

Spatial upsampling of sparse head-related transfer function sets by directional equalization - Influence of the spherical sampling scheme

Johannes M. AREND^{(1)(2)†}, Christoph PÖRSCHMANN⁽¹⁾

⁽¹⁾Institute of Communications Engineering, TH Köln, D-50679 Cologne, Germany,

⁽²⁾Audio Communication Group, TU Berlin, D-10587 Berlin, Germany

[†]Corresponding author, E-mail: johannes.arend@th-koeln.de

Abstract

Many immersive audio applications rely on a dense set of head-related transfer functions (HRTFs). However, often only measurements on a specific sparse grid are available. To obtain dense HRTF sets from sparse measurements, one common approach is to apply spatial interpolation in the spherical harmonics (SH) domain. However, the SH representation of sparse HRTF sets is order-limited, leading to spatial aliasing and truncation errors. In a recent publication, we presented the so-called SUPDEq method (Spatial Upsampling by Directional Equalization) for spatial upsampling of sparse HRTF sets. The approach is based on a directional equalization of the sparse set prior to the spherical Fourier transform to remove direction-dependent temporal and spectral components. This significantly reduces the spatial complexity of the sparse set, allowing for an enhanced interpolation at reduced SH orders. In this study we investigate how different spherical sampling schemes affect the performance of common SH interpolation and the SUPDEq method. For this, we compare spatially upsampled HRTF sets originally based on sparse equiangular, Gaussian, Lebedev, and Fliege grids at various SH orders to a reference. The influence of the different grids are assessed spectrally, temporally, and with localization models. Keywords: Head-Related Transfer Functions (HRTFs), Spatial Audio, Spherical Harmonics

1 INTRODUCTION

Human sound source localization is based on binaural cues, i.e. interaural time differences (ITDs) and interaural level differences (ILDs) between both ears, as well as on monaural cues, i.e. spectral distortions of the incoming sound caused mainly by the listener's pinna, head, and torso. Head-related transfer functions (HRTFs) contain these binaural and monaural cues and thus describe the sound incidence from a source to both ears [5].

For headphone-based virtual acoustic environments (VAEs), a set of HRTFs is essential. Ideally, such a set should include individual HRTFs for a large number of directions, typically measured on a sphere around a listener. However, measuring so-called dense sets of individual HRTFs requires special equipment, experience in handling the equipment and, depending on the measurement approach, can also be time-consuming (see e.g. [6]). For this reason, it seems appealing to measure only a small number of HRTFs on a sparse spherical sampling grid with a simplified measurement system, and to apply a specific interpolation or spatial upsampling method afterwards to generate a dense HRTF set with perhaps thousands of directions.

One popular approach for spatial upsampling is interpolation in the spherical harmonics (SH) domain. For this, an HRTF set captured on a spherical sampling scheme (also simply called spatial grid) is first transformed to the SH domain applying the spherical Fourier transform (SFT). The resulting spatially continuous representation of the HRTF set in the SH domain allows for interpolation, i.e. an HRTF for any desired direction can be obtained by means of the respective inverse spherical Fourier transform (ISFT) [13]. However, the SH representation and interpolation of sparse HRTF sets suffers from so-called sparsity errors, which is a combination of spatial aliasing and truncation errors [3]. For this reason, various pre- and post-processing methods have been proposed to reduce the sparsity error and thus to improve SH interpolation of sparse HRTF sets (see e.g. [7]).

Within this scope, we presented the SUPDEq method (SUPDEq - Spatial Upsampling by Directional Equalization) as a pre- and post-processing approach allowing improved SH interpolation of sparse HRTF sets [11].

In the respective paper, we examined the performance of the SUPDEq method regarding spectral and temporal features as well as concerning modeled localization performance of reconstructed HRTFs and showed that the approach clearly outperforms common SH interpolation in terms of these features. However, as the analysis was based only on the Lebedev sampling scheme [9], the present paper now provides further evaluation investigating the influence of various (sparse) spherical sampling schemes on the performance of the SUPDEq method. As the aliasing error strongly depends on the sampling scheme [3], this evaluation is of particular interest to ensure the general applicability of the proposed upsampling method.

In this paper, we therefore compare spatially upsampled HRTF sets originally based on sparse equiangular, Gaussian, Lebedev, and Fliege grids [13, Ch. 3][9][8] at various spatial orders N to a dense reference HRTF set. Similar to the evaluation in [11], we assess the impact of the grids on the spatially upsampled HRTF sets spectrally, temporally, and by means of localization models. To anticipate some of the results, the analysis showed that the sampling scheme has very little, if any, influence on the performance of the SUPDEq method.

2 SPHERICAL SAMPLING SCHEMES

A set of HRTFs is commonly measured at discrete points on a surrounding sphere according to a spherical sampling scheme. Such a full-spherical HRTF set can be transformed to the SH domain with the discrete SFT. The sampling schemes investigated in this paper provide closed-form expressions to calculate SH coefficients, whereas SH coefficients for arbitrary sampling configurations can be computed by an inversion of the respective SH matrix [13, Ch. 3]. The latter is however not further discussed here. Given a spherical sampling scheme L with a closed-form expression, the spherical HRTF set $H(\omega, \Omega_q)$ for the left and right ear (indices for left and right are omitted here and in the following for ease of display) can be described in the SH domain by the SH coefficients $h_{nm}(\omega)$ that are computed with the discrete SFT [13, p. 58]

$$\hat{h}_{nm}(\omega) = \sum_{q=1}^{Q_L} \beta_q H(\omega, \Omega_q) [Y_n^m(\Omega_q)]^*, \quad (1)$$

with the temporal frequency ω , the Q_L directions $\Omega_q = \{(\phi_1, \theta_1), \dots, (\phi_{Q_L}, \theta_{Q_L})\}$ at azimuth ϕ and elevation θ , and the sampling weights β_q depending on the sampling scheme L . The notation $(\cdot)^*$ denotes complex conjugation and Y_n^m are the complex SH functions of order n and degree m . The ISFT can be applied to recover $H(\omega)$ at arbitrary angles allowing for SH interpolation [13, p. 17]

$$\hat{H}(\omega, \Omega) = \sum_{n=0}^N \sum_{m=-n}^n h_{nm}(\omega) Y_n^m(\Omega), \quad (2)$$

where N is the spatial order (also referred to as SH order). As discrete sampling of a function with infinite order induces spatial aliasing and truncation errors, the SH coefficients are only error-free up to a specific scheme-dependent order N_L . If the function sampled on the sphere is strictly order-limited, a sampling scheme providing a sufficient order N_L results in $h_{nm}(\omega) = \hat{h}_{nm}(\omega)$. Similar, $H(\omega) = \hat{H}(\omega)$ holds if N is chosen appropriately.

The maximum resolvable order of the sampling scheme N_L is generally defined by the number of directions (or sampling points) Q_L and by the way the sampling points are distributed around the surface of the sphere. This relationship can be expressed by $Q_L \geq \eta(N+1)^2$, with η describing the efficiency of the sampling scheme [12]. The SH order of an HRTF set however increases as the frequency increases, following the relation $N \sim kr$, with k the wavenumber and r the radius of a sphere surrounding the head [11][3]. Assuming an average human head radius of $r = 8.75 \text{ cm}$, a minimum SH order $N = 32$ is required to perform a nearly perfect SFT, ISFT, and thus SH interpolation of HRTFs for frequencies up to 20 kHz .

In research, various schemes have been developed in order to sample the sphere with the highest possible accuracy and efficiency. A good overview on different sampling approaches in the context of spatial audio can be found for example in [13, Ch. 3]. For this study, we focused on four different frequently applied schemes, namely the equiangular, Gaussian, Lebedev, and Fliege grids. The equiangular grids have a uniform distribution

of samples along ϕ and θ , with both angles sampled at $2(N+1)$ locations, requiring $4(N+1)^2$ samples in total [12]. The Gaussian grids require only $2(N+1)^2$ samples, as the elevation θ is only sampled at $(N+1)$ locations, resulting in a nearly-uniform distribution of samples along both angles [12]. However, the equiangular and the Gaussian sampling schemes do not provide uniform distributions of sample points on the surface of the sphere. The Lebedev and Fliege schemes however offer nearly-uniform distribution of samples around the surface of the sphere, with the advantage that even less sample points are required to reach a specific SH order. Thus, the Lebedev grids require approximately $1.3(N+1)^2$ samples whereas the Fliege grids only require $(N+1)^2$ sample points [12][13, Ch. 3]. Figure 1 shows the four introduced grids on the sphere, exemplarily of SH order $N = 7$, resulting in 256 points for the equiangular grid (a), 128 points for the Gaussian grid (b), 86 points for the Lebedev grid (c), and 64 points for the Fliege grid (d).

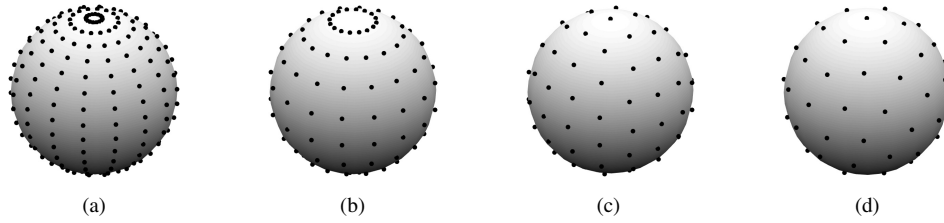


Figure 1. Equiangular (a), Gaussian (b), Lebedev (c), and Fliege (d) sampling schemes of SH order $N = 7$.

3 SPATIAL UPSAMPLING BY DIRECTIONAL EQUALIZATION (SUPDEq)

The following section gives a brief overview of the basic concept behind the SUPDEq method, as illustrated in the block diagram in Fig. 2. Further details on the implementation and evaluation can be found in [11]. Basically, the approach aims at enhanced SH interpolation and spatial upsampling of sparse HRTF sets. To achieve this, a sparse HRTF set $H(\omega, \Omega_s)$ measured at S sampling points $\Omega_s = \{(\phi_1, \theta_1), \dots, (\phi_S, \theta_S)\}$ is equalized directionally by spectral division with an appropriate equalization dataset $D_{EQ}(\omega, \Omega_s)$ before the SFT:

$$H_{EQ}(\omega, \Omega_s) = H(\omega, \Omega_s) / D_{EQ}(\omega, \Omega_s). \quad (3)$$

As a rather good and established approximation of a human head, direction-dependent rigid sphere transfer functions for an incident plane wave [13, p. 44] are used as the equalization dataset. The spherical head model should match the respective human head as best as possible. As a first and easy to implement approach, the radius of the sphere is calculated according to the physical dimensions of the head [1] and the ears are positioned at $\phi = \pm 90^\circ$ and $\theta = 0^\circ$ on the sphere. Using a spherical head model also has the advantage that it can be described analytically, which allows the calculation of the rigid sphere transfer functions at high SH orders $N_{high} \geq 32$. The directional equalization described in Eq. (3) significantly reduces the spatial complexity of the sparse HRTF set, therefore minimizing the required SH order for the SFT. The reason for the decrease of the SH order is that on the one hand, the equalization leads to a time-alignment of the HRTFs, similar to a re-centering, and on the other hand, direction-dependent influences of the sphere or the head are compensated.

After equalization, the equalized HRTF set $H_{EQ}(\omega, \Omega_s)$ is transformed to the SH domain with the SFT (Eq. (1)) at a low SH order N_{low} according to the maximum resolvable SH order of the sparse sampling scheme. Then, an upsampled (equalized) HRTF set $\hat{H}_{HRTF, EQ}$ is calculated on a dense sampling grid $\Omega_d = \{(\phi_1, \theta_1), \dots, (\phi_D, \theta_D)\}$, with $D \gg S$ using the ISFT (Eq. (2)). Finally, HRTFs are reconstructed with a subsequent de-equalization by spectral multiplication with an appropriate de-equalization dataset D_{DEQ} :

$$\hat{H}_{DEQ}(\omega, \Omega_d) = \hat{H}_{EQ}(\omega, \Omega_d) \cdot D_{DEQ}(\omega, \Omega_d). \quad (4)$$

Again, rigid sphere transfer functions for an incident plane wave can be used as the de-equalization dataset. In general, the de-equalization recovers energies at higher SH orders that were transformed to lower orders by

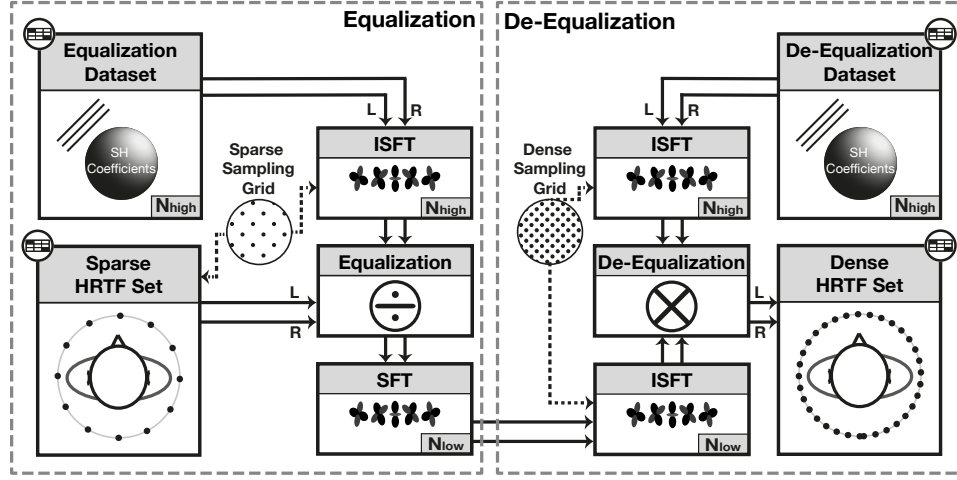


Figure 2. Block diagram of the SUPDEq method. Left panel: A sparse HRTF set is equalized on the corresponding sparse sampling grid and then transformed to the SH domain with $N = N_{low}$. Right panel: The equalized set is de-equalized on a dense sampling grid, resulting in a dense HRTF set.

the equalization. Similar as described in Sec. 2, $H = \hat{H}_{DEQ}$ holds if, in this case, N_{low} is sufficient for the SFT of H_{EQ} and N_{high} is chosen appropriately. Otherwise, spatial aliasing and truncation errors occur, resulting in $H \approx \hat{H}_{DEQ}$. The following section now analyzes the performance of the SUPDEq method with respect to the sparse sampling scheme underlying the input HRTF set in comparison to common SH interpolation.

4 INFLUENCE OF THE SAMPLING SCHEME

In our previous publication [11], we investigated the performance of the SUPDEq method for two different dummy heads, but only for Lebedev grids of different SH orders. To further ensure the general applicability of the SUPDEq method, the present paper now focuses on the influence of the spherical sampling scheme underlying the sparse input HRTF set. As a reference set, we used HRTFs of a Neumann KU100 dummy head that were measured on a Lebedev grid with 2702 sampling points [4]. This reference HRTF set was transformed to SH domain at $N = 35$, further referred to as $h_{REF,nm}$. The various sparse HRTF sets required as input data were generated by spatial subsampling of the reference set $h_{REF,nm}$ to the respective sparse equiangular, Gaussian, Lebedev, or Fliege grids of (limited) SH orders $N = 1 - 15$ applying the ISFT. Next, these sparse HRTF sets were spatially upsampled to a dense sampling grid (again the Lebedev grid with 2702 sampling points, further abbreviated Lebedev₂₇₀₂), applying the SUPDEq method as well as (order-limited) SH interpolation without any pre- or post-processing before or after the SFT/ISFT. The upsampled dense HRTF sets were then again transformed to SH domain at $N = 35$, resulting in SH coefficients further referred to as $h_{DEQ,nm}$ and $h_{OL,nm}$, with *DEQ* standing for de-equalized and *OL* for (strictly) order-limited. The de-equalized and order-limited HRTFs, as hereinafter referred to, were then obtained via the ISFT at the direction required for the respective analysis method. The optimal radius for the rigid sphere model used for (de-)equalization was calculated based on the dimensions of the Neumann KU100 dummy head [1], leading to $r = 9.19$ cm.

4.1 Spectral differences

As a first error measure, we analyzed the spectral differences between $h_{REF,nm}$ and $h_{DEQ,nm}$ or $h_{OL,nm}$ as a function of the SH order N on various test sampling grids with T sampling points $\Omega_t = \{(\phi_1, \theta_1), \dots, (\phi_T, \theta_T)\}$.

The frequency-dependent spectral differences per sampling point were calculated in dB as

$$\Delta g(\omega, \Omega_t) = 20 \lg \frac{|H_{REF}(\omega, \Omega_t)|}{|H_{TEST}(\omega, \Omega_t)|}, \quad (5)$$

where H_{REF} is the left ear HRTF extracted from $h_{REF,nm}$ and H_{TEST} is the left ear HRTF extracted from $h_{OL,nm}$ or $h_{DEQ,nm}$ at the sampling point Ω_t . Then, the absolute value of $\Delta g(\omega, \Omega_t)$ was averaged across all sampling points Ω_t to obtain the frequency-dependent measure $\Delta G_f(\omega)$ (in dB)

$$\Delta G_f(\omega) = \frac{1}{n_{\Omega_t}} \sum_{\Omega_t=1}^{n_{\Omega_t}} |\Delta g(\omega, \Omega_t)|, \quad (6)$$

and across ω and Ω_t , resulting in a single value ΔG (in dB) describing the spectral difference

$$\Delta G = \frac{1}{n_{\Omega_t}} \frac{1}{n_{\omega}} \sum_{\Omega_t=1}^{n_{\Omega_t}} \sum_{\omega=1}^{n_{\omega}} |\Delta g(\omega, \Omega_t)|. \quad (7)$$

Figure 3(a) shows the spectral differences ΔG across N over the full audio bandwidth for the four different sampling schemes applying the SUPDEq method or order-limited interpolation. The test sampling grid Ω_t was the reference Lebedev₂₇₀₂ grid. Independent of the sampling scheme, SUPDEq processing results in about 2 dB less spectral differences than order-limited interpolation. The Fliege scheme however has distinct outliers at $N = 10$ and $N = 12$ for both upsampling methods. Interestingly, exactly at these orders, some of the calculated weights are negative, which is something Fliege and Maier could not explain [8]. Applying the SFT according to Eq.(1), the negative weights most probably lead to a phase shift in the complex SH coefficients, certainly resulting in reconstruction errors when transformed back with the ISFT. Apart from these outliers, the spectral differences for the four different sampling schemes are pretty similar, indicating that the performance of the SUPDEq method is independent of the sampling scheme. For order-limited interpolation, the equiangular scheme leads to slightly higher spectral differences than the other schemes.

Figure 3(b) illustrates the frequency-dependent spectral differences, exemplarily at $N = 7$, for the four different sampling schemes applying the SUPDEq method or order-limited interpolation. As before, Ω_t was the Lebedev₂₇₀₂ grid. It can be seen that the spectral differences are significantly smaller for the SUPDEq method than for order-limited interpolation, and furthermore that order-limited interpolation leads to a sharp increase in spectral differences above the spatial aliasing frequency. Regarding the SUPDEq method, the Fliege scheme performs a little worse than the three other schemes, but overall there is only a marginal influence of the sampling scheme on the performance of the method. Furthermore, the equiangular scheme induces slightly higher spectral differences than the other schemes when applying order-limited interpolation.

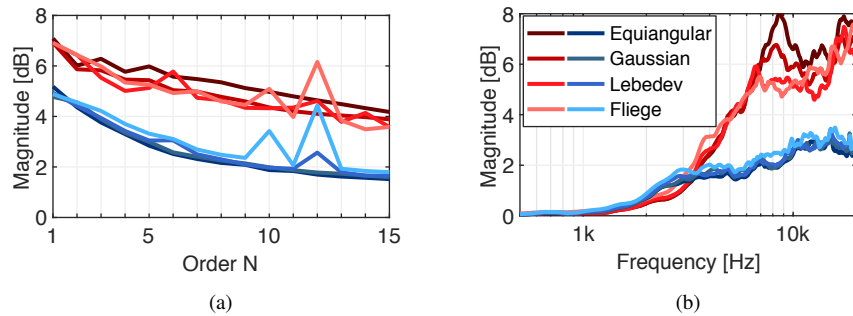


Figure 3. Spectral differences in dB (left ear) between reference HRTF set and order-limited (red) or de-equalized (blue) HRTF sets for four different sampling schemes (color saturation). The test grid Ω_t was always the Lebedev₂₇₀₂ grid. (a) Spectral differences ΔG across N over the full audio bandwidth. (b) Frequency-dependent spectral differences $\Delta G_f(\omega)$ at $N = 7$.

4.2 Binaural cues

Next, we compared the ILDs and ITDs of the reference HRTF set to those of order-limited or de-equalized sets, again with respect to different sampling schemes. For this, HRTFs in the horizontal plane ($\theta = 0^\circ$) with an angular spacing of $\phi = 1^\circ$ were extracted from the reference set $h_{REF,nm}$ and, depending on N , from the respective order-limited or de-equalized set $h_{OL,nm}$ and $h_{DEQ,nm}$. The broadband ILDs were then calculated as the ratio between the energy of the left and right ear HRIR (HRIR, the time-domain equivalent of an HRTF). The ITDs were calculated by means of a threshold-based onset detection on the ten times up-sampled and low-pass filtered HRIRs (10th order Butterworth low-pass at 3kHz).

Figure 4 illustrates the calculated ILDs and ITDs of the reference HRTF set as well as of the order-limited and de-equalized sets, again exemplarily at $N = 7$. As can be seen in Fig. 4(a), the ILDs of the de-equalized HRTFs are in good agreement with the reference and mostly unaffected by the sampling scheme. Overall, the Fliege grid shows the most notable deviations, especially at lateral directions. At these directions, also the Gaussian scheme provides clear deviations from the reference, whereas the equiangular and Lebedev schemes show only slight differences over the entire angular range. In contrast, the ILDs of the order-limited HRTFs (see Fig. 4(b)) differ significantly from the reference. However, there is also only a rather weak influence of the sampling scheme. Regarding the ITDs, Fig. 4(c) and (d) illustrate that there is virtually no influence of the sampling scheme, regardless of the upsampling method. Thus, at $N = 7$, the ITDs of the de-equalized HRTFs (see Fig. 4(c)) as well as of the order-limited HRTFs (see Fig. 4(d)) are in good agreement with the reference.

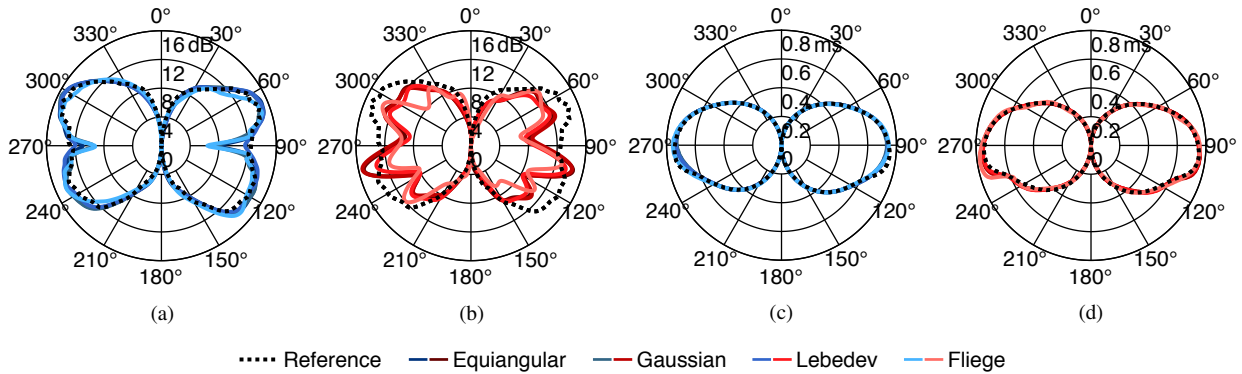


Figure 4. ILDs (a), (c) and ITDs (b), (d) in the horizontal plane of the reference (black) HRTFs as well as of the order-limited (red) or de-equalized (blue) HRTFs for four sampling schemes (color saturation) at $N = 7$.

4.3 Localization performance

To conclude the analysis, we compared the localization performance of order-limited and de-equalized HRTFs with respect to the sampling scheme applying two different auditory models from the Auditory Modeling Toolbox [14]. To assess the localization performance in the median sagittal plane, we used the model from Baumgartner et al. [2], which provides estimates for the polar RMS error (PE) as well as for the quadrant error rate (QE) based on monaural spectral cues. To evaluate the performance in the horizontal plane, we applied the model from May et al. [10], which estimates the azimuthal position of a sound source based on binaural cues. By comparing the intended and the estimated source position, a lateral error (LE) can be calculated. To calculate the error measures, first the performance of $h_{REF,nm}$, $h_{OL,nm}$, and $h_{DEQ,nm}$ was determined for each sampling scheme as a function of N . To estimate median plane localization performance, we used a test sampling grid Ω_t with $\phi = \{0^\circ, 180^\circ\}$ and $-30^\circ \leq \theta \leq 90^\circ$ in steps of 1° , and assumed a median listener sensitivity of $S = 0.76$. To estimate the horizontal plane localization performance, we applied a test sampling grid with $\phi = \pm 90^\circ$ in steps of 5° . As final error measures, the absolute polar error difference (in degree)

$$\Delta PE = |PE_{REF} - PE_{TEST}|, \quad (8)$$

the absolute quadrant error difference (in percent)

$$\Delta QE = |QE_{REF} - QE_{TEST}|, \quad (9)$$

as well as the absolute lateral error difference (in degree)

$$\Delta LE = \frac{1}{T} \sum_{t=1}^T |LE_{REF}(\Omega_t) - LE_{TEST}(\Omega_t)|, \quad (10)$$

were calculated for each sampling scheme and order N , with the subscripts *REF* and *TEST* as defined above.

In the horizontal plane (see Fig. 5(a)), the order-limited interpolation leads to an error increase at low orders $N \leq 4$. Similar to previous results, the Fliege grid performs worst here, especially at these low orders. In contrast, the SUPDEq method leads to hardly any increase in lateral error over the entire tested range of N , no matter which sampling scheme was applied. This shows that even at low orders, upsampling with the SUPDEq method always results in sufficient binaural cues.

In the median sagittal plane (see Fig. 5(b) and (c)), order-limited interpolation leads to considerably higher errors over the entire range of N . Obviously, the high-frequency deviations in order-limited HRTFs badly affects the monaural spectral cues. Overall, the Fliege and the Lebedev grids seem to perform worse, but it is difficult to see a clear trend besides a general decrease in error with increasing order N . The SUPDEq method however amplifies the polar error only slightly at low orders $N \leq 2$. Once again, the Fliege grid tends to lead to a higher increase in error than the other grids, both for the polar error as well as for the quadrant error rate. The other grids perform more or less the same, with only slight increases in polar error and quadrant error rate at $N \geq 2$, indicating that spectral cues are only marginal impaired.

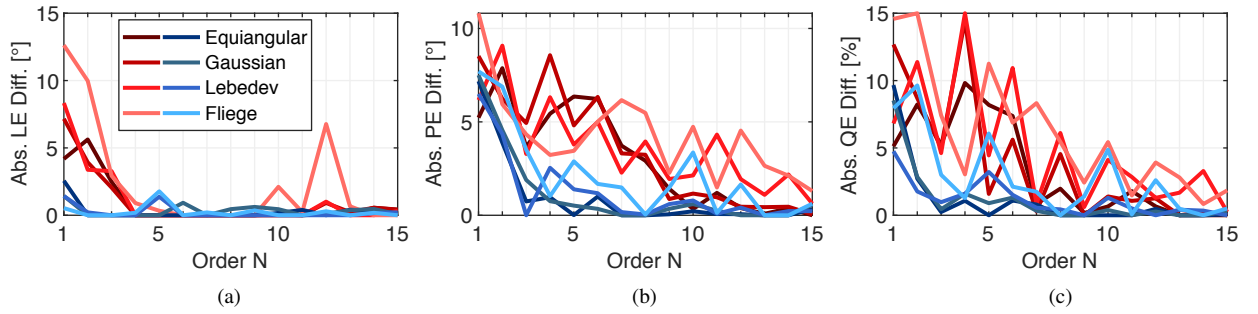


Figure 5. Absolute lateral error difference ΔLE (a), polar error difference ΔPE (b), and quadrant error difference ΔQE (c) across N for four different sampling schemes (color saturation) applying order-limited interpolation (red) or the SUPDEq method (blue).

5 CONCLUSION

This paper presented further evaluation of the SUPDEq method for spatial upsampling of sparse (individual) HRTF sets by investigating the influence of the spherical sampling scheme on the performance of the method. The study compared spatially upsampled HRTF sets originally based on sparse equiangular, Gaussian, Lebedev, and Fliege grids at various spatial orders N to a dense reference HRTF set, applying the SUPDEq method as well as common SH interpolation for upsampling. The analysis of spectral features, binaural cues, and localization performance revealed that the influence of the sampling scheme on the results of SUPDEq processing

is only marginal. Overall, only the Fliege scheme tended to perform a little worse than the other three tested schemes. With order-limited interpolation, the sampling scheme affected the examined features slightly stronger. The results of this study confirm or at least increase the general applicability of the SUPDEq method regarding the sampling scheme of the input HRTF set. Thus, the SUPDEq method might be applied with any sparse HRTF set measured on a proper full-spherical sampling grid. However, we only examined sampling schemes providing a closed-form expression in this paper. Therefore, further tests with arbitrary sampling schemes could be performed, even though it seems that given a reasonable sparse sampling scheme providing a well-conditioned (inverse) SH matrix, the results will be quite similar. Furthermore, listening experiments could be performed to analyze the perceptual influence of the sampling scheme, although the analysis in this paper suggests that the perceptual influence might be marginal. A Matlab-based implementation of the SUPDEq method is available on <https://github.com/AudioGroupCologne/SUPDEq>. The research presented in this paper was funded by the German Federal Ministry of Education and Research (BMBF 03FH014IX5-NarDasS).

REFERENCES

- [1] V. R. Algazi, C. Avendano, and R. O. Duda. Estimation of a Spherical-Head Model from Anthropometry. *J. Audio Eng. Soc.*, 49(6):472–479, 2001.
- [2] R. Baumgartner, P. Majdak, and B. Laback. Modeling sound-source localization in sagittal planes for human listeners. *J. Acoust. Soc. Am.*, 136(2):791–802, 2014.
- [3] Z. Ben-Hur, D. L. Alon, B. Rafaely, and R. Mehra. Loudness stability of binaural sound with spherical harmonic representation of sparse head-related transfer functions. *EURASIP J. Audio, Speech, Music Process.*, 2019(5):1–14, 2019.
- [4] B. Bernschütz. A Spherical Far Field HRIR / HRTF Compilation of the Neumann KU 100. In *Proc. 39th DAGA*, pages 592–595, 2013.
- [5] J. Blauert. *Spatial Hearing*. MIT Press, Cambridge, MA, 1996.
- [6] R. Bomhardt, M. de la Fuente Klein, and J. Fels. A high-resolution head-related transfer function and three-dimensional ear model database. *Proc. Meet. Acoust.*, 29(1):1–11, 2017.
- [7] F. Brinkmann and S. Weinzierl. Comparison of head-related transfer functions pre-processing techniques for spherical harmonics decomposition. In *Proc. Audio Eng. Soc. Conf. Audio for Virtual and Augmented Reality*, pages 1–10, 2018.
- [8] J. Fliege and U. Maier. The distribution of points on the sphere and corresponding cubature formulae. *SIAM J. Numer. Anal.*, 19(2):317–334, 1999.
- [9] V. I. Lebedev. Spherical quadrature formulas exact to orders 2529. *Siberian Math. J.*, 18(1):132–142, 1977.
- [10] T. May, S. Van De Par, and A. Kohlrausch. A probabilistic model for robust localization based on a binaural auditory front-end. *IEEE Trans. Audio, Speech, Lang. Process.*, 19(1):1–13, 2011.
- [11] C. Pörschmann, J. M. Arend, and F. Brinkmann. Directional Equalization of Sparse Head-Related Transfer Function Sets for Spatial Upsampling. *IEEE Trans. Audio, Speech, Lang. Process.*, 27(6):1060–1071, 2019.
- [12] B. Rafaely. Analysis and Design of Spherical Microphone Arrays. *IEEE Trans. Speech, Audio Process.*, 13(1):135–143, 2005.
- [13] B. Rafaely. *Fundamentals of Spherical Array Processing*. Springer-Verlag, Berlin Heidelberg, 2015.
- [14] P. Søndergaard and P. Majdak. The Auditory Modeling Toolbox. In J. Blauert, editor, *The Technology of Binaural Listening*, pages 33–56. Springer-Verlag, Berlin Heidelberg, 2013.