

CONTINUOUS MEASUREMENT OF SPATIAL ROOM IMPULSE RESPONSES USING A NON-UNIFORMLY MOVING MICROPHONE

Nara Hahn and Sascha Spors*

Institute of Communications Engineering
University of Rostock, Germany

ABSTRACT

The room impulse responses at multiple receiver positions can be measured efficiently with a continuously moving microphone. The acoustic system is periodically excited by a self-orthogonal signal, called perfect sequence, and the microphone captures the sound field on a pre-defined path. As shown in recent studies by the authors, the captured signal constitutes a spatio-temporal sampling of the sound field, and the impulse responses can be obtained by a spatial interpolation. So far, a uniformly moving microphone was mainly considered for the measurement of spatial room impulse responses. In this paper, the method is applied to non-uniformly moving microphones thereby addressing more general cases. The proposed method is evaluated by numerical simulations where the spatial room impulse responses on a circle are measured using a microphone with a fluctuating angular speed. The accuracy of the impulse responses are compared for varying interpolation orders.

1. INTRODUCTION

The spatial structure of a sound field can be captured by measuring the impulse responses or by recording the acoustic signals at multiple receiver positions. For instance, the spatial room impulse responses are measured with a microphone array for sound field analysis or spatial sound reproduction [1, 2]. Similarly, a large number of acoustic impulse responses are measured to investigate the directional properties of sound sources, receivers, or the combinations of these [3, 4, 5]. Since the acquired spatial resolution scales with the number of receiver points, a large number of measurements are required, typically ranging from several hundreds to thousands. During such a measurement, identical processes are repeated many times, while giving a slight change to the receiver position or the source/receiver orientation.

Instead of performing the measurement in a static configuration, a dynamic set-up can be considered in order to accelerate the process. For example, head-related impulse responses are measured either by continuously rotating the head-and-torso simulator, or by moving the loudspeaker on a circle [6, 7]. Several approaches for such a continuous measurement have been introduced and used for different types of spatial impulse responses [8, 6, 9, 10, 11, 12]. It was pointed out in [13] that the signal captured by the microphone can be interpreted as being spatio-temporally sampled from the sound field. Therefore, the time-varying system identification problem can be regarded as a sound field interpolation problem.

Recently, the authors carried out a series of studies on the continuous measurement of impulse responses on a circle [13, 14,

15]. Constant angular speed was considered which constitutes an equiangular sampling on the circle. Based on the spatial bandwidth of the sound field in the circular harmonics domain, an anti-aliasing condition for the angular speed was derived [13]. The individual impulse responses were computed using a spatial interpolation, and the results were compared with static measurements [13] as well as with other existing approaches for dynamic measurement [14]. It was shown that higher-order interpolations generally achieve better performance provided that the anti-aliasing condition is fulfilled [15]. If the microphone exceeds the maximum allowable angular speed, the accuracy of the impulse responses decreases due to spatial aliasing.

In this paper, the same approach is applied for non-uniformly moving microphones. The microphone is still assumed to move on a pre-defined trajectory but the speed is varied over time. It is of interest to find out the relation of non-uniform motion and the resulting distribution of spatial sampling points. The impulse responses are estimated by using an interpolation of different orders. The performance of the proposed approach is examined by numerical simulations.

2. SPATIAL SAMPLING

For a continuous measurement, the system is typically excited by a periodic perfect sequence $\psi(n) = \psi(n + N)$ which exhibits a self-orthogonality [16],

$$\sum_{m=0}^{N-1} \psi(n+m)\psi(n) = \sigma_{\psi}^2 \sum_{\mu \in \mathbb{Z}} \delta(m + \mu N), \quad (1)$$

where N denotes the period of $\psi(n)$, σ_{ψ}^2 the energy within one period, and $\delta(\cdot)$ the Dirac delta function. To avoid temporal aliasing, the period N has to be longer than the longest impulse response of the system [7]. For brevity, $\sigma_{\psi}^2 = 1$ is assumed in the remainder.

Assuming a finite impulse responses (FIR) model, the time-domain sound field $p(\mathbf{x}, n)$ reads

$$p(\mathbf{x}, n) = \sum_{k=0}^{N-1} \psi(n-k)h(\mathbf{x}, k), \quad (2)$$

where $\mathbf{x} = [x, y, z]^T$ denotes the position, n the discrete-time index, and $h(\mathbf{x}, k)$ the impulse response. Note that $p(\mathbf{x}, n)$ exhibits the same periodicity as the excitation signal $\psi(n)$.

For a noiseless measurement, the impulse response is equal to the length- N circular cross-correlation of $\psi(n)$ with the signal re-

*This research was supported by a grant of the Deutsche Forschungsgemeinschaft (DFG) SP 1295/7-1.

ceived at \mathbf{x} ,

$$h(\mathbf{x}, m) = \sum_{n=0}^{N-1} p(\mathbf{x}, n) \psi(n - m) \quad (3)$$

$$= \sum_{n=0}^{N-1} p(\mathbf{x}, n + n_0) \psi(n + n_0 - m), \quad (4)$$

which can be proven by inserting (3) into (2) and exploiting (1). The second equality (4) states that the cross-correlation can be computed for any N consecutive samples of $p(\mathbf{x}, n)$ and $\psi(n)$.

Assume that an omni-directional microphone moves on a path defined by $\tilde{\mathbf{x}}(n)$. The signal captured by the microphone then reads

$$s(n) = p(\tilde{\mathbf{x}}(n), n), \quad (5)$$

which is a slice of the sound field [6]. By exploiting the time-periodicity of $p(\mathbf{x}, n)$, the captured signal $s(n)$ can be decomposed into N sequences,

$$s_\nu(l) = s(\nu + lN) \quad (6)$$

$$= p(\tilde{\mathbf{x}}(\nu + lN), \nu + lN) \quad (7)$$

$$= p(\tilde{\mathbf{x}}(\nu + lN), \nu), \quad (8)$$

for $\nu = 0, \dots, N-1$. The ν -th sequence thus constitutes a spatial sampling of the sound field at time $n = \nu$. This is shown in Fig. 1 where the microphone moves on a circular trajectory. Note that the distribution of the sampling positions \bullet differs for each ν . If the total number of samples is denoted by L , the effective number of spatial samples is $\frac{L}{N}$.

3. SPATIAL INTERPOLATION

Since only one sample of $p(\mathbf{x}, n)$ is captured at each position $\tilde{\mathbf{x}}(n)$, (3) cannot be directly used for the computation of $h(\mathbf{x}, n)$. Instead, the original sound field has to be interpolated from $s(n)$ [10],

$$\hat{p}(\mathbf{x}, \nu) = \sum_{l \in \mathbb{Z}} g_{\nu l}(\mathbf{x}) p(\tilde{\mathbf{x}}(\nu + lN), \nu) \quad (9)$$

$$= \sum_{l \in \mathbb{Z}} g_{\nu l}(\mathbf{x}) s_\nu(l), \quad (10)$$

for $\nu = 0, \dots, N-1$, where $\hat{p}(\mathbf{x}, \nu)$ denotes the estimate of the sound field and $g_{\nu l}(\mathbf{x})$ the interpolation coefficients.

The process of computing the impulse responses is depicted in Fig. 2. A negative delay z^ν is applied to $s(n)$ and $\tilde{\mathbf{x}}(n)$, followed by a factor-of- N decimation. In practice, this part can be performed quite simply by using `reshape` in Matlab or `numpy.reshape` in Python. The sequences $s_\nu(l)$ will be stacked in a matrix (array). For a set of target positions \mathbf{x}_k , the sound field is estimated by interpolating the individual sequences. While one-dimensional interpolation suffices for linear or circular trajectories, two- or three-dimensional interpolations might be desirable depending on the geometry of the microphone path. The estimates $\hat{p}(\mathbf{x}_k, \nu)$ obtained from the ν -th sequences then compose N consecutive samples of the sound field. Finally, $\hat{h}(\mathbf{x}_k, n)$ is obtained by computing the circular cross-correlation of $\hat{p}(\mathbf{x}_k, n)$ and $\psi(n)$, or equivalently the circular convolution of $\hat{p}(\mathbf{x}_k, n)$ and $\psi(-n)$ as indicated by $*$ in Fig. 2.

It is implicitly assumed in Fig. 2 that $s(n)$ and $\tilde{\mathbf{x}}(n)$ have the same sampling frequency, which is not necessarily the case in practice. For instance, an optical motion tracking system exhibits a rate

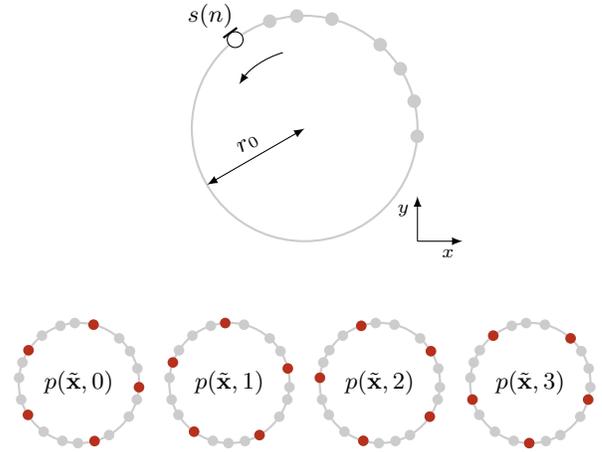


Figure 1: Spatio-temporal sampling in a continuous measurement. The microphone \circ moves on a circle of radius r_0 at a non-uniform angular speed (top). The gray circles \bullet indicate the locations where the signal is captured. For a periodic excitation ($N = 4$), the decimated signal $s_\nu(l), \nu = 0, \dots, 3$, indicated by \bullet , corresponds to the spatial sampling of the sound field $p(\tilde{\mathbf{x}}, \nu)$ (bottom).

in the order of 100 Hz, while typical audio sampling rate is much higher, e.g. $f_s = 44.1$ kHz. Though, as far as the microphone position is sampled with no spatial aliasing, it can be upsampled to the same sampling rate.

The trajectory and the speed of the microphone have to be chosen carefully by taking the spatial bandwidth of the sound field into account. The spatial distribution of each $\tilde{\mathbf{x}}_\nu(l)$ must be sufficiently dense so that the interpolation accuracy is not affected by spatial aliasing. In [13], an anti-aliasing condition was introduced for the continuous measurement on a circle with a uniformly moving microphone. Based on the approximated spatial bandwidth of the sound field in the circular harmonics domain, the maximum allowable speed Ω_0 reads [13, Eq. (14)],

$$\Omega \leq \Omega_0 = \frac{c}{r_0 N}, \quad (11)$$

where c denotes the speed of sound and r_0 the radius. It was observed in [15] that higher-order interpolation generally improves the accuracy of the impulse responses, if the anti-aliasing condition is fulfilled and the noise is not severe. In this particular case, the periodic sinc interpolation is the analytic solution [6, 14].

Although a uniformly moving microphone is of interest from a theoretical perspective, it is not trivial to perform such a measurement in practice. If either the total duration within a rotation $L = \frac{2\pi}{\Omega} \times f_s$ or the effective number of sampling points $\frac{L}{N}$ is not exactly an integer, the sampled positions exhibit a non-uniform distribution on the circle. In such cases, the sinc interpolation cannot be used straightforwardly, and the anti-aliasing condition (11) does not apply any more. In this paper, the Lagrange interpolation is used instead, which considers also non-uniform sampling points.

4. EVALUATION

In this section, a continuous measurement of spatial impulse responses is simulated. The sound field of a Dirac-shaped plane wave

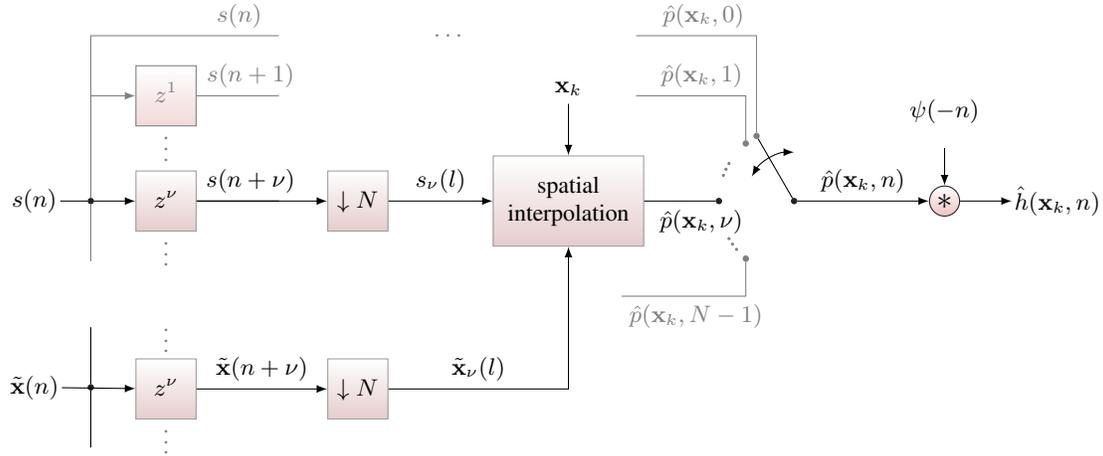


Figure 2: Block diagram illustrating the time-varying system identification in a continuous measurement. The captured signal $s(n)$ and position $\tilde{\mathbf{x}}(n)$ is decomposed into N sequences, which are obtained by applying a negative delay and a factor-of- N decimation. The time-domain sound field $\hat{p}(\mathbf{x}_k, n)$ is estimated by interpolating the sampled values. The corresponding time-variant impulse response $\hat{h}(\mathbf{x}_k, n)$ is finally obtained by convolving (denoted by $*$) $\hat{p}(\mathbf{x}_k, n)$ with $\psi(-n)$.

is considered that propagates parallel to the xy -plane with an angle of $\phi_{pw} = 270^\circ$. Assuming a free-field condition, the continuous-time impulse response on the circle reads

$$h(\phi, \tau) = \delta\left(\tau - \frac{r_0}{c} \cos(\phi - \phi_{pw})\right), \quad (12)$$

where ϕ denotes the polar angle of the microphone. In the simulation, the sampling frequency is $f_s = 8$ kHz and the speed of sound is assumed to be $c = 343$ m/s. Non-integer delays are implemented by sinc functions windowed by a Blackman window of length 64 [17]. The plane wave is driven by a periodic perfect sweep with $N = 800$ (0.1 s) [18]. A microphone moves on a circle with radius of $r_0 = 0.5$ m. The polar angle of the microphone is

$$\phi(t) = \Omega t + \alpha \sin 2\pi t, \quad (13)$$

where Ωt corresponds to the uniform part and $\alpha \sin 2\pi t$ to a periodic fluctuation determined by the parameter α . In Fig. 3 (left), the first few seconds of $\phi(t)$ are shown. The straight line ($\alpha = 0$ indicated by \rightarrow) corresponds to a uniform motion. According to (11), the anti-aliasing angular speed (11) is $\Omega_0 = 49.1^\circ \cdot \text{s}^{-1}$ for this configuration. Although this only holds for a uniformly moving microphone, it is considered as a reference. The simulation is performed for a sufficiently lower angular speed $\Omega = 22.5^\circ \cdot \text{s}^{-1}$. The total length of the captured signal is $L = \frac{360}{\Omega} \times f_s = 128000$ for all conditions. To interpolate the spatial samples, Lagrange polynomials of different orders are used [19, Eq. (25.2.2)],

$$g_{\nu l}(\phi) = \prod_{j \neq l} \frac{\phi - \phi_{\nu j}}{\phi_{\nu l} - \phi_{\nu j}},$$

where $\phi_{\nu l}$ denotes the polar angle in which the $(\nu + Nl)$ -th sample is captured. The impulse responses are computed for 360 equiangular points on the circle, $\phi_k = 0, 1, \dots, 359^\circ$. The accuracy of individual impulse responses is examined by the (normalized) system distance defined as

$$\text{SD}(\phi_k) = \left(\frac{\sum_{n=0}^{N-1} |h(\phi_k, n) - \hat{h}(\phi_k, n)|^2}{\sum_{n=0}^{N-1} |h(\phi_k, n)|^2} \right)^{\frac{1}{2}} \quad (14)$$

where $\hat{h}(\phi_k, n)$ denotes the estimate of $h(\phi_k, n)$.

In general, the system distance exhibits angular dependencies (not shown here). The accuracy increases around $\phi_k = 90, 270^\circ$ where the time-of-arrival $\frac{r_0}{c} \cos(\phi(t) - \phi_{pw})$ is piecewise constant and the system is nearly time invariant. In Fig. 3 (center), the system distance $\text{SD}(\phi_k)$ is averaged over ϕ_k to compare the overall performance between different conditions. For $\alpha = 0$ (uniform case), the higher the interpolation order, the more accurate are the impulse responses. The system distance converges to -67 dB. This is not the case for $\alpha \neq 0$ where a slight increase in $\text{SD}(\phi_k)$ is observed for an interpolation order of 32. This suggests that the time-variability of the system has to be taken into account in choosing the interpolation order. Unnecessarily high interpolation could rather degrade the performance.

While stronger fluctuations (larger α) generally lead to an increased system distance, an exception is observed for $\alpha = 12$. Despite the strongest fluctuation, the performance is significantly better than $\alpha = 4, 8$ and almost comparable to $\alpha = 2$. This can be explained by examining the distribution of the spatial sampling points as follows. For every ν , the decimated angles $\phi_\nu(l)$ are sorted in ascending order and the difference between consecutive angles $\Delta\phi$ is computed. The distribution of the angular spacings $\Delta\phi$ is shown in Fig. 3 (right). The distribution for $\alpha = 0$ has a Dirac shape, since $\Delta\phi = \frac{360N}{L} = 2.25^\circ$ is constant. For $\alpha = 2, 4$, $\Delta\phi$ exhibits two peaks on each side of the distribution. The spread increases with α which agrees with the standard deviations indicated by σ in Fig 3 (right). The left peak is attributed to the change of direction where the angular speed (slope of $\phi(t)$) is close to 0. Here, the spatial samples are closely distributed and thus resulting in a small $\Delta\phi$. The peak on the right is related to the maximum speed where the slope of $\phi(t)$ is steepest, i.e. $t = 0, 1, 2$ s in Fig. 3 (left). The large $\Delta\phi$ causes an increase in interpolation errors. Interestingly, for $\alpha = 12$, the right peak disappears and the distribution is more or less uniform in $\Delta\phi \in [0, 4]$. The fluctuation is so strong such that the sampled angles at the maximum speed coincide with those at the minimum speed. The dense distribution in the former compensates the sparse distribution in the latter. This is the ben-

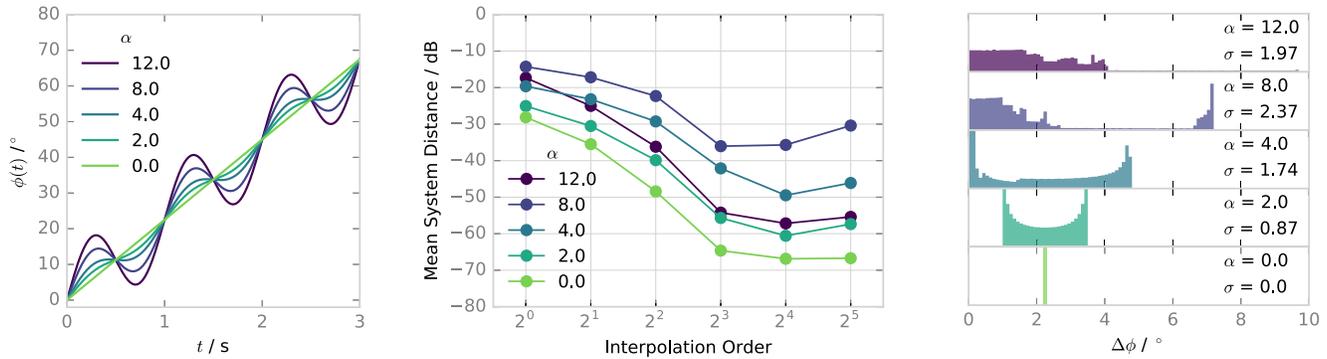


Figure 3: Continuous measurement on a circle using a non-uniformly moving microphone. Left: The polar angle of the microphone is $\phi(t) = \Omega t + \alpha \sin 2\pi t$ with $\Omega = 22.5^\circ \cdot \text{s}^{-1}$ and $\alpha = 0, 2, 4, 8, 12$. Center: For each α , the impulse responses are estimated by using the Lagrange polynomials of different orders $2^0, \dots, 2^5$. The system distance $\text{SD}(\phi_k)$ is averaged over ϕ_k . Right: Histograms showing the distribution of the angular spacings $\Delta\phi$ between the spatial samples. The standard deviation of the distribution is denoted by σ .

efit of the proposed approach where the spatial distribution of the captured sound field is explicitly considered.

5. CONCLUSION

The continuous measurement method based on spatial interpolation was applied to non-uniformly moving microphones. The signal captured by the microphone was considered as spatio-temporally sampled from the sound field. The original sound field was estimated by interpolating the non-uniform samples. The proposed approach was validated by numerical simulations where the spatial room impulse responses were measured on a circle. It was shown that the continuous measurement technique can be applied in more general cases, for instance, when the position of the microphone cannot be controlled perfectly.

The perceptual properties of the impulse responses are still under investigation. Once the associated physical attributes are identified, it is expected that the measurement speed can be further increased without causing perceivable degradations. The extension to two- or three-dimensional cases is also left as future work.

6. REFERENCES

- [1] B. Rafaely, *Fundamentals of Spherical Array Processing*. Springer, 2015.
- [2] A. Kuntz, *Wave Field Analysis Using Virtual Circular Microphone Arrays*. Verlag Dr. Hut, 2009.
- [3] F. Zotter, “Analysis and Synthesis of Sound-radiation with Spherical Arrays,” Ph.D. dissertation, University of Music and Performing Arts, Graz, Austria, 2009.
- [4] H. Wierstorf, M. Geier, A. Raake, and S. Spors, “A Free Database of Head Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances,” in *Proc. of 130th Audio Engineering Society (AES) Convention*, London, UK, May 2011.
- [5] H. Morgenstern, B. Rafaely, and F. Zotter, “Theory and Investigation of Acoustic Multiple-input Multiple-output Systems Based on Spherical Arrays in a Room,” *The Journal of the Acoustical Society of America (JASA)*, vol. 138, no. 5, pp. 2998–3009, 2015.
- [6] T. Ajdler, L. Sbaiz, and M. Vetterli, “Dynamic Measurement of Room Impulse Responses Using a Moving Microphone,” *The Journal of the Acoustical Society of America (JASA)*, vol. 122, no. 3, pp. 1636–1645, 2007.
- [7] C. Antweiler and G. Enzner, “Perfect Sequence LMS for Rapid Acquisition of Continuous-azimuth Head Related Impulse Responses,” in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, Oct. 2009, pp. 281–284.
- [8] E. M. Hulsebos, “Auralization Using Wave Field Synthesis,” Ph.D. dissertation, Delft University of Technology, Delft, The Netherlands, 2004.
- [9] C. Antweiler, A. Telle, P. Vary, and G. Enzner, “Perfect Sweep NLMs for Time-variant Acoustic System Identification,” in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012.
- [10] N. Hahn and S. Spors, “Identification of Dynamic Acoustic Systems by Orthogonal Expansion of Time-variant Impulse Responses,” in *Proc. of the 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, Athens, Greek, May 2014.
- [11] F. Katzberg, R. Mazur, M. Maass, P. Koch, and A. Mertins, “Measurement of Sound Fields Using Moving Microphones,” in *Proc. of the 42nd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, New Orleans, USA, Mar. 2017.
- [12] J.-G. Richter and J. FelsHahn, “Influence of Continuous Subject Rotation During HRTF Measurements,” in *Proc. of the 43rd German Annual Conference on Acoustics (DAGA)*, Kiel, Germany, Mar. 2017.
- [13] N. Hahn and S. Spors, “Continuous Measurement of Impulse Responses on a Circle Using a Uniformly Moving Microphone,” in *Proc. of the European Signal Processing Conference (EUSIPCO)*, Nice, France, Aug. 2015.
- [14] —, “Comparison of Continuous Measurement Techniques for Spatial Room Impulse Responses,” in *Proc. of the European Signal Processing Conference (EUSIPCO)*, Budapest, Hungary, Aug. 2016.
- [15] —, “Spatial Aliasing in Continuous Measurement of Spatial Room Impulse Responses,” in *Proc. of the 43rd German Annual Conference on Acoustics (DAGA)*, Kiel, Germany, Mar. 2017.
- [16] H. D. Lüke, *Korrelationskssignale*. Springer, 1992.
- [17] T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine, “Splitting the Unit Delay,” *Signal Processing Magazine, IEEE*, vol. 13, no. 1, pp. 30–60, 1996.
- [18] N. Aoshima, “Computer-generated Pulse Signal Applied for Sound Measurement,” *The Journal of the Acoustical Society of America (JASA)*, vol. 69, no. 5, pp. 1484–1488, 1981.
- [19] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Courier Corporation, 1964.