng. Soc.*, 50(4):249–262, 2002.

G. Theile. On the standardization of the frequency response of high-quality studio headphones. *J. Audio Eng. Soc.*, 34(12):956–969, 1986.

APPENDICES

A List of files

The measurements with different microphones have been packed into individual zip files. Furthermore, measurements with receivers on the stage are available in separate packages. The full database includes the files listed below. In addition to the responses, each zip file also includes the readme.txt. The files can be downloaded from `http://www.acoustics.hut.fi/projects/poririrs/`

File	**Contents**	**Channels**	**Src**	**Rcv**	**Microphone**
binaural.zip	s1_r1_b.wav	L,R	S1	R1	HATS
	s1_r2_b.wav	L,R	S1	R2	HATS
	s1_r3_b.wav	L,R	S1	R3	HATS
	s1_r4_b.wav	L,R	S1	R4	HATS
	s2_r1_b.wav	L,R	S2	R1	HATS
	s2_r2_b.wav	L,R	S2	R2	HATS
	s2_r3_b.wav	L,R	S2	R3	HATS
	s2_r4_b.wav	L,R	S2	R4	HATS
	s3_r1_b.wav	L,R	S3	R1	HATS
	s3_r2_b.wav	L,R	S3	R2	HATS
	s3_r3_b.wav	L,R	S3	R3	HATS
	s3_r4_b.wav	L,R	S3	R4	HATS
bin_dfeq.zip	s1_r1_bd.wav	L,R	S1	R1	HATS, d.f. EQ
	s1_r2_bd.wav	L,R	S1	R2	HATS, d.f. EQ
	s1_r3_bd.wav	L,R	S1	R3	HATS, d.f. EQ
	s1_r4_bd.wav	L,R	S1	R4	HATS, d.f. EQ
	s2_r1_bd.wav	L,R	S2	R1	HATS, d.f. EQ
	s2_r2_bd.wav	L,R	S2	R2	HATS, d.f. EQ
	s2_r3_bd.wav	L,R	S2	R3	HATS, d.f. EQ
	s2_r4_bd.wav	L,R	S2	R4	HATS, d.f. EQ
	s3_r1_bd.wav	L,R	S3	R1	HATS, d.f. EQ
	s3_r2_bd.wav	L,R	S3	R2	HATS, d.f. EQ
	s3_r3_bd.wav	L,R	S3	R3	HATS, d.f. EQ
	s3_r4_bd.wav	L,R	S3	R4	HATS, d.f. EQ
cardioid.zip	s1_r1_c.wav	L,R	S1	R1	Pearl
	s1_r2_c.wav	L,R	S1	R2	Pearl
	s1_r3_c.wav	L,R	S1	R3	Pearl
	s1_r4_c.wav	L,R	S1	R4	Pearl
	s2_r1_c.wav	L,R	S2	R1	Pearl
	s2_r2_c.wav	L,R	S2	R2	Pearl
	s2_r3_c.wav	L,R	S2	R3	Pearl
	s2_r4_c.wav	L,R	S2	R4	Pearl
	s3_r1_c.wav	L,R	S3	R1	Pearl
	s3_r2_c.wav	L,R	S3	R2	Pearl
	s3_r3_c.wav	L,R	S3	R3	Pearl
	s3_r4_c.wav	L,R	S3	R4	Pearl

File	Contents	Channels	Src	Rcv	Microphone
omni.zip	s1_r1_o.wav	L,R	S1	R1	DPA 4006
	s1_r2_o.wav	L,R	S1	R2	DPA 4006
	s1_r3_o.wav	L,R	S1	R3	DPA 4006
	s1_r4_o.wav	L,R	S1	R4	DPA 4006
	s2_r1_o.wav	L,R	S2	R1	DPA 4006
	s2_r2_o.wav	L,R	S2	R2	DPA 4006
	s2_r3_o.wav	L,R	S2	R3	DPA 4006
	s2_r4_o.wav	L,R	S2	R4	DPA 4006
	s3_r1_o.wav	L,R	S3	R1	DPA 4006
	s3_r2_o.wav	L,R	S3	R2	DPA 4006
	s3_r3_o.wav	L,R	S3	R3	DPA 4006
	s3_r4_o.wav	L,R	S3	R4	DPA 4006
omni_p.zip	s1_p1_o.wav	L,R	S1	P1	DPA 4006
	s1_p2_o.wav	L,R	S1	P2	DPA 4006
	s1_p3_o.wav	L,R	S1	P3	DPA 4006
	s2_p1_o.wav	L,R	S2	P1	DPA 4006
	s2_p2_o.wav	L,R	S2	P2	DPA 4006
	s2_p3_o.wav	L,R	S2	P3	DPA 4006
	s3_p1_o.wav	L,R	S3	P1	DPA 4006
	s3_p2_o.wav	L,R	S3	P2	DPA 4006
	s3_p3_o.wav	L,R	S3	P3	DPA 4006
porifigs.pdf	The accompanying document (Merimaa et al., 2005)				
poriref.pdf	This document				
readme.txt	Short summary of the files available for downloading				
sirr.zip	s1_r1_sr.wav	L,R,C,LS,RS	S1	R1	SIRR
	s1_r2_sr.wav	L,R,C,LS,RS	S1	R2	SIRR
	s1_r3_sr.wav	L,R,C,LS,RS	S1	R3	SIRR
	s1_r4_sr.wav	L,R,C,LS,RS	S1	R4	SIRR
	s2_r1_sr.wav	L,R,C,LS,RS	S2	R1	SIRR
	s2_r2_sr.wav	L,R,C,LS,RS	S2	R2	SIRR
	s2_r3_sr.wav	L,R,C,LS,RS	S2	R3	SIRR
	s2_r4_sr.wav	L,R,C,LS,RS	S2	R4	SIRR
	s3_r1_sr.wav	L,R,C,LS,RS	S3	R1	SIRR
	s3_r2_sr.wav	L,R,C,LS,RS	S3	R2	SIRR
	s3_r3_sr.wav	L,R,C,LS,RS	S3	R3	SIRR
	s3_r4_sr.wav	L,R,C,LS,RS	S3	R4	SIRR
sndfld.zip	s1_r1_sf.wav	W,X,Y,Z	S1	R1	SoundField
	s1_r2_sf.wav	W,X,Y,Z	S1	R2	SoundField
	s1_r3_sf.wav	W,X,Y,Z	S1	R3	SoundField
	s1_r4_sf.wav	W,X,Y,Z	S1	R4	SoundField
	s2_r1_sf.wav	W,X,Y,Z	S2	R1	SoundField
	s2_r2_sf.wav	W,X,Y,Z	S2	R2	SoundField
	s2_r3_sf.wav	W,X,Y,Z	S2	R3	SoundField
	s2_r4_sf.wav	W,X,Y,Z	S2	R4	SoundField
	s3_r1_sf.wav	W,X,Y,Z	S3	R1	SoundField
	s3_r2_sf.wav	W,X,Y,Z	S3	R2	SoundField
	s3_r3_sf.wav	W,X,Y,Z	S3	R3	SoundField
	s3_r4_sf.wav	W,X,Y,Z	S3	R4	SoundField

File	Contents	Channels	Src	Rcv	Microphone
sndfld_p.zip	s1_r1_sf.wav	W,X,Y,Z	S1	P1	SoundField
	s1_r2_sf.wav	W,X,Y,Z	S1	P2	SoundField
	s1_r3_sf.wav	W,X,Y,Z	S1	P3	SoundField
	s2_r1_sf.wav	W,X,Y,Z	S2	P1	SoundField
	s2_r2_sf.wav	W,X,Y,Z	S2	P2	SoundField
	s2_r3_sf.wav	W,X,Y,Z	S2	P3	SoundField
	s3_r1_sf.wav	W,X,Y,Z	S3	P1	SoundField
	s3_r2_sf.wav	W,X,Y,Z	S3	P2	SoundField
	s3_r3_sf.wav	W,X,Y,Z	S3	P3	SoundField

All responses are saved as 24 bit wav files with a sampling rate of 48 kHz. For a description of the source and receiver position see Section 2 (Table 2).

A.1 Microphone and channel specifications

The microphone directions are specified in a Cartesian coordinate system. The positive X axis (front) for each receiver position in the audience area is defined as the direction towards the stage perpendicular to the backs of the seats in the corresponding position. For the receiver positions on the stage, the front is defined as the direction perpendicular to and pointing towards the front wall behind the stage. Positive Y and Z directions are defined as left and up, respectively, when facing the front.

DPA 4006
: Pair of DPA Type 4006 omnidirectional microphones facing the front with a distance of 10 cm between them in the left-right direction. L = leftmost microphone, R = rightmost microphone.

HATS
: Brüel and Kjær HATS dummy head custom fitted with DPA Type 4053 microphones, facing the front. L = left ear input, R = right ear input.

HATS, d.f. EQ
: Same as HATS, but diffuse field equalized (see Section 5.2).

Pearl
: Pearl TL-4 stereo condenser microphone with cardioid directivity patterns. L = cardioid facing to the left, R = cardioid facing to the right.

SIRR
: SoundField B-format signals processed with the SIRR method (see Section 5.3) for reproduction with a standard 5.0 loudspeaker setup (ITU-R BS.775-1, 1994). Loudspeakers for the corresponding channels should be located in the horizontal plane equidistant from a listener with the following azimuthal angles: left (L): 30°, right (R): −30°, center (C): 0°, left surround (LS): 110°, and right surround (RS): −110°.

SoundField
: B-format signals from a SoundField MKV microphone system. W is an omnidirectional signal and X, Y, and Z have figure-of-eight directivity patterns with positive directions aligned with the corresponding positive coordinate axes.

B Copyright