COLOURATION IN 2.5D LOCAL WAVE FIELD SYNTHESIS USING SPATIAL BANDWIDTH-LIMITATION

Fiete Winter,[1*] *Christoph Hold,*[2] *Hagen Wierstorf,*[3] *Alexander Raake,*[4] *Sascha Spors*[1]

[1]University of Rostock, Institute of Communications Engineering, Rostock, D-18119, Germany
[2]Technische Universität Berlin, Assessment of IP-based Applications, Berlin, D-10587, Germany
[3]University of Surrey, Centre for Vision, Speech and Signal Processing, Guildford, GU2 7XH, UK
[4]Technische Universität Ilmenau, Audiovisual Technology Group, Ilmenau, D-98693, Germany

ABSTRACT

Sound Field Synthesis techniques, such as Wave Field Synthesis aim at a physically accurate reproduction of a desired sound field inside an extended listening area. This area is surrounded by loudspeakers individually driven by their respective driving signals. Due to practical limitations, artefacts impair the synthesis accuracy resulting in a perceivable change in timbre compared to the desired sound field. Recently, an approach for so-called Local Wave Field Synthesis was published which enhances the reproduction accuracy in a limited region by applying a spatial bandwidth limitation in the circular/spherical harmonics domain to the desired sound field. This paper reports on a listening experiment comparing conventional Sound Field Synthesis techniques with the mentioned approach. Also the influence of the different parametrisations for Local Wave Field Synthesis is investigated. The results show that the enhanced reproduction accuracy in Local Wave Field Synthesis leads to an improvement with regard to the perceived colouration.

1. INTRODUCTION

Sound Field Synthesis (SFS) subsumes techniques which aim to synthesise a desired sound field within a target region in a physically accurate manner. Two well-known representatives are Near-Field-Compensated Higher Order Ambisonics (NFC-HOA) [1] and Wave Field Synthesis (WFS) [2]. A distribution of (up to hundreds) loudspeakers surrounding the region are employed to achieve the synthesis in practice. As the density of the distribution is limited for practical reasons, spatial aliasing artefacts impair the synthesis accuracy. Techniques for Local Sound Field Synthesis (LSFS) are useful for applications, where the listener's position is restricted to a small region of interest or is tracked using a suitable technology. They focus on a more accurate reproduction inside an area which is smaller than the area surrounded by the loudspeaker distribution. Among other approaches [3, 4, 5, 6, 7], the authors have published a method for Local Wave Field Synthesis (LWFS) [8] which expands the desired sound field into circular/spherical harmonics around a chosen expansion centre. Spatial aliasing is reduced by truncating the expansion and reproducing the resulting sound field with conventional WFS. Former is referred to as spatial bandwidth limitation.

A discussion about the actual definition of timbre and colouration in [9, Sec. 2] comes to the conclusion, that timbre is obviously a multidimensional percept, and colouration describes the distance between two points within that timbral space. One point serves as the reference, which may be presented explicitly to the listener or has been built up internally by prior listening experience. In the context of surround sound, timbre has been identified as a significant aspect for the overall sound quality humans perceive [10]. It has further been shown for conventional WFS [9], that spatial aliasing leads to a perceivable colouration between the reproduced and the desired sound field. The colouration becomes stronger the sparser the loudspeaker distribution. As LSFS enhances the synthesis accuracy around the listener's position, the question arises, if the perceived colouration can be reduced by such techniques as well.

Within this work, the results of a listening experiment comparing the colouration introduced by the mentioned LWFS method, conventional SFS techniques, and classical stereophony is presented. Further, the impact of different parametrisations of LWFS is investigated. The paper is organised as follows: We briefly describe the approach for LWFS in Sec. 2 and explain the experimental setup in Sec. 3. In Sec. 4, the results of the experiment are discussed. Sec. 5 closes this contribution with conclusions.

2. LOCAL WAVE FIELD SYNTHESIS USING SPATIAL BANDWIDTH-LIMITATION

The fundamental task in SFS including LSFS and LWFS is to synthesise a desired sound field $P(\mathbf{x}, \omega)$ inside a defined region. An ensemble of loudspeakers positioned at discrete positions $\mathbf{x}_0$ outside this region is driven by individual driving signals $D(\mathbf{x}_0, \omega)$. The angular frequency $\omega = 2\pi f$ is defined via the time frequency f. The resulting superposition of all loudspeakers constitutes the reproduced sound field. The driving signals have to be chosen such that the reproduced and the desired sound field coincide. In $2^1/_2$-dimensional (2.5D) synthesis scenarios [11, Sec. 2.3], reproduction and the loudspeaker positions are restricted to the horizontal plane, i.e $z = z_0 = 0$.

2.1. Driving Signal

Any two-dimensional sound field $P(\mathbf{x}, \omega)$ which is source-free inside the region of interest may be expanded into plane waves with their propagation directions $\mathbf{n}_{\mathrm{pw}} = [\cos\phi_{\mathrm{pw}}, \sin\phi_{\mathrm{pw}}, 0]^{\mathrm{T}}$ continuously distributed along the unit circle [12, Eq. (2.246)]. The resulting coefficients $\bar{P}(\phi_{\mathrm{pw}}, \mathbf{x}_{\mathrm{c}}, \omega)$ of this expansion additionally depend on the expansion centre $\mathbf{x}_{\mathrm{c}}$. In practice, the continuous plane wave expansion is discretised to N_{pw} angles distributed equiangularly along the unit circle. In the LWFS approach, each plane wave is synthesised using conventional WFS and weighted by its

*Correspondence should be adressed to fiete.winter@uni-rostock.de

978-1-5386-1631-4/17/$31.00 ©2017 IEEE

respective coefficient. The resulting driving function reads

$$D(\mathbf{x}_0,\omega) \approx \sqrt{\mathrm{j}\frac{\omega}{\mathrm{c}}}\,\frac{1}{N_{\mathrm{pw}}} \times \sum_{m=0}^{N_{\mathrm{pw}}-1} \breve{P}(m\Delta\phi_{\mathrm{pw}},\mathbf{x}_{\mathrm{c}},\omega)\, D_{\mathrm{pw}}^{\mathrm{WFS}}\left(\mathbf{x}_0',m\Delta\phi_{\mathrm{pw}},\omega\right), \tag{1}$$

where the angular step size is defined as $\Delta\phi_{\mathrm{pw}} = {}^{2\pi}/_{N_{\mathrm{pw}}}$. The 2.5D WFS driving function $D_{\mathrm{pw}}^{\mathrm{WFS}}$ for a single plane wave is given by [13, Eq. (2.177)]. The loudspeaker position $\mathbf{x}_0' = \mathbf{x}_0 - \mathbf{x}_{\mathrm{c}}$ is given in a shifted coordinate frame with expansion centre as its origin. The regular circular expansion with its coefficient $\breve{P}_\mu(\mathbf{x}_{\mathrm{c}},\omega)$ can be utilised as another possibility to expand a sound field $P(\mathbf{x},\omega)$. An approximative conversion formula between the two representations is given by [12, Eq. (4.91)]

$$\breve{P}(\phi_{\mathrm{pw}},\mathbf{x}_{\mathrm{c}},\omega) \approx \sum_{\mu=-M}^{M} \mathrm{j}^{\mu}\breve{P}_\mu(\mathbf{x}_{\mathrm{c}},\omega)\mathrm{e}^{+\mathrm{j}\mu\phi_{\mathrm{pw}}}\,. \tag{2}$$

Here, the actual spatial bandwidth limitation is applied to the desired sound field, as the infinite summation is truncated to $2M+1$ coefficients. M is usually referred to as the spatial bandwidth. For a virtual point source located at $\mathbf{x}_{\mathrm{s}}' = \rho_{\mathrm{s}}'[\cos\phi_{\mathrm{s}}', \sin\phi_{\mathrm{s}}', 0]^{\mathrm{T}}$, the circular expansion coefficients are given as [14, Eq. (13)]

$$\breve{P}_\mu(\mathbf{x}_{\mathrm{c}},\omega) = \hat{S}(\omega)\frac{\mathrm{j}^{|\mu|-\mu}}{4\pi}\left(-\mathrm{j}\frac{\omega}{\mathrm{c}}\right)h_{|\mu|}^{(2)}\left(\frac{\omega}{\mathrm{c}}\rho_{\mathrm{s}}'\right)\mathrm{e}^{-\mathrm{j}\mu\phi_{\mathrm{s}}'}, \tag{3}$$

where $\hat{S}(\omega)$ incorporates the spectrum of the source signal emitted by the point source. The nth order spherical Hankel function of second kind and the speed of sound are denoted by $h_n^{(2)}(\cdot)$ and c, respectively. Since a point source is a three-dimensional sound field, its circular basis expansion is derived using an approximation. For more information about the discrete-time implementation of this approach, the reader is referred to [14].

2.2. Influence of the Parametrisation

In the following, the influence of the number of plane waves N_{pw}, spatial bandwidth M and the expansion centre $\mathbf{x}_{\mathrm{c}}$ is discussed. As shown in Fig. 1, a circular array consisting of $L = 56$ loudspeakers with radius of 1.5 m is used to synthesise a point source located at $[0, 2.5, 0]^{\mathrm{T}}$ m. It is evident from Fig. 1b that conventional WFS introduces visible aliasing artefacts which are due to the finite number of loudspeakers and the infinite spatial bandwidth of the WFS driving function. The artefacts are stronger for positions near the virtual point source. They lead to significant fluctuations in the magnitude spectrum of the reproduced sound field, cf. WFS in Fig. 2 (left). Reducing M to a finite value by using LWFS, cf. Fig. 1c, eventually leads to a circular region of high accuracy around the expansion centre. Its maximally possible radius can be approximated by $r_M(f) = {}^{Mc}/_{2\pi f}$ [15, Eq. (2.41)], which states an inversely proportional relation to f. If the expansion centre is co-located with the centre of the loudspeaker array, the optimal spatial bandwidth is given via $M = \lfloor {}^{(L-1)}/_{2} \rfloor$ [15, Eq. (4.26)], where the floor operator $\lfloor\cdot\rfloor$ rounds its argument towards the next smaller integer. Choosing a higher value for M will lead to aliasing inside the region. However, if the expansion centre is shifted towards the position of the virtual point source, the spatial aliasing artefacts become stronger and decrease the accuracy inside the previously aliasing-free region, compare Fig. 1c and 1e. This is also indicated by the magnitude spectra

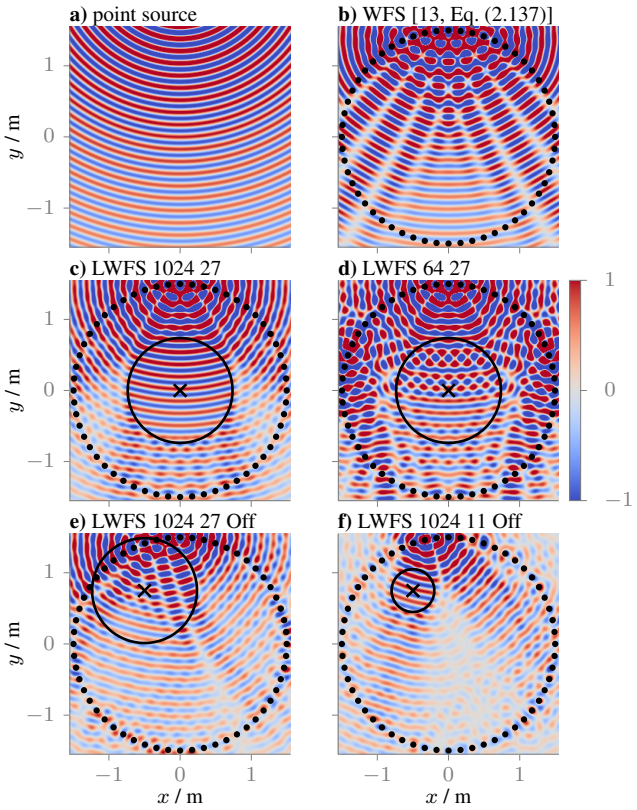

Figure 1: The top left plot shows a monochromatic ($f = 2$ kHz) point source at $[0, 2.5, 0]^{\mathrm{T}}$ m. It serves as the desired sound field for the SFS approaches, whose reproduced sound fields are shown in the other plots. A circular array of 56 loudspeakers (black dots) was used. For LWFS, the first value corresponds to N_{pw}, the second one to M. The expansion centre $\mathbf{x}_{\mathrm{c}}$ (black cross) is set to $[0, 0, 0]^{\mathrm{T}}$ m for **c)** and **d)**, and to $[-0.5, 0.75, 0]^{\mathrm{T}}$ m for **e)** and **f)**. The black circle around it indicates the circular area of radius $r_M(f)$.

for LWFS ($N_{\mathrm{pw}} = 1024$) shown Fig. 2 (left), where the spectral deviations are stronger for the position closer to the virtual point source. It is hence necessary to further reduce M in order to avoid aliasing, compare Fig. 1e and 1f. As a drawback, the radius of the area of high accuracy is further decreased for the same frequency. Since the region shrinks further with increasing frequency, $r_M(f)$ eventually gets near or even smaller than the human head radius for frequencies inside the audible range. Fig. 2 (right) shows, that small values of M cause a significant loss of magnitude at high frequencies. It is hence evident, that the optimal choice for M compromises between aliasing artefacts and the maximum possible $r_M(f)$, as both increase with M. For a more detailed analysis of the spatial occurrence of spatial aliasing artefacts in LSFS, see [16, Sec. IV-D].

Similar to the spatial sampling introduced by the finite number of employed loudspeakers, the discretisation of the plane wave decomposition to N_{pw} equiangular coefficients describes an additional sampling process. Hence, a too coarse sampling introduces spatial aliasing to the reproduced sound field which is observed in Fig. 1d. Contrary to the number of loudspeakers, however, the only limiting factor is the computational effort which grows linearly with N_{pw}. In Fig. 2 (left), different values for N_{pw} are compared for two $\mathbf{x}_{\mathrm{c}}$: as expected, the spatial aliasing leads to rapid fluctuations which start at lower frequencies the lower N_{pw} is. Again, more

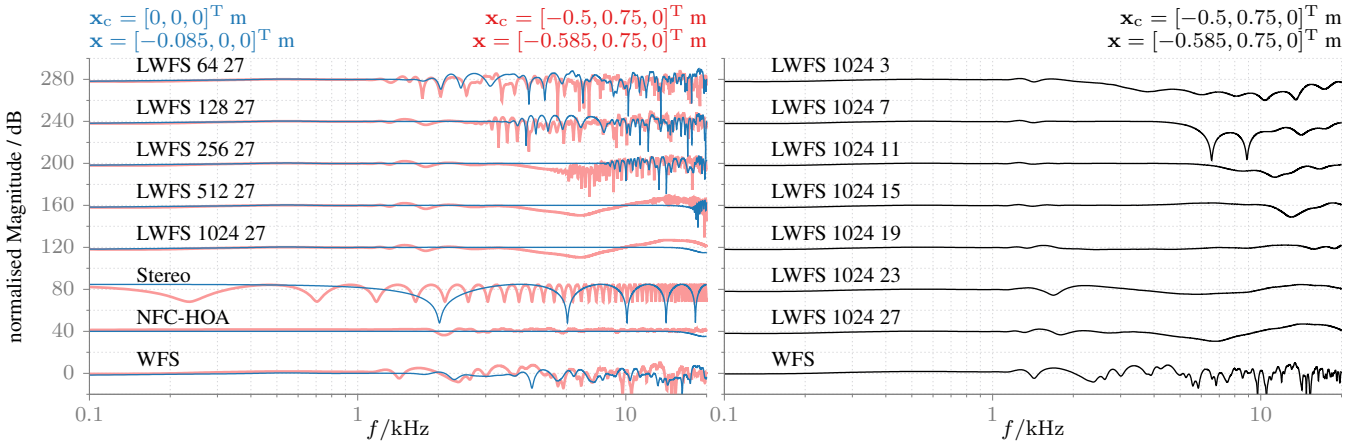

Figure 2: The magnitude spectra of the reproduced sound fields are evaluated by a virtual omnidirectional microphone positioned at $\mathbf{x}$. The left plot shows the spectra for the 8 reproduction methods used in the first (blue) and second part (red) of the listening experiment. The right plot shows the spectra for the methods used in the third part. For LWFS, the first value corresponds to N_{pw}, the second one to M. All spectra are normalised to the sound field of the desired point source evaluated at $\mathbf{x}$. The expansion centre for LWFS is given by $\mathbf{x}_c$.

aliasing is introduced for an expansion centre which is closer to the point source. Vice versa, more plane waves are necessary for this position in order to suppress the aliasing.

3. EVALUATION METHODS

The colouration experiment was performed using an approach for non-individual binaural synthesis [17] which allows for the simulation of sound sources and whole loudspeaker arrays via headphones. The originally used Head Related Transfer Functions (HRTFs) [18] were low frequency corrected allowing for a better interpolation of missing sound source positions. For details, see [19]. The head tracking was deactivated to avoid changes in timbre due to head rotation. A discussion on the limitations of binaural synthesis for colouration experiments is given in [9, Sec. 4]. It is assumed within this study that the binaural synthesis approximately has the same influence on all presented conditions neglecting possible interaction effects. It thus can be used to investigate the difference in timbre between the conditions.

The experiment was conducted separately at two facilities. At the University of Rostock, it took place in a $86\,\mathrm{m}^3$ acoustically damped room (Audio laboratory, Institute of Communications Engineering). At the Technische Universität Berlin, the experiment was conducted in a $54\,\mathrm{m}^3$ acoustically damped listening room (room Pinta, Telefunken Building). In both cases, the listeners wore open headphones (AKG K601). In a separate room, a computer equipped with a sound card (Focusrite Scarlett 2i2, 1st Gen. in Rostock and RME Hammerfall DSP MADI + Behringer HA4700 Powerplay Pro-XL in Berlin) was used for audio playback. The signals travelled through an analogue cable of approximately 6 m length to the headphones inside the listening room.

11 and 9 listeners were recruited for the experiment in Rostock and Berlin, respectively. The age of the participants ranged from 19 to 60 years with an average of approximately 34 years. All test participants self-reported normal hearing.

3.1. Stimuli

The reference condition for the listeners to judge colouration against was a binaurally simulated point source positioned at $[0, 2.5, 0]^T$ m, cf. Fig. 1a. WFS, NFC-HOA, and LWFS were employed to synthesise this point source with a binaurally simulated, circular array of 56 loudspeakers with 1.5 m radius, centred at the coordinate origin. For NFC-HOA, the 2.5D driving function for a point source implemented by IIR-filters [20] is used. The experiment was split into three parts: the listening position was set to the coordinate origin for the first and to $[-0.5, 0.75, 0]^T$ m for second and the third part. The latter will be referred to as the off-centre position. The positions are shown in Fig. 1c-f. In the first two parts, the conditions were WFS, NFC-HOA and five times LWFS, with N_{pw} ranging from 64 to 1024 on a base-2 logarithmic scale. The spatial bandwidth M was kept constant at 27 for NFC-HOA and LWFS. Additionally, a stereophonic setup with the loudspeakers positioned at $[\pm 1.4, 2.5, 0]^T$ m and the phantom source panned to the centre between both loudspeakers was included. This results into 8 conditions for each part in total. The third part included one WFS and seven LWFS conditions with $N_{\mathrm{pw}} = 1024$ and M varying from 3 to 27 on a linear scale. In all parts, the expansion centre $\mathbf{x}_c$ in LWFS was set to the respective listening position. The listener was always oriented towards the point source except for the off-centre stereo condition, where the head pointed to the nearest loudspeaker. For all conditions, the magnitude spectra of the reproduced sound fields near the listening position are plotted in Fig. 2.

For the dry source signal $\hat{S}(\omega)$, a pink noise pulse train with a pulse duration of 900 ms (cosine-shaped fade-in/fade-out of 50 ms, each) and a pause of 500 ms was used. The second signal was a female speech sample of eight seconds length. In order to avoid loudness differences among the conditions, a loudness model [21, 22] was used to adjust the loudness of all conditions to the reference. For the loudness estimation, the noise stimulus was used as a source signal. The used implementation of the model is part of [23].

3.2. Procedure

This experiment used a Multiple Stimulus with Hidden Reference and Anchor (MUSHRA) test paradigm [24]. The subjects were asked to use a Graphical User Interface (GUI) with one slider per condition to assess the respective colouration compared to the explicitly given reference signal on a continuous scale ranging from *no difference* to *very different*. Within each run, the 8 conditions,

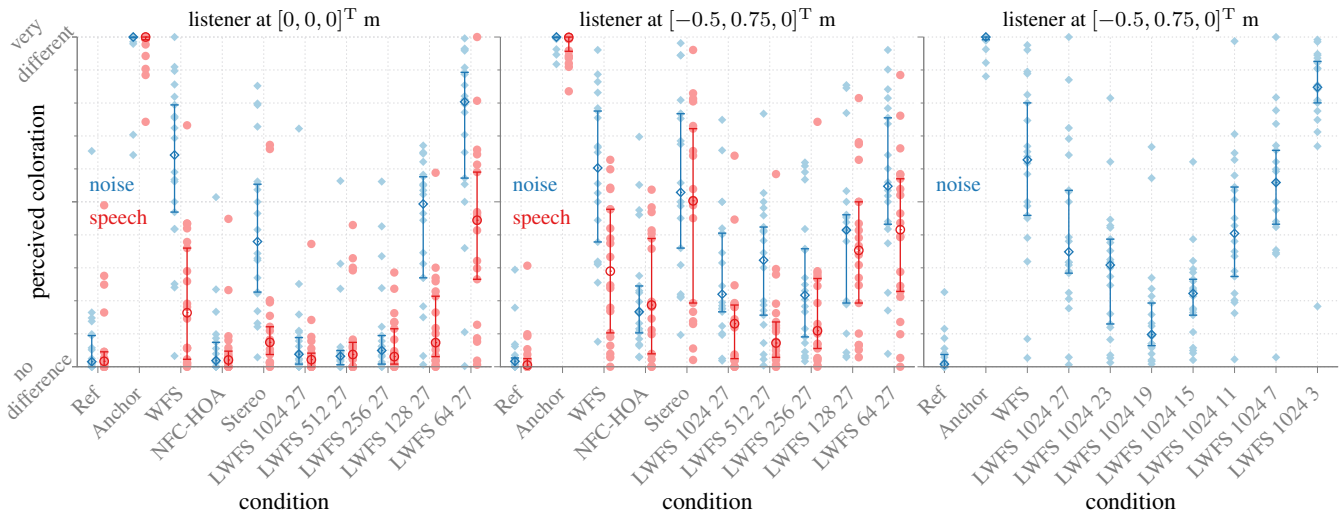

Figure 3: The plots show the colouration ratings for the three parts of the experiment (left to right). Results for individual participants $r_l^{p,c}$ are plotted with blue diamonds and red circles for the noise and the speech signal, respectively. The bold diamonds/circles and whiskers indicate the sample median $\tilde{r}^{p,c}$ and the corresponding confidence interval of the population median, respectively.

the hidden reference and a lower anchor had to be rated. The reference signal filtered by a 2nd order Butterworth high pass with a cutoff-frequency of 5 kHz served as the lower anchor. For all three parts, the noise impulse train was used as the source signal. The first two parts were repeated once using the speech signal instead, which results in five MUSHRA runs in total. The order of runs and the arrangement of the conditions in the GUI were randomised. An additional run using a music stimulus and the conditions from part one had to be passed a-priori for training. During a single run, the listener could switch instantaneously between the conditions as often as desired. The stimuli were looped.

3.3. Data Analysis

The experiment results in a three-dimensional dataset $r_l^{p,c}$ of MUSHRA ratings, where l and p correspond to one of the 20 listeners and to one of the five runs, respectively. One of the 10 stimuli per part is denoted by c. As the number of participants is relatively small, normal distribution of the data cannot be assumed ruling out several parametric statistic methods for the data analysis [25, Sec. 2.2]. Instead, non-parametric approaches assuming as little as possible about the underlying distribution are used. The 20 ratings for each (p, c) are ordered in ascending order to get the respective order statistics with $r_{(i)}^{p,c}$ denoting the ith smallest rating. The sample median $\tilde{r}^{p,c} := \frac{1}{2}(r_{(10)}^{p,c} + r_{(11)}^{p,c})$ serves as a good point estimator for population median. For its confidence interval the method given in [26, Sec. 5.2.1] is used: for a sample size of 20, the chosen interval $[r_{(6)}^{p,c}, r_{(15)}^{p,c}]$ corresponds to a confidence level of at least 0.9586.

4. RESULTS AND DISCUSSION

The results are summarised in Fig. 3. For all five runs, the hidden reference and the anchor have been rated as not different and very different from the reference, respectively. For the first two parts, the perceived colouration for the speech source signal is generally shifted towards *no difference* compared to the noise signal. This is not further surprising as the speech spectrum does not have significant energy at high frequencies, where most of the artefacts introduced by the different SFS techniques are present, cf. Fig. 2. For the centre listening position, cf. Fig. 3 (left), the perceived colouration of LWFS increases as the number of plane waves decreases. This indicates, that sufficient resolution in (1) has to be provided in order to avoid additional colouration. NFC-HOA and LWFS with $N_{\mathrm{pw}} = 1024$, 512, and 256 achieve a transparent presentation for the speech and the noise stimulus. This observation agrees with the respective spectra shown in Fig. 2 (left), where the artefacts only occur at high frequencies. In Fig. 3 (middle), the corresponding ratings for the off-centre listening position are plotted. Especially for the noise signal, previously transparent presentations also suffer from colouration now. This is related to the suboptimal choice of the spatial bandwidth M leading to spectral aliasing artefacts for this position. The influence of M can be observed in Fig. 3 (right). As already discussed in Sec. 2.2, the optimal choice M postulates a trade-off between spatial aliasing artefacts and the size of the region of accurate synthesis. Up to $M = 19$, the perceived colouration decreases as the spectral magnitude loss decreases, cf. Fig. 2 (right). However, even for this optimum value a LWFS is not transparent. For M beyond 19, colouration increases again due to significant aliasing contributions. As conventional WFS can be regarded as technique with infinite M, its rating fits into this pattern.

5. CONCLUSION

For the given loudspeaker setup and the desired point source, LWFS is able to decrease perceived colouration compared to conventional WFS. The influence of the different parameters in LWFS on the colouration agrees with the expectations raised by prior analysis of the physical properties of the reproduced sound field. However, for the investigated off-centre listening position, no parametrisation leads to a fully transparent presentation of the point source. Future work has to tackle additional modifications of the presented LWFS approach. As most promising, smoother truncation windows than the rectangular window used here may be applied to the circular expansion of the desired sound field to achieve spatial bandwidth limitation. Also other approaches for LSFS, e.g. [6], have to be included into the investigation on colouration.

6. REFERENCES

[1] J. Daniel, "Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format," in *Proc. of 23rd Intl. Aud. Eng. Soc. Conf. on Signal Processing in Audio Recording and Reproduction*, 2003.

[2] A. J. Berkhout, "A Holographic Approach to Acoustic Control," *J. Aud. Eng. Soc.*, vol. 36, no. 12, pp. 977–995, 1988.

[3] E. Corteel, C. Kuhn-Rahloff, and R. Pellegrini, "Wave Field Synthesis Rendering with Increased Aliasing Frequency," in *Proc. of 124th Aud. Eng. Soc. Conv.*, Amsterdam, The Netherlands, 2008.

[4] J. Hannemann and K. D. Donohue, "Virtual Sound Source Rendering Using a Multipole-Expansion and Method-of-Moments Approach," *J. Aud. Eng. Soc.*, vol. 56, no. 6, pp. 473–481, 2008.

[5] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, Apr. 2009, pp. 93–96.

[6] S. Spors and J. Ahrens, "Local Sound Field Synthesis by Virtual Secondary Sources," in *Proc. of 40th Intl. Aud. Eng. Soc. Conf. on Spatial Audio*, Tokyo, Japan, 2010.

[7] S. Spors, K. Helwani, and J. Ahrens, "Local sound field synthesis by virtual acoustic scattering and time-reversal," in *Proc. of 131st Aud. Eng. Soc. Conv.*, New York, USA, 2011.

[8] N. Hahn, F. Winter, and S. Spors, "Local Wave Field Synthesis by Spatial Band-Limitation in the Circular/Spherical Harmonics Domain," in *Proc. of 140th Aud. Eng. Soc. Conv.*, Paris, France, June 2016, pp. 1–12.

[9] H. Wierstorf, C. Hohnerlein, S. Spors, and A. Raake, "Coloration in Wave Field Synthesis," in *Proc. of 55th Intl. Aud. Eng. Soc. Conf. on Spatial Audio*, Helsinki, Finland, Aug. 2014.

[10] F. Rumsey, S. Zieliński, R. Kassier, and S. Bech, "On the relative importance of spatial and timbral fidelities in judgements of degraded multichannel audio quality," *J. Acoust. Soc. Am.*, vol. 118, no. 2, pp. 968–976, 2005.

[11] E. N. G. Verheijen, "Sound Reproduction by Wave Field Synthesis," Ph.D. dissertation, Delft University of Technology, 1997.

[12] A. Kuntz, "Wave Field Analysis Using Virtual Circular Microphone Arrays," Ph.D. dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2009.

[13] F. Schultz, "Sound Field Synthesis for Line Source Array Applications in Large-Scale Sound Reinforcement," Ph.D. dissertation, University of Rostock, 2016.

[14] F. Winter, N. Hahn, and S. Spors, "Time-Domain Realisation of Model-Based Rendering for 2.5D Local Wave Field Synthesis Using Spatial Bandwidth-Limitation," in *2017 25th European Signal Processing Conference (EUSIPCO)*, Kos Island, Greece, Aug. 2017, *accepted*.

[15] J. Ahrens, *Analytic Methods of Sound Field Synthesis*, ser. T-Labs Series in Telecommunication Services. Berlin Heidelberg: Springer-Verlag, 2012.

[16] F. Winter, J. Ahrens, and S. Spors, "On Analytic Methods for 2.5-D Local Sound Field Synthesis Using Circular Distributions of Secondary Sources," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 5, pp. 914–926, 2016.

[17] F. Winter, H. Wierstorf, and S. Spors, "Improvement of the reporting method for closed-loop human localization experiments," in *Proc. of 142nd Aud. Eng. Soc. Conv.*, Berlin, Germany, May 2017.

[18] H. Wierstorf, M. Geier, and S. Spors, "A Free Database of Head Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances," in *Proc. of 130th Aud. Eng. Soc. Conv.*, London, UK, 2011.

[19] V. Erbes, M. Geier, H. Wierstorf, and S. Spors, "Free Database of Low Frequency Corrected Head-Related Transfer Functions and Headphone Compensation Filters," in *Proc. of 142nd Aud. Eng. Soc. Conv.*, Berlin, Germany, May 2017.

[20] S. Spors, V. Kuscher, and J. Ahrens, "Efficient realization of model-based rendering for 2.5-dimensional near-field compensated higher order Ambisonics," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, USA, 2011.

[21] "Procedure for the computation of loudness of steady sounds," American National Standards Institute, New York, USA, Standard ANSI S3.4-2007, 2007.

[22] B. C. J. Moore, B. R. Glasberg, and T. Baer, "A model for the prediction of thresholds, loudness, and partial loudness," *J. Audio Eng. Soc*, vol. 45, no. 4, pp. 224–240, 1997.

[23] GENESIS, "Loudness Toolbox 1.0," jan 2010. [Online]. Available: http://genesis-acoustics.com/en/loudness_online-32.html

[24] "Method for the subjective assessment of intermediate quality level of audio systems," International Telecommunication Union Radiocommunication Assembly, Recommendation ITU-R BS.1534-3, 2015.

[25] T. Sporer, J. Liebetrau, and S. Schneider, "Statistics of MUSHRA revisited," in *Proc. of 127th Aud. Eng. Soc. Conv.*, New York, USA, Oct. 2009.

[26] G. J. Hahn and W. Q. Meeker, *Statistical Intervals: A Guide for Practitioners*. John Wiley & Sons, 1991.