Simulative investigation of required spatial source resolution in directional room impulse response measurements

Johannes Klein, Michael Vorländer

SIMULATIVE INVESTIGATION OF REQUIRED SPATIAL SOURCE RESOLUTION IN DIRECTIONAL ROOM IMPULSE RESPONSE MEASUREMENTS

Johannes Klein, Michael Vorländer
Institute of Technical Acoustics, RWTH Aachen University, Aachen, Germany

ABSTRACT

In classical and standardized room acoustic measurements, the directivity of sources is either not considered, or specified to be omni-directional for inter-measurement comparability. These room impulse response measurements allow for the analysis of certain aspects of an acoustic scene. Parameters such as Early Decay Time attempt to reflect the listening impression, while others such as the reverberation time serve as physical descriptors and for the evaluation of rooms for different purposes.

The explicit non-consideration of the source and receiver directivity waives an abundance of information, which is essential for a complete description of the spatial composition of an acoustic scene, e.g. for an auralization. On the one hand, a more complete description raises the degree of realism in auralizations, presumably intensifying the immersion of human listeners. On the other hand, the freedom of arbitrary source-receiver directivity combinations opens up the possibility of directional acoustical analysis like scanning single reflection paths within a room. For the measurement of directional room impulse responses several measurement methods and instruments have been implemented in the past. Most of them aim either at fast measurements or at measurements of a high spatial resolution. The compromise of obtaining a sufficiently high resolution in an acceptable time usually is often disregarded due to the missing definition of a sufficiently high resolution.

Before the development of suitable methods and instruments, the importance of directivity in an acoustical scene has to be determined. This entails the question if and up to which complexity the measurement of directional room impulse responses offers an advantage for room representations in auralizations and for parametric room acoustic descriptions. A sufficiently high resolution can be defined in many ways, either in regard to just noticeable differences for human listeners, or objectively regarding the effects on technical parameters. This work investigates the effect of the source directivity resolution on room acoustical parameters. Since real measurements contain too many uncontrollable influences such as time variances and sources that can radiate the required directivity with a sufficient precision do not yet exist, this investigation is done using room acoustic simulations.

First, the modeling of a suitable artificial directivity will be explained. The spatial resolution will be denoted by the corresponding spherical harmonic order. The goal is a directivity with a minimum beam width for each given spherical harmonic order without strong side lobes. This characteristic represents the worst case for each resolution. The generated source directivity will then be used in hybrid ray tracing and image source room acoustic simulations in two rooms of different size and acoustic property. This approach allows a more generally valid statement about the impact of the source directivity. The results will be discussed using the impact on objective room acoustic parameters as an indicator for the required spatial source resolution. Subjective parameters will be considered as an outlook toward the impact on the human perception. The findings are meant to aid the design of measurement instruments for directional room impulse response measurement in reasonable measurement times with a sufficiently high spatial resolution.

1. INTRODUCTION

The measurement of *Room Impulse Responses* (RIRs) with arbitrary source directivity promises an advantage for room representations in auralizations and for parametric room acoustic descriptions. However, before the development of suitable measurement methods and instruments for such *Directional Room Impulse Response* (DRIR) measurements, it has to be investigated if and up to which complexity [1] they are required.

This study compares the results of simulations in several acoustic scenes with the complexity of the source directivity as the variable. The definition of complexity is ambiguous, hence the worst case for each level of complexity is defined and used in the simulations. The resulting DRIRs are analyzed regarding their temporal structure and resulting room acoustic parameters to evaluate a threshold for the required source directivity complexity either in each room or over a wide range of rooms.

[1] *complexity* refers to the spatial detail and resolution of a directivity.

doi:10.25836/sasp.2019.24

2. FUNDAMENTALS

Some definitions and fundamentals are required to facilitate the understanding of the investigation.

2.1 Spherical Harmonic Base

The complex *Spherical Harmonic* (SH) base functions can be used to decompose any azimuth (φ) and elevation (ϑ) angle-dependent spatial function $f(\vartheta, \varphi)$ on the unit sphere into its fundamentals. This study uses the complex SH base functions $Y_n^m(\vartheta, \varphi)$ containing the associated Legendre functions P_n^m [1] defined as [2] [2]

$$Y_n^m(\vartheta, \varphi) = \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}} \cdot P_n^m(\cos(\vartheta)) \cdot \mathrm{e}^{\mathrm{j}m\varphi}. \tag{1}$$

2.2 Spherical Wave Spectrum

The transformation of a spatial function into its fundamentals results in a coefficient function $\hat{f}_{nm}$, denoting the share of each fundamental in *order* (n) and *degree* (m). This coefficient function is called *Spherical Wave Spectrum* (SWS). The *Spherical Harmonic Transform* (SHT) is defined as [2]

$$\hat{f}_{nm} = \mathcal{S}\{f(\vartheta, \varphi)\} = \oint_{S^2} f(\vartheta, \varphi) \cdot \overline{Y_n^m(\vartheta, \varphi)} \mathrm{d}\Omega. \tag{2}$$

The operation yields a precise coefficient function, if the integral can be solved in an exact way. The *Inverse Spherical Harmonic Transform* (ISHT) is defined as [2]

$$f(\vartheta, \varphi) = \mathcal{S}^{-1}\{\hat{f}_{nm}\} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \hat{f}_{nm} \cdot Y_n^m(\vartheta, \varphi). \tag{3}$$

A finite summation limit can be applied (*truncation*), resulting in a less precise spatial function.

2.3 Dirac on a Sphere

The completeness relation for the SH is using the Kronecker delta δ as [2]

$$\sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_n^m(\vartheta, \varphi) \overline{Y_n^m(\vartheta', \varphi')} = \delta^{(\vartheta', \varphi')}(\vartheta, \varphi). \tag{4}$$

Analogous to the Dirac impulse in the time domain, the function differs from zero only in the one case of $\vartheta = \vartheta', \varphi = \varphi'$. Comparing Eq. (4) with Eq. (3) yields the SWS coefficients for the spatial Dirac on the unit sphere

$$\delta_{nm}^{(\vartheta', \varphi')} = \overline{Y_n^m(\vartheta', \varphi')}. \tag{5}$$

The spatial Dirac integrates to unity [2], and the sifting property holds [2]. The spatial Dirac is the "1"-element of the spatial convolution, which for SWS is defined as [4]

$$\mathcal{S}\{f(\Omega) * g(\Omega)\} = 2\pi \sqrt{\frac{4\pi}{2n+1}} \hat{f}_{nm} \hat{g}_{n0}. \tag{6}$$

with the coefficients $\hat{g}_{n0}$ only defined for a degree $m = 0$.

[2] This definition omits the explicit mention of the Condon-Shortley phase $(-1)^m$, due to its inclusion in the associated Legendre function [3].

2.4 Exterior Problems

In exterior problems all sources are confined to a volume with a radius r_0. The exterior is source-free, and the Sommerfeld radiation condition is met. With the Hankel function of the second kind h_n, the extrapolation of SWS coefficients $\hat{c}_{nm}$ from r_0 to a radius r is defined as [2]

$$\hat{c}_{nm}(kr) = \hat{c}_{nm}(kr_0) \frac{h_n(kr)}{h_n(kr_0)}. \tag{7}$$

2.5 Order-Far-Field

The extrapolation frequency-dependently annihilates high orders. While passing through the near-field, this effect is stronger for coefficients for high orders at low frequencies compared to coefficients of low orders, distorting the directivity. Starting at an *order-far-field* distance, these order-relative differences in attenuation become negligible.

2.6 Order Truncation

Order truncation in the SWS domain introduces artifacts. Their magnitude is determined by the highest retained order and the truncation method. Predictors of substantial orders for a specific combination of wave number k and source radius r can be found in [5–7], the most common being the *kr-limit* [8]

$$n_{\mathrm{trunc}} = \lfloor kr \rfloor. \tag{8}$$

The coefficients can be cut-off (*rect* function) or faded out using for example a Hann function [9].

2.7 Vibrating Polar Cap

The vibrating polar cap describes a uniform surface vibration of a north-pole sphere section limited by the elevation aperture angle α. The SWS coefficients for the vibration are [6]

$$\hat{u}_{nm} = \begin{cases} \sqrt{\pi}(1-\beta), & n = 0, \\ \delta_{m0} \sqrt{\frac{\pi}{2n+1}} \left[P_{n-1}(\beta) - P_{n+1}(\beta) \right], & n > 0, \end{cases} \tag{9}$$

with $\beta = \cos(\alpha)$ and the Legendre functions P_{n-1} and P_{n+1}. The acoustic radiation impedance with the speed of sound c and the density ρ_0 connects the vibration coefficients and the sound pressure at the boundary surface [6]

$$\hat{p}_{nm}(r_0) = -\mathrm{j}\rho_0 c \frac{h_n(kr_0)}{h_n'(kr_0)} \hat{u}_{nm}(r_0). \tag{10}$$

The coefficients $\hat{p}_{nm}$ can be rotated to any (ϑ', φ') using the spherical convolution and the Dirac given in Eqs. (5) and (6). The coefficients can be extrapolated to larger radii using Eq. (7). The Hankel functions introduce a frequency-dependent order distribution, further diminishing coefficient values for high orders in low frequencies.

doi:10.25836/sasp.2019.24

3. ARTIFICIAL SOURCE DIRECTIVITY GENERATION

The source directivity needs to be scalable and represent the worst case for each level of complexity. Here, the worst case is defined as directivity with a frequency dependent minimum-width main-lobe. The directivity is artificially generated in the SWS domain to ensure a consistent worst case directivity. Two methods for generating such a directivity are discussed in this section.

3.1 Acoustic Beam

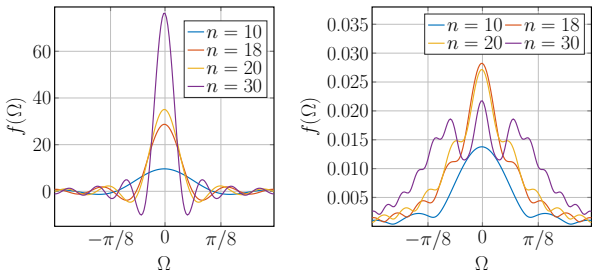
The acoustic beam is based on the spatial Dirac. It can be freely positioned using the spherical convolution in Eq. (6). It can simply be extrapolated using Eq. (7) (cf. Fig. 1).

3.2 Spherical Cap

The spherical cap is widely used as a model for transducers on a spherical array [2, 6, 10, 11]. In its original form it describes the same actively vibrating surface at all frequencies, contradicting the actual behavior of loudspeakers [12, 13] and the worst case defined above.

3.3 Order Limitation

The orders allowed by the kr-limit still contain a sub-range of high orders that is deformed significantly stronger on the way to the order-far-field than the lower orders. The amplitude of the Dirac functions on the source surface in Fig. 1a rises and the beam-width narrows monotonously with the truncation order, this is not true after the extrapolation in Fig. 1b. The peak value of the Dirac with a truncation order of 20 is lower than that of order 18, while side-lobes start to rise. At a truncation order of 30 the side-lobe amplitude almost reaches that of the further attenuated main-lobe.



(a) Source surface. **(b)** Extrapolation.

Figure 1: Amplitude of n-truncated Dirac functions over angular distance to the center-axis direction. Extrapolation $0.15\text{m} \rightarrow 100\text{m}$ for $f = 22.05\text{kHz}$.

Using the first zero-crossings of the first derivative of the Dirac peak value over the truncation order, a stricter *kr-root-limit* limit can experimentally be determined. This prevents the unwanted deformation of the order-far-field

$$n_{\sqrt{kr}_{1,s,e}} = \begin{cases} \lfloor s \cdot \sqrt{k \cdot r_s} - 1 \rfloor, & k \cdot r_s \geq \sqrt{\frac{1}{s}} \\ e, & k \cdot r_s < \sqrt{\frac{1}{s}}. \end{cases} \quad (11)$$

4. SIMULATION SETTINGS AND PARAMETERS

This section presents the methodology, the specific source directivity, and the room models for the simulations.

4.1 Simulation Methods

A hybrid model of ray tracing and image sources, implemented in the room acoustic simulation software RAVEN [14, 15] is used for all simulations. The image source order is set to 2 and $1 \cdot 10^6$ ray tracing particles are deployed. The ray tracing reflection pattern and the impulse response generation Poisson sequence are fixed. The directivity is defined as magnitude spectra at 65160 points of a regular horizontal and vertical 1° sampling scheme, allowing for the representation of an order of 50. The monaural receiver directivity is chosen to be omnidirectional. The binaural simulations use the head-related transfer-functions of the ITA artificial head [16] and at the same receiver center positions.

The surface material information and the geometrical models for the simulation are taken from the Benchmark for Room Acoustical Simulation (BRAS) database [17]. The material coefficients are fitted[3] for all models to match the measured mean T30 values in the BRAS database at the ANSI center frequencies [18] using the ITA-Toolbox[4] [19] and RAVEN. Below 50Hz the absorption values are spline-interpolated.

4.2 Simulated Sources

An acoustic beam source with a source radius $r_\text{s} = 0.15\text{m}$, limited to $\sqrt{kr}_{1,\frac{64}{25},1}$, smoothed with a Hann function and extrapolated to $r = 100\text{m}$ is used for the simulations. Only the absolute value is transferred to the simulation. Simulations are done for a main-lobe alignment along all 6 spatial axes. Fig. 2 shows the directivity for selected frequencies.

4.3 Simulation Rooms

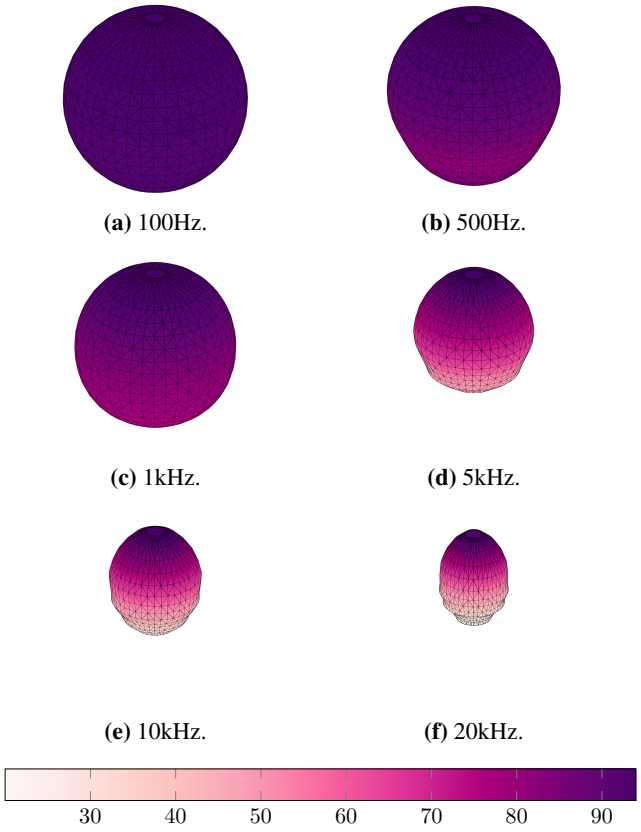
Two different rooms from the BRAS database are used as simulation models. The simulation and measurement data is obtained for all permutations of 2 loudspeaker and 5 receiver positions. The measured room characteristics from the BRAS database are briefly presented in this section.

4.3.1 Medium Concert Hall (Scene 10)

The medium concert hall has a surface area of 2720m^2 and a volume of 3319m^3. It is designed for classical music concerts. The mean T30 is 1.184s, resulting in a Schroeder frequency of 37.77Hz. The reverberation time is very constant for the frequency range up to 1kHz and decays to higher frequencies. The T30 reverberation times for all positions show a mean standard deviation of $s = 0.05\text{s}$ between 100Hz and 10kHz. The mean broadband EDT over all positions is 0.8545s. The energy ratios indicate a high clarity and depend strongly on the measurement position.

[3] The starting values for the fitting process are the *fitted estimates* included in [17].

[4] Git commit SHA dffe6d84524aca91cec8bcbe55f582fa8bb421e8.



(a) 100Hz. **(b)** 500Hz. **(c)** 1kHz. **(d)** 5kHz. **(e)** 10kHz. **(f)** 20kHz.

Figure 2: Balloon plot of the absolute values of acoustic beam source directivity aligned with +z.

4.3.2 Large Auditorium (Scene 11)

The large auditorium has a $5811\mathrm{m}^2$ surface area and a volume of $8658\mathrm{m}^2$. It is designed for lectures. The mean T30 over all positions is 1.956s, resulting in a Schroeder frequency of 30.06Hz. The T30 decays over the frequency with a slight elevation around 200Hz. The T30 reverberation times for all positions show a mean standard deviation of $s = 0.064$s between 100Hz and 10kHz. The mean broadband EDT over all positions is 1.3367s. The energy ratios show a high variance.

5. RESULTS

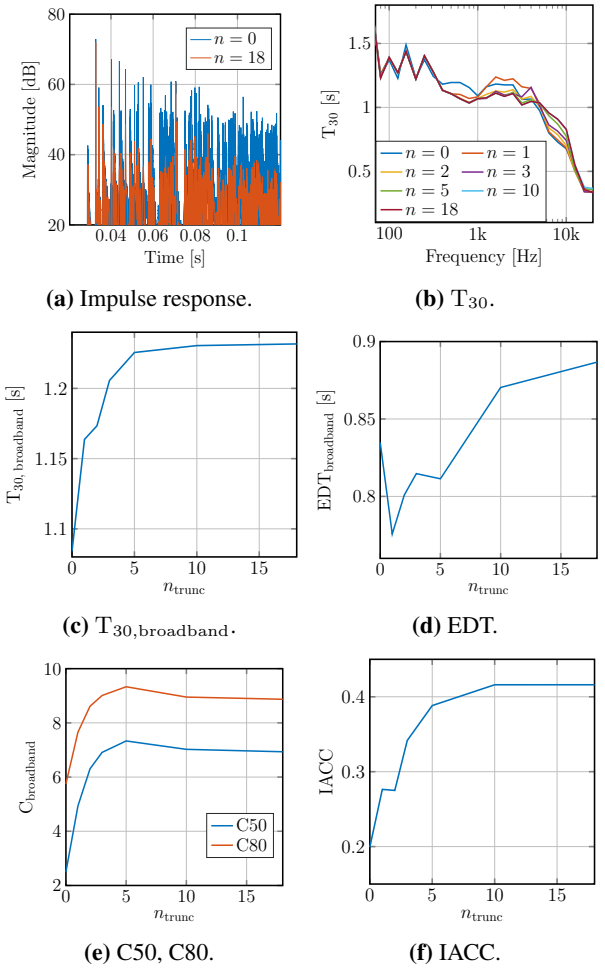
For reasons of brevity, only a small selection of the simulation results is presented. The measurement point numbering and axes correspond to the BRAS database.

5.1 Medium Concert Hall (Scene 10)

The results for simulations with a forward (+x) facing source at position 1 and a receiver at position 3 are shown in Fig. 3. Fig. 4 shows the results for a simulation with an upward (+z) facing source at the same position and a receiver at measurement position 2.

The RIR in Fig. 3a shows, that the measurement position 3 is still well in the beam of the forward facing source, while measurement position 2 is clearly not for the upward facing beam, as seen in Fig. 4a. Especially the behavior of the energy ratios in Figs. 3e and 4e is affected by this difference. While for measurement position 3 the clarity rises with the directivity due to the disappearance of room reflections, it falls with the order for position 2 as soon as it is not within the beam aperture.

The behavior of the other parameters is similar for both simulations, while the values differ strongly. The broadband T30 is prolonged by 0.13s or 12% for the first simulation and by 0.5 or 50% for the second (cf. Figs. 3c and 4c). The EDT rises about 0.1s or 12% for simulation 1 and by 0.5s or 50% for simulation 2 (cf. Figs. 3d and 4d). The *Inter-Aural Cross Correlation* (IACC) rises by roughly 0.23 for both simulations. The narrow-band T30s in Figs. 3b and 4b show some variance for frequencies up to 10kHz. Most curves reach a saturation around the truncation order 10, only the EDT changes up to the truncation order 18.



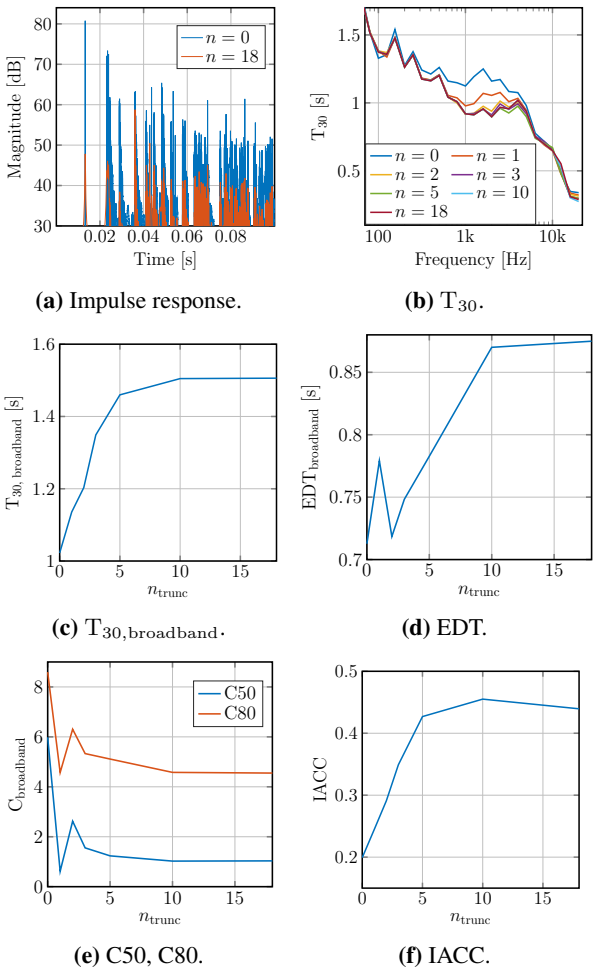
(a) Impulse response. **(b)** T_{30}. **(c)** $\mathrm{T}_{30,\mathrm{broadband}}$. **(d)** EDT. **(e)** C50, C80. **(f)** IACC.

Figure 3: Simulation results. Scene 10, source position 1, direction +x, measurement point 3.

5.2 Large Auditorium (Scene 11)

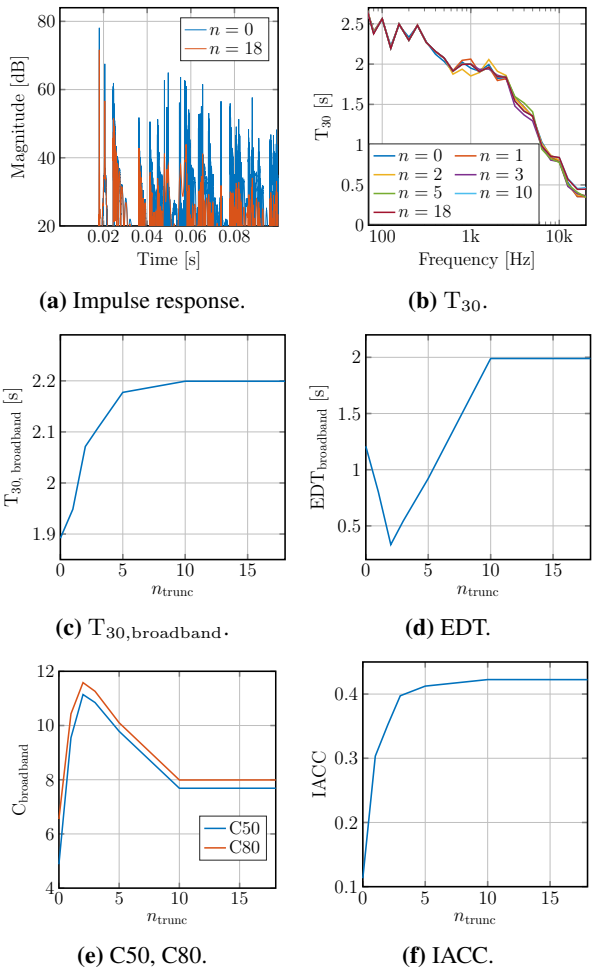
Fig. 5 shows the simulation results for the large auditorium with an acoustic beam source in forward (+x) orientation at source position 1, recorded with a receiver at position 4. Fig. 6 shows a backward (-x) oriented beam source at the same position, recorded by a receiver at position 3.

doi:10.25836/sasp.2019.24



(a) Impulse response. **(b)** T_{30}. **(c)** $T_{30,broadband}$. **(d)** EDT. **(e)** C50, C80. **(f)** IACC.

Figure 4: Simulation results. Scene 10, source position 1, direction +z, measurement point 2.



(a) Impulse response. **(b)** T_{30}. **(c)** $T_{30,broadband}$. **(d)** EDT. **(e)** C50, C80. **(f)** IACC.

Figure 5: Simulation results. Scene 11, source position 1, direction +x, measurement point 4.

While position 4 is still well in the forward beam (cf. Fig. 5a), the direct sound is nearly missing at position 3 for the backward beam, as seen in Fig. 6a. This is also without any doubt an audible effect.

Both T30 simulations in Figs. 5b and 6b are again only slightly affected. For the simulation with the forward facing source, the T30 rises by 0.3s (16%, Fig. 5c), the EDT by 1.5s (300%, Fig. 5d), and the IACC by 0.31. For the simulations with the backward facing source, the T30 rises by 0.3s (15%, Fig. 6c), the EDT by 0.6s (38%, Fig. 6d), and the IACC by 0.4.

For the measurement point in the source beam, the energy ratios rise as long the point is within the aperture (cf. Fig. 5e), while any other truncation order than 0 for the backward facing source means a lesser clarity, as seen in Fig. 6f.

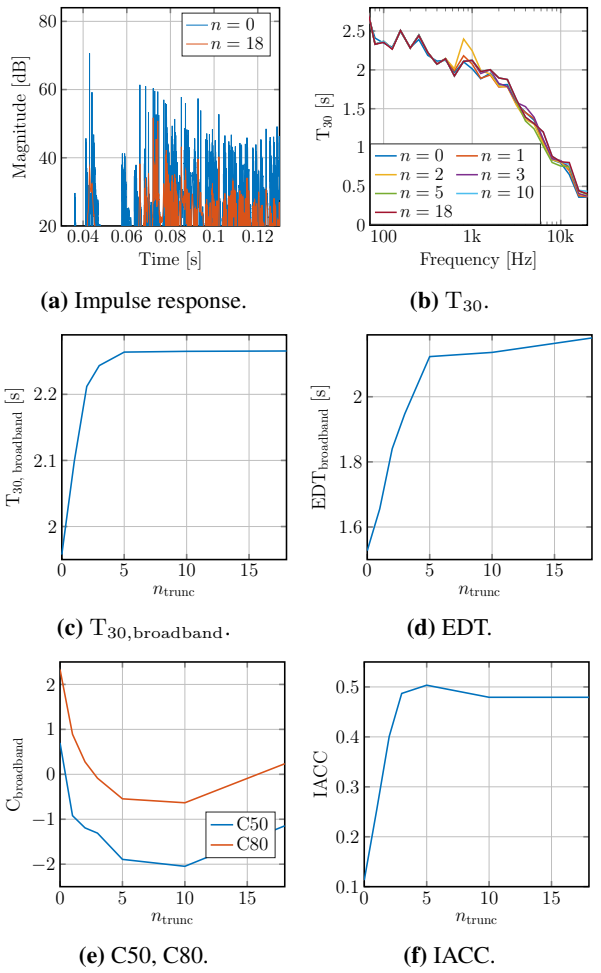
6. CONCLUSION

The simulations use the worst-case directivity for a common source with a radius $r_s = 0.15$m. The source radius dictates the respective truncation order at every specific frequency.

The simulations show a significant directivity impact on the RIR and some room acoustic parameters. The temporal structure of the RIR is affected by the absence of reflections and the direct sound level by the orientation of the source, which can be considered as easily audible. Especially the C50 and C80 energy ratios yield information about the relative positioning of the source beam and the receiver position. The IACC generally profits from a more diffuse field without many prominent reflections. The T reverberation times are not as affected by the directivity, since they do not regard the direct sound and the early reflections. The EDT rises by 50%-300% due to a changing source directivity. Since the EDT is considered a measure for the perception of reverberation, this is a clear indicator that the source directivity matters for RIR measurements and the perception of acoustic scenes.

In the simulations in can be seen that most parameters show an order-saturation, since the beam width changes minimally and so do the parameters. The effect is illustrated for the RIR of the simulation in the large auditorium, for the backward facing source and receiver position 3 in Fig. 7. The RIR correlation finds a saturation above a truncation order of 10. This should also be taken into account during the design of instruments for DRIR measurements.

doi:10.25836/sasp.2019.24



(a) Impulse response. **(b)** T_{30}.

(c) $T_{30,broadband}$. **(d)** EDT.

(e) C50, C80. **(f)** IACC.

Figure 6: Simulation results. Scene 11, source position 1, direction -x, measurement point 3.

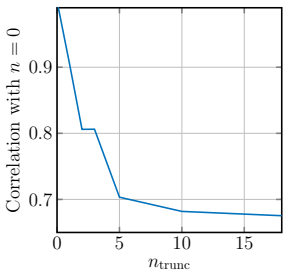


Figure 7: Scene 11. Correlation of the simulated RIRs for several orders with RIR for $n = 0$.

7. REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. Courier Corporation, 1964.

[2] E. G. Williams, *Fourier Acoustics*. Academic press, 1999.

[3] G. B. Arfken, H.-J. Weber, and F. E. Harris, *Mathematical Methods for Physicists*. Elsevier, 2013.

[4] J. R. Driscoll and D. M. Healy, "Computing Fourier Transforms and Convolutions on the 2-Sphere," *Advances in Applied Mathematics*, vol. 15, no. 2, 1994.

[5] N. A. Gumerov and R. Duraiswami, "Computation of Scattering from N Spheres Using Multipole Reexpansion," *J. Acoust. Soc. Am.*, vol. 112, no. 6, 2002.

[6] B. Rafaely, *Fundamentals of Spherical Array Processing*. Springer Berlin Heidelberg, 2015.

[7] M. Müller-Trapet, M. Pollow, and M. Vorländer, "Spherical Harmonics as a Basis for Quantifying Scattering and Diffusing Objects," in *Proc. Forum Acusticum*, (Aalborg, Denmark), 2011.

[8] R. Duraiswami, D. N. Zotkin, and N. A. Gumerov, "Interpolation and Range Extrapolation of HRTFs," in *Proc. IEEE ICASSP*, (Montreal, Canada), 2004.

[9] A. V. Oppenheim and R. W. Schafer, *Discrete-time signal processing*. Pearson, 2010.

[10] M. Pollow and G. K. Behler, "Variable Directivity for Platonic Sound Sources Based on Spherical Harmonics Optimization," *Acta Acustica united with Acustica*, vol. 95, no. 6, 2009.

[11] F. Zotter and R. Höldrich, "Modeling Radiation Synthesis with Spherical Loudspeaker Arrays," in *Proc. ICA*, (Madrid, Spain), 2007.

[12] F. J. M. Frankort, "Vibration Patterns and Radiation Behavior of Loudspeaker Cones," *J. Audio Eng. Soc*, vol. 26, no. 9, 1978.

[13] W. Klippel and J. Schlechter, "Distributed Mechanical Parameters of Loudspeakers, Part 1: Measurements," *J. Audio Eng. Soc*, vol. 57, no. 7/8, 2009.

[14] D. Schröder and M. Vorländer, "RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments," in *Proc. Forum Acusticum*, (Aalborg, Denmark), 2011.

[15] D. Schröder, *Physically Based Real-Time Auralization of Interactive Virtual Environments*. Logos Berlin, 2011.

[16] A. Schmitz, "Ein neues digitales Kunstkopfmeßsystem," *Acustica*, vol. 81, no. 4, 1995.

[17] L. Aspöck, F. Brinkmann, D. Ackermann, S. Weinzierl, and M. Vorländer, "BRAS - Benchmark for Room Acoustical Simulation," 2019. `http://dx.doi.org/10.14279/depositonce-6726.2`.

[18] American National Standards Institute, "ANSI S1.11-2004: Specification for Octave-Band and Fractional-Octave-Band Analog an Digital Filters," 2004.

[19] M. Berzborn, R. Bomhardt, J. Klein, J.-G. Richter, and M. Vorländer, "The ITA-Toolbox," in *Proc. DAGA*, (Kiel, Germany), 2017.

doi:10.25836/sasp.2019.24