

# On Fractional Delay Interpolation for Local Wave Field Synthesis

Fiete Winter and Sascha Spors

**Abstract**—Wave Field Synthesis aims at the accurate reproduction of a sound field inside an extended listening area which is surrounded by individually driven loudspeakers. Recently a Local Wave Field Synthesis technique has been published which utilizes focused sources as a distribution of virtual loudspeakers in order to increase the reproduction accuracy in a particular local region. Similar to conventional Wave Field Synthesis, this technique relies heavily on delaying and weighting the input signals of the virtual sound sources. As these delays are in general not an integer multiple of the input signals' sample rate, delay interpolation is necessary. This paper analyses in how far the accuracy of the delay interpolation influences the spectral properties of the synthesised sound field. The results show, that an upsampling of the virtual source's input signal is an computationally efficient tool which leads to a significant increase of accuracy.

**Index Terms**—fractional delay filter, delay interpolation, sound field synthesis, wave field synthesis, local wave field synthesis, Lagrange interpolation

## I. INTRODUCTION

Sound Field Synthesis (SFS) techniques synthesize a desired acoustic scenario within an extended listening area. Wave Field Synthesis (WFS) is a well established representative. In theory, WFS creates a reproduction of a virtual wave field using a continuous distribution of acoustic sources. A limited number (up to hundreds) of individually driven loudspeakers placed at discrete positions around the listening area realizes this distribution in practical implementations. The finite spatial resolution of this discretization may induce spatial aliasing artefacts to the reproduced wave field and therefore limits the synthesis accuracy. Current setups for WFS do not allow for an accurate synthesis within the extended area for the full audible frequency range up to 20 kHz.

For application scenarios where the listeners' position is restricted to a smaller region of interest, Local Sound Field Synthesis (LSFS) techniques are useful. They aim at a more accurate synthesis within a (local) area which is smaller than the area surrounded by the loudspeaker array. This improvement in terms of accuracy comes at the cost of stronger artefacts outside the local listening area. Among other approaches [1]–[4] for LSFS, a technique [5] has been proposed which utilizes focused sources as virtual loudspeakers surrounding the local listening area. Analogue to conventional SFS, these virtual loudspeakers are driven by a suitable SFS technique in order to reproduce the desired sound field within the local listening area. The focused sources are then synthesized by the real loudspeaker setup. It has been shown in [5], that WFS is a computationally efficient tool for implementing this LSFS technique: The resulting driving signal essentially consists of

a weighted sum of delayed versions of the virtual source's input signal.

The required delays are generally not an integer multiple of the sample period of the discretised source signal. It is hence necessary to interpolate signal values between the known samples. The reproduction system has furthermore to cope with time varying delays in dynamic scenarios, e.g. moving virtual sources. Extensive studies w.r.t. an efficient implementation of such interpolation techniques in the context of conventional WFS have been conducted by [6]. Perceptual experiments for stationary WFS scenarios were conducted by Ahrens et al. [7]: At a sampling frequency of 44.1kHz, the test subjects were not able to distinguish between a method rounding the delays to nearest sample position and interpolation techniques of higher accuracy.

This paper presents an analysis of the effects of imperfect fractional delay (FD) interpolation on the reproduction accuracy of Local Wave Field Synthesis (LWFS) for stationary scenarios. The accuracy is evaluated by the spectral properties of the synthesized sound field. We investigate the interaction of delay interpolation and spatial aliasing. We furthermore compare different standard FD filters with regard to their applicability to this synthesis technique.

This paper is organized as follows: In the next section, the theory of Local Wave Field Synthesis is revisited. Selected methods for delay interpolation are presented in Sec. III. Sec. IV presents an analysis of the spectral properties of the sound field reproduced using LWFS with various delay interpolation techniques. A conclusion is in given Sec. V.

## II. LOCAL WAVE FIELD SYNTHESIS

### A. Basic Theory

LSFS aims at the synthesis of a desired sound field  $S(\mathbf{x})$  within a defined listening area  $\Omega_1 \subseteq \Omega_0$  (cf. yellow area in Fig. 1). The dependency on the angular frequency  $\omega$  is omitted for brevity's sake. A distribution of loudspeakers is positioned along the boundary  $\partial\Omega_0$  as so-called secondary sources. The sound field emitted by an individual secondary source is modelled by a monopole point source. It is given by the three-dimensional Green's free field function [8, Eq. (8.41)]  $G(\mathbf{x}|\mathbf{x}_0)$  with  $\mathbf{x}_0 \in \partial\Omega_0$ . Each secondary source is driven by its individual driving function  $D_0(\mathbf{x}_0)$  such that the superposition of all secondary sources yields the reproduced sound field

$$P(\mathbf{x}) = \oint_{\partial\Omega_0} D_0(\mathbf{x}_0)G(\mathbf{x}|\mathbf{x}_0) dA_0 \stackrel{!}{=} S(\mathbf{x}) \forall \mathbf{x} \in \Omega_1. \quad (1)$$

Thereby,  $dA_0 = dA_0(\mathbf{x}_0)$  is a suitably chosen differential boundary element. Some SFS techniques, such as WFS, allow for the reproduction of so-called focused sources  $G_{\text{fs}}(\mathbf{x}|\mathbf{x}_{\text{fs}})$  which approximate the sound field of a monopole point source  $G(\mathbf{x}|\mathbf{x}_{\text{fs}})$  located inside  $\Omega_0$ . A set of focused sources is utilized as a virtual secondary source distribution, which is driven like a real loudspeaker setup. The virtual secondary sources are distributed along  $\partial\Omega_1$  (cf. Fig. 1). The loudspeakers' driving function is given as [9, Eq. (9)]

$$D_0(\mathbf{x}_0) = \oint_{\partial\Omega_1} D_1(\mathbf{x}_1) D_{\text{fs}}(\mathbf{x}_0|\mathbf{x}_1) dA_1, \quad (2)$$

where  $D_1(\mathbf{x}_1)$  denotes the driving function for each virtual secondary source to reproduce  $S(\mathbf{x})$  inside  $\Omega_1$ . The driving function to reproduce a particular focused source located at  $\mathbf{x}_1 \in \partial\Omega_1$  is denoted by  $D_{\text{fs}}(\mathbf{x}_0|\mathbf{x}_1)$ .

### B. Driving Functions

Many practical setups restrict the reproduction to the horizontal plane using circular or rectangular loudspeaker distributions. For such a two-dimensional scenario, theory requires line sources instead of point sources. This dimensionality mismatch is usually subsumed under 2.5D synthesis and leads to a systemic deviation of the amplitude decay between the reproduced and the desired sound field. In WFS the stationary phase approximation [10, Eq. (A2)] is utilized to derive the driving functions for 2.5D synthesis. For a focused source located at  $\mathbf{x}_1$  it is given by [11, Eq. (A.14)]

$$D_{\text{fs}}(\mathbf{x}_0|\mathbf{x}_1) = \sqrt{\frac{-jk}{2\pi}} \sqrt{\frac{|\mathbf{x}_0 - \mathbf{x}_{\text{ref}}|}{|\mathbf{x}_0 - \mathbf{x}_{\text{ref}}| - |\mathbf{x}_1 - \mathbf{x}_0|}} \times a_{\text{fs}}(\mathbf{x}_0|\mathbf{x}_1) \frac{(\mathbf{x}_1 - \mathbf{x}_0)^{\text{T}} \mathbf{n}_0}{|\mathbf{x}_1 - \mathbf{x}_0|^{3/2}} e^{+jk|\mathbf{x}_1 - \mathbf{x}_0|}, \quad (3)$$

whereas  $\mathbf{x}_{\text{ref}}$  defines the reference position, where the amplitude of the reproduced sound field matches the virtual one for high temporal frequencies. The wavenumber  $k$  is thereby defined as the quotient of  $\omega$  and the speed of sound  $c$ . The secondary source selection criterion  $a_{\text{fs}}(\mathbf{x}_0|\mathbf{x}_1)$  ensures that only secondary sources which contribute to the main propagation direction  $\mathbf{n}_1$  are active. Within this treatise a monopole point source  $G(\mathbf{x}|\mathbf{x}_{\text{ps}})$  emitting a source signal  $\hat{S}(\omega)$  will serve as the desired sound field  $S(\mathbf{x})$ . The respective 2.5D driving function is given as [12, Eq. (3.10)]

$$D_1(\mathbf{x}_1) = \hat{S}(\omega) \sqrt{\frac{jk}{2\pi}} \sqrt{\frac{|\mathbf{x}_1 - \mathbf{x}_{\text{ref}}|}{|\mathbf{x}_1 - \mathbf{x}_{\text{ref}}| + |\mathbf{x}_1 - \mathbf{x}_{\text{ps}}|}} \times a_{\text{ps}}(\mathbf{x}_1|\mathbf{x}_{\text{ps}}) \frac{(\mathbf{x}_1 - \mathbf{x}_{\text{ps}})^{\text{T}} \mathbf{n}_1}{|\mathbf{x}_1 - \mathbf{x}_{\text{ps}}|^{3/2}} e^{-jk|\mathbf{x}_1 - \mathbf{x}_{\text{ps}}|}, \quad (4)$$

where  $a_{\text{ps}}(\mathbf{x}_1|\mathbf{x}_{\text{ps}})$  denotes the respective selection criterion for a virtual point source. It is evident from the definitions in (3) and (4), that  $D_1(\mathbf{x}_1)$  and  $D_{\text{fs}}(\mathbf{x}_0|\mathbf{x}_1)$  consist of three essential components: The exponential terms state a delay depending on the geometry. Furthermore, a position independent pre-filter  $\sqrt{\pm jk/2\pi}$  is applied. The remaining terms can be subsumed

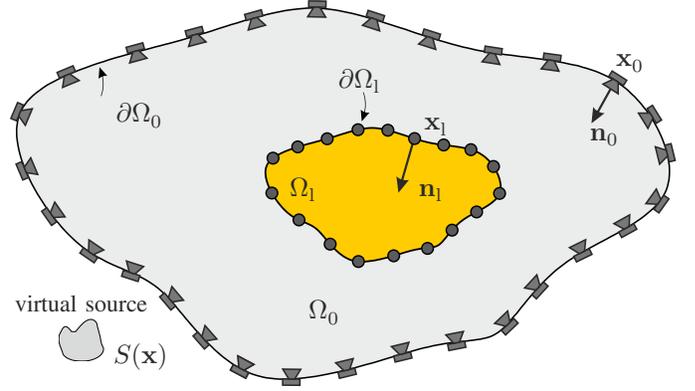


Fig. 1: The secondary sources are indicated by the loudspeaker symbols, while the focused sources, i.e. virtual secondary sources, are marked by black circles.

under a geometry dependent weighting. Hence,  $D_0(\mathbf{x}_0)$  may be expressed as

$$D_0(\mathbf{x}_0) = \hat{S}(\omega) \underbrace{\frac{k}{2\pi}}_{H_{\text{pre}}(\omega)} \underbrace{\oint_{\partial\Omega_1} w(\mathbf{x}_0, \mathbf{x}_1)}_{\text{weighting}} \underbrace{e^{-j\omega\tau(\mathbf{x}_0, \mathbf{x}_1)}}_{\text{delaying}} dA_1 \quad (5)$$

where the dependencies on  $\mathbf{x}_{\text{ps}}$  and  $\mathbf{x}_{\text{ref}}$  are skipped for brevity's sake. The multiplication of the two pre-filters in (3) and (4) yields  $H_{\text{pre}}(\omega)$ . Applying the inverse Fourier transform to (5) yields

$$d_0(\mathbf{x}_0, t) = h_{\text{pre}}(t) * \oint_{\partial\Omega_1} w(\mathbf{x}_0, \mathbf{x}_1) \hat{s}(t - \tau(\mathbf{x}_0, \mathbf{x}_1)) dA_1 \quad (6)$$

and discloses, that the driving function is a weighted superposition of the delayed source signal filtered by  $h_{\text{pre}}(t)$ .

### C. Spatial and Temporal Discretization

A continuous secondary source distribution cannot be implemented with today loudspeaker technology. Hence, a limited number of loudspeakers has to be placed at discrete points on the boundary  $\partial\Omega_0$ . It is furthermore necessary to discretise the virtual secondary source distribution as well due to computational limitations. Hence, the integrals in (1) and (2) transform to sums over finite sets of positions  $\mathbf{x}_0 \in \mathcal{X}_0$  and  $\mathbf{x}_1 \in \mathcal{X}_1$ , respectively. It is known, that this discretisation may lead to spatial aliasing in the reproduced sound field which degrades the reproduction accuracy most prominently at high temporal frequencies. As a rule of thumb, spatial aliasing decreases if the distance between adjacent (virtual) secondary sources decrease.

The temporal sampling of the continuous signal  $\hat{s}(t)$  is conveniently modelled by multiplying it with a Dirac comb. The sampled signal  $\hat{s}_s(t)$  is given as a sequence of samples  $\hat{s}[n] := \hat{s}(nT_s)$ , where  $T_s$  defines the sampling period as the reciprocal of the sample rate  $f_s$ . Thus, the spatially and temporally discretised driving function is given as

$$d_0[\mathbf{x}_0, n] = h_{\text{pre}}[n] * \sum_{\mathbf{x}_1 \in \mathcal{X}_1} w(\mathbf{x}_0, \mathbf{x}_1) \hat{s} \left[ n - \frac{\tau(\mathbf{x}_0, \mathbf{x}_1)}{T_s} \right]. \quad (7)$$

It is evident from this equation that the number of delay operations for one virtual sound source grows linearly with the number of secondary sources  $N_0 := |\mathcal{X}_0|$  and the number of focused sources  $N_1 := |\mathcal{X}_l|$ .

### III. FRACTIONAL DELAY FILTERS

In order to determine the value of a delayed input signal  $\hat{s}_s(t - \tau)$ , the continuous signal  $\hat{s}(t)$  has to be reconstructed from the input sequence  $\hat{s}[n]$  and evaluated at  $t - \tau$ . It is known from sampling theory, that this is achieved by filtering the sequence with an ideal reconstruction filter  $h_{\text{ideal}}(t)$ . This filter is an ideally bandlimited function with a cut-off frequency of  $f_s/2$ . The samples of its discretized impulse response are therefore given by the sinus cardinalis  $h_{\text{ideal}}[n, \tau/T_s] = \text{sinc}(n - \tau/T_s)$ . This filter has a constant phase and group delay of  $\tau/T_s$  over the whole frequency range. However, as this filter is noncausal and of infinite length it is not suited for practical implementations. It has to be reasonably approximated by

$$h_{\text{ideal}} \left[ n, \frac{\tau}{T_s} \right] \approx \delta[n - d_{\text{int}}] * h_{\text{frac}}[n, d_{\text{frac}}]. \quad (8)$$

whereas  $\tau = (d_{\text{int}} + d_{\text{frac}})T_s$  and  $d_{\text{int}} \in \mathbb{Z}$  and  $d_{\text{frac}} \in \mathbb{Q}$ . The FD filter is given as  $h_{\text{frac}}[n, d_{\text{frac}}]$ . The accuracy of the delay interpolation is evaluated by the discrete time frequency-domain error function [13, Eq. (15)]

$$E(e^{j\omega T_s}) = H_{\text{frac}}(e^{j\omega T_s}) - H_{\text{ideal}}(e^{j\omega T_s}). \quad (9)$$

The optimal integer delay  $d_{\text{int}}$  w.r.t. accuracy of the delay interpolation is in general not the closest integer of  $\tau/T_s$ , as a specific FD filter might introduce additional constraints to the fractional part  $d_{\text{frac}}$ .

#### A. Lagrange Interpolation

The optimal filter w.r.t. a maximally flat  $E(e^{j\omega T_s})$  at  $\omega = 0$  [13, Eq. 37] is given by the  $N$ th-order Lagrange interpolator [13, Eq. 42]

$$h_{\text{frac}}[n, d_{\text{frac}}] = \prod_{\substack{m=0 \\ m \neq n}}^N \frac{d_{\text{frac}} - m}{n - m}, \quad n = 0, 1, \dots, N. \quad (10)$$

It was shown by Välimäki [14, Sec. 3.3.6], that the accuracy of this filter is best for  $(N-1)/2 \leq d_{\text{frac}} \leq (N+1)/2$  and hence

$$d_{\text{frac}} = \frac{\tau}{T_s} - d_{\text{int}} = \frac{\tau}{T_s} - \text{round} \left( \frac{\tau}{T_s} - \frac{N}{2} \right). \quad (11)$$

Another well-known reconstruction method from analog-to-digital converters, named zero-order hold (ZOH), is closely related to the Lagrange interpolator of zeroth order: Instead of rounding  $\tau/T_s$  to nearest integer delay (cf. 11 for  $N = 0$ ) ZOH always rounds up the next larger integer delay. ZOH can therefore be regarded as a Lagrange interpolator with a sub-optimal choice of  $d_{\text{int}}$  w.r.t. accuracy. It however does not lead to causality issues as  $d_{\text{int}}$  never underestimates the actual delay.

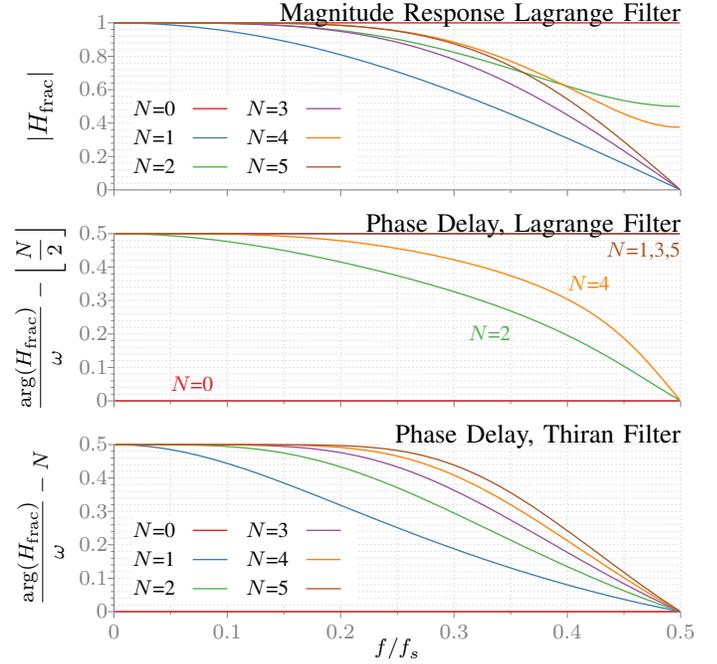


Fig. 2: The upper two graphs show the magnitude response and the phase delay of Lagrange interpolation filters of different order  $N$  for a fractional delay of  $d_{\text{frac}} = 0.5 + \lfloor N/2 \rfloor$ . The bottom depicts the phase delay for Thiran filters for a fractional delay of  $d_{\text{frac}} = 0.5 + N$ .

#### B. Thiran's Allpass Infinite Impulse Response (IIR) Filter

The benefit from using IIR allpass filters compared to Finite Impulse Response (FIR) filters are a lower number of multiplications needed [13, p. 46] and their unit magnitude over all frequencies. However, their drawback becomes evident, if the filter coefficients have to be changed over time, e.g. due to a time-varying delay: The state variables of the filter store intermediate results of the filtering. As these might belong to a former coefficient set, so-called *transient errors* are likely to occur [14, Sec. 3.5]. Although such problems also occur for FIR filters, their handling is more challenging for IIR filters [14, Sec. 3.5.2].

The coefficients of an  $N$ th-order allpass filter are given by its z-transform

$$H_{\text{frac}}(z) = \frac{z^{-N} A(z^{-1})}{A(z)} = \frac{\sum_{n=0}^N a_n z^{n-N}}{\sum_{n=0}^N a_n z^{-n}}. \quad (12)$$

Thiran's Filter [15] is the only known IIR FD Filter, whose coefficients can be given in closed-form [13, p. 49], namely

$$a_n = (-1)^n \binom{N}{n} \prod_{m=0}^N \frac{d_{\text{frac}} - N - m}{d_{\text{frac}} - N + m + n}, \quad n = 0, 1, \dots, N. \quad (13)$$

The filter is optimal w.r.t. a maximally flat group delay  $E(e^{j\omega T_s})$  at  $\omega = 0$ . It was experimentally shown in [14, Sec. 3.4.3/4] that this filter is stable for  $d_{\text{frac}} > N - 1$  and reaches close-to-optimal accuracy for  $N - 1/2 \leq d_{\text{frac}} \leq N + 1/2$ , hence

$$d_{\text{frac}} = \frac{\tau}{T_s} - d_{\text{int}} = \frac{\tau}{T_s} - \text{round} \left( \frac{\tau}{T_s} \right) + N. \quad (14)$$

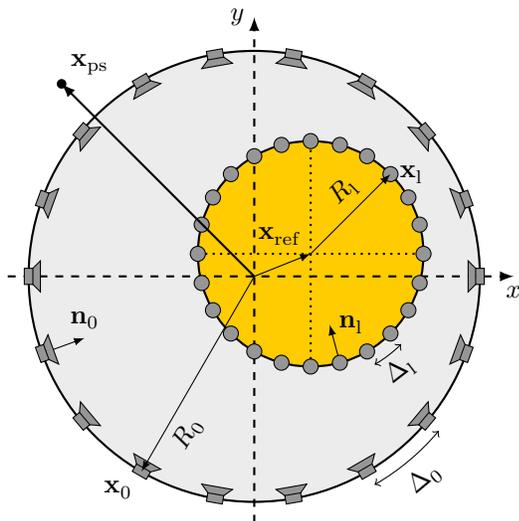


Fig. 3: Reproduction setup for evaluation

### C. Upsampling of the Input Signals

As the number of delay operation increases linearly with  $N_0$  and  $N_1$ , treating each delay operation separately leads to tremendous computational effort, especially for high filter orders. It was mentioned by [16, Sec. 3] in the context of WFS, that it is sensible to apply a delay independent pre-processing to the input signals. Among other techniques, an upsampling of the source signal was suggested, as the FD filters presented in Sec. III-A and III-B achieve the highest accuracy for low frequencies. The individual delay operation is then applied to the upsampled signal using FD filters of lower order.

Increasing the sample rate about an integer ratio  $R$  is efficiently implemented using polyphase structures. Within in this treatise, a linear-phase FIR filter with  $R \cdot 64$  taps is used to interpolate the upsampled signal. It is designed using the algorithm of [17].

## IV. EVALUATION

### A. Experimental Setup

A circular secondary source distribution of  $R_0 = 1.5$  m radius with an equi-angular spacing of  $\Delta_0 = 2\pi/N_0$  is used (see. Fig. 3). The distribution is centred at coordinates' origin. The local listening area is bounded by a circle of radius  $R_1 = 0.3$  m centred at  $\mathbf{x}_{\text{ref}}$ . The focused sources are positioned along its boundary with an equi-angular spacing of  $\Delta_1 = 2\pi/N_1$ . The desired sound field  $S(\mathbf{x})$  is given as a virtual point source located at  $\mathbf{x}_{\text{ps}} = [0, 2.5, 0]^T$  m. A reproduction setup with  $N_0 = 256$  loudspeakers driven by LWFS incorporating  $N_1 = 256$  focused sources is used as a reference setup. Although such parametrisation is infeasible for many practical applications it allows for studying the effect of fractional delays isolated from spatial aliasing. The sample rate  $f_s$  is set to 44.1 kHz. As the reference, i.e. high accuracy, FD method an upsampling of the input signal about  $R = 8$  together with a 9th-order Lagrange interpolator for the individual delay operation is chosen.

### B. Spectral Properties of Reproduced Sound Field

1) *Influence of FD Filter without Spatial Aliasing:* In order to investigate the influence of the FD filters on the spectral properties of the reproduced sound field, the sound field's spectrum is evaluated at  $\mathbf{x}_{\text{ref}}$ . Fig. 4 shows it for different FD methods and two different  $\mathbf{x}_{\text{ref}}$ : As expected, deviations of the magnitude from the reference method filter are generally reduced when using a filter with higher complexity. Note, that the response for the reference method is not flat at low frequencies as the driving functions of LWFS are based on high-frequency approximations. For the Lagrange filter of third order, the lowpass characteristics of its magnitude response shown in Fig. 2 lead to a lowpass filtering of the reproduced sound field. As Thiran filters have a unit magnitude, the reproduced sound field is mainly influenced by phase distortions (cf. bottom graph in Fig. 2). Hence, more complex interference patterns are observed in Fig. 4. For both filter types an upsampling of the input signal about a factor of 2 leads to significant increase of accuracy. The sub-optimal choice of the delay's integer part (cf. Sec. III-A) of ZOH negatively effects its accuracy compared to zeroth-order Lagrange interpolator.

2) *Interaction between FD Filters and Spatial Aliasing:* A comparison of ZOH and the reference FD method for different reproduction setups is shown in Fig. 5: As the number of loudspeakers and/or the number of focused source decreases, more artefacts are introduced due to spatial aliasing. These artefacts are recognizable by strong amplitude fluctuations. For  $N_l = 128$ , ZOH introduces stronger fluctuations compared to the reference FD filter. For  $N_l = 64$  however, it is hard to determine, whether reference filter outperforms the ZOH, as the aliasing artefacts dominate the spectra. These findings agree with the results of the listening test for WFS [7], which showed that the test subjects were not able to distinguish ZOH from any higher order interpolation scheme. As sound fields reproduced by typical WFS setups only remain aliasing-free for frequency below 2 kHz, this indistinguishability is most probably related to spatial aliasing.

## V. CONCLUSION & FUTURE WORK

This paper presents a comparison of different delay interpolation techniques for LWFS. The study focuses on the reproduction of stationary scenes. The results serve as a prerequisite for the synthesis of dynamic scenarios where delay interpolation is mandatory to avoid audible artefacts. Independent of the FD filter type, an oversampling of the input signal significantly improves the reproduction accuracy. In terms of computational complexity this is an important finding, as the computational effort of oversampling is independent of the number of required delay operations. After upsampling the signals by a factor of two, a FD filter of third order is suitable. The artefacts of delay interpolation superimpose to the spatial aliasing artefacts. Since both are most prominent for high frequencies, the effects of delay interpolation seem to be not so relevant for higher frequencies. The presented evaluation shows that a real-time realisation of listener-tracked

LWFS is feasible using resampling and low-order FD filters. Further work includes the evaluation of dynamic scenarios and listening tests.

## VI. ACKNOWLEDGEMENTS

This research has been supported by EU FET grant TWO!EARS, ICT-618075.

## REFERENCES

- [1] 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.
- [2] 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.
- [3] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, apr 2009, pp. 93–96.
- [4] 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: Audio Engineering Society (AES), 2011.
- [5] 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.
- [6] A. Franck, "Efficient Algorithms for Arbitrary Sample Rate Conversion with Application to Wave Field Synthesis," Ph.D. dissertation, Technische Universität Ilmenau, 2011.
- [7] J. Ahrens, M. Geier, and S. Spors, "Perceptual Assessment of Delay Accuracy and Loudspeaker Misplacement in Wave Field Synthesis," in *128th Convention of the Audio Engineering Society*, London, UK, 2010.
- [8] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. London: Academic Press, 1999.
- [9] F. Winter and S. Spors, "Physical Properties of Local Wave Field Synthesis using Linear Loudspeaker Arrays," in *Audio Engineering Society 138th Convention*, Warsaw, Poland, 2015.
- [10] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, 1993.
- [11] E. N. G. Verheijen, "Sound Reproduction by Wave Field Synthesis," Ph.D. dissertation, Delft University of Technology, 1997.
- [12] E. W. Start, "Direct Sound Enhancement by Wave Field Synthesis," Ph.D. dissertation, Delft University of Technology, 1997.
- [13] T. I. Laakso, V. Valimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay [FIR/all pass filters design]," *Signal Processing Magazine, IEEE*, vol. 13, no. 1, pp. 30–60, jan 1996.
- [14] V. Välimäki, "Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters," Ph.D. dissertation, Helsinki University of Technology, 1998.
- [15] J.-P. Thiran, "Recursive digital filters with maximally flat group delay," *Circuit Theory, IEEE Transactions on*, vol. 18, no. 6, pp. 659–664, nov 1971.
- [16] A. Franck, K. Brandenburg, and U. Richter, "Efficient Delay Interpolation for Wave Field Synthesis," in *Proc. of 125th Aud. Eng. Soc. Conv.*, San Francisco, USA, 2008.
- [17] J. H. McClellan and T. W. Parks, "A personal history of the Parks-McClellan algorithm," *Signal Processing Magazine, IEEE*, vol. 22, no. 2, pp. 82–86, 2005.

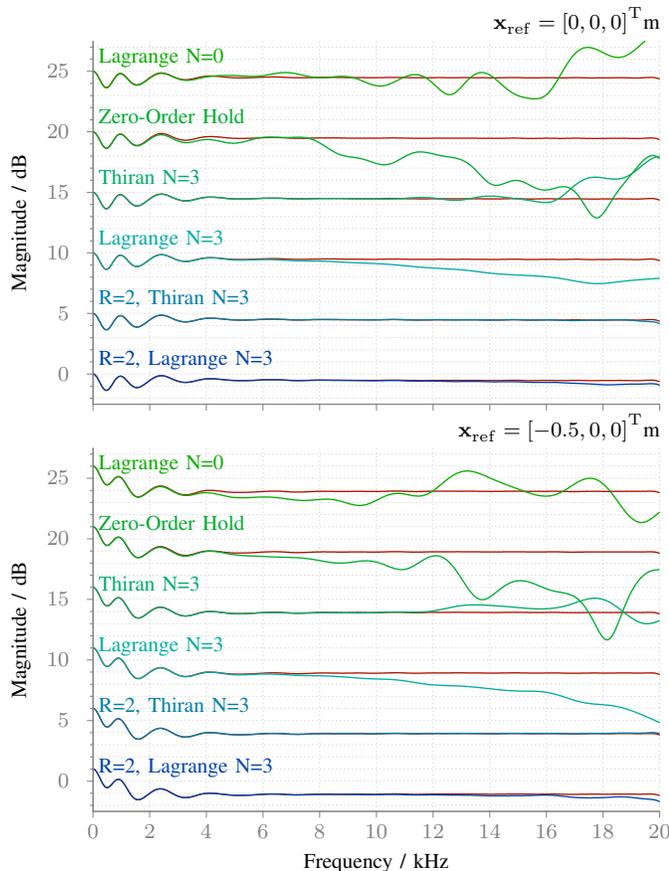


Fig. 4: Magnitude response of the reproduced sound field at  $\mathbf{x}_{\text{ref}}$  using the reference reproduction setup described in Sec. IV-A in combination with different FD interpolation methods (blue to green). The reference interpolation method is shown in red. All magnitude responses have been normalized to their respective values at  $f = 0$  Hz and are shifted incrementally by 5 dB to improve visibility.

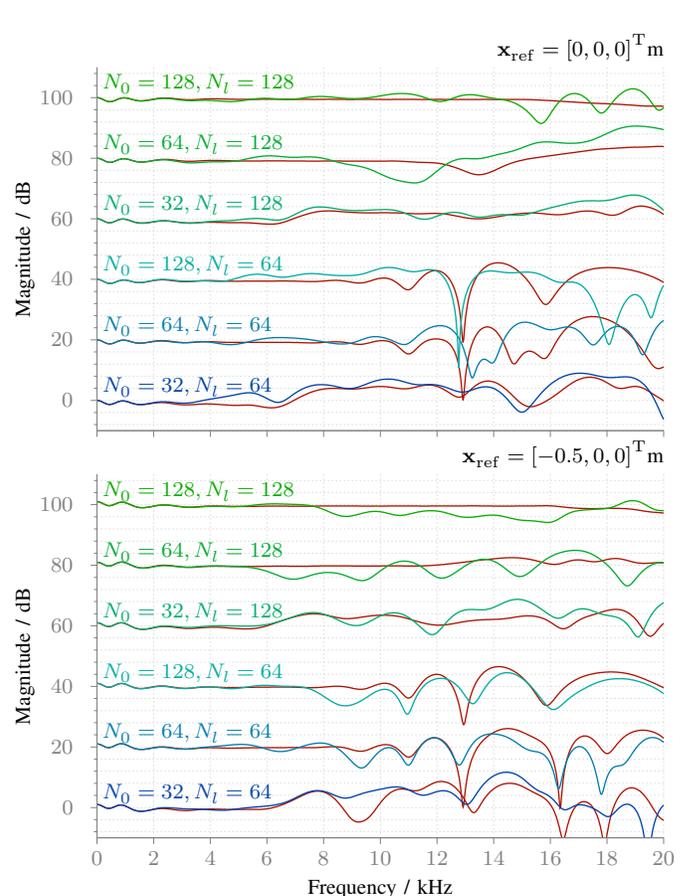


Fig. 5: Magnitude response of the reproduced sound field at  $\mathbf{x}_{\text{ref}}$  for ZOH (blue to green) and the reference FD method (red) for different reproduction setups. All magnitude responses have been normalized to their respective values at  $f = 0$  Hz and are shifted incrementally by 20 dB to improve visibility.