



# Real-time sound synthesis of pass-by noise: comparison of spherical harmonics and time-varying filters

Mansour Alkmim<sup>1,2</sup> , Guillaume Vandernoot<sup>3</sup> , Jacques Cuenca<sup>1,\*</sup> , Karl Janssens<sup>1</sup>, Wim Desmet<sup>2,4</sup>, and Laurent De Ryck<sup>1</sup> 

<sup>1</sup>Siemens Digital Industries Software, Interleuvenlaan 68, 3001 Leuven, Belgium

<sup>2</sup>KU Leuven, Department of Mechanical Engineering, Celestijnenlaan 300 B, 3001 Leuven, Belgium

<sup>3</sup>Siemens Digital Industries Software, 150 Avenue de la Republique, 92320 Chatillon, France

<sup>4</sup>DMMS Core Lab, Flanders Make, Gaston Geenslaan 8, 3001 Heverlee, Belgium

Received 3 May 2023, Accepted 16 June 2023

**Abstract** – This paper proposes and compares two sound synthesis techniques to render a moving source for a fixed receiver position based on indoor pass-by noise measurements. The approaches are based on the time-varying infinite impulse response (IIR) filtering and spherical harmonics (SH) representation. The central contribution of the work is a framework for realistic moving source sound synthesis based on transfer functions measured using static far-field microphone arrays. While the SHs require a circular microphone array and a free-field propagation (delay, geometric spread), the IIR filtering relies on far-field microphones that correspond to the propagation path of the moving source. Both frameworks aim to provide accurate sound pressure levels in the far-field that comply with standards. Moreover, the frameworks can be extended to additional sources and filters (e.g. sound barriers) to create different moving source scenarios by removing the room size constraint. The results of the two sound synthesis approaches are preliminary evaluated and compared on a vehicle pass-by noise dataset and it is shown that both approaches are capable of accurately and efficiently synthesize a moving source.

**Keywords:** Sound synthesis, Moving source, Spherical harmonics, Time-varying filters

## 1 Introduction

Due to the associated health risks, international regulations specify the maximum allowed noise level for road vehicles [1, 2]. Automotive companies and original equipment manufacturers need to ensure that the vehicles' target noise levels comply with the imposed regulations in the early design stages since major changes are no longer possible at the end of the development. Hence, the ability to predict noise levels as well as sound quality criteria from early phases of vehicles' development has become a recurrent request.

The indoor pass-by noise (PBN) test [3] provides sound pressure levels comparable to those from outdoor tests and has notable advantages [4]. The vehicle placed above a drum roller in a hemi-anechoic chamber is fully controllable and therefore able to produce consistent and repeatable results. Even though the required number of sensors is considerably larger than for exterior PBN, the complexity of the measurement is reduced since there is no need for wireless communication, light barriers, speed radar,

weather stations, and telemetry systems. Additionally, the component-based transfer path analysis combined with the acoustic source quantification (ASQ) technique [5] allows for the characterization of individual components on a test bench and the assembly of the components virtually, enabling the quantification of the sound pressure level for homologation purposes as well as listening tests for different vehicle configurations at earlier stages of design.

Besides the noise level quantification, automotive companies and original equipment manufacturers are increasingly interested in the subjective assessment of the exterior sounds produced by the vehicles, particularly with the resurgence of electric vehicles and the corresponding need for acoustic vehicle alerting systems. This requires the virtual vehicle assembly to be realistic, both quantitatively and perceptually.

A poor synthesis reconstruction can drastically affect any perceptual cues from the audio and affects the sound quality assessments. Many attempts of the accurate synthesis of a traffic event have been proposed using simulation techniques such as time-domain finite differences [6], pseudo-spectral methods [7] and binaural impulse responses [8]. In the first two, the time-domain simulations are still

\*Corresponding author: [jacques.cuenca@siemens.com](mailto:jacques.cuenca@siemens.com)

rather simplistic and computationally expensive for real-time applications. In the latter, the overlap-add and cross-fade approaches did not generate a satisfactory result since clicks, artifacts, and sound modulations were present in the synthesized audio. Measured sources and transfer functions were employed in [9] to develop general traffic events, but the result showed audible clicks and a smearing effect in the time-frequency spectrum. Other measured-based approaches have been proposed [10–14] which rely on the decomposition of recorded signal into propulsion and tire noise components. The sound propagation was done using inverse short-time Fourier transform or by point-to-point sound propagation model such [15]. Alternatively, Pieren et al. [16] proposed a fully synthetic but realistic synthesis and Yang [17] presented a framework for traffic scene auralization. The propagation model was implemented using the Unity3D game engine based on the image-source model with multi-tap time-varying delay lines.

The approaches proposed in this paper have notable differences from previous techniques. The first one is the use of indoor laboratory-controlled near-field sources contribution from individual components. The measurements are readily available from the well-established indoor pass-by noise test [3]. The second difference is the use of infinite impulse responses and spherical harmonics to address the issue of moving sources without the generation of clicks and artifacts. Both methods have been initially presented in past conference publications [18, 19] and are here detailed and compared. The main motivation of developing such techniques is to accurately synthesize moving events for subjective audio assessment. Since the microphone array is statically positioned, additional processing is required to obtain the indoor PBN time signal that is comparable to the time signal obtained in the exterior PBN test. The challenge therein is the correct mixing of the far-field individual signals to recreate the continuous moving effect.

The paper is organized as follows: the measurement procedure for the indoor PBN is presented in Section 2. Two real-time sound synthesis techniques are presented in Section 3, namely the time-varying IIR filter (TVIIR) and the spherical harmonics (SH) representation. To ensure the accuracy and naturalness of the output audio synthesis, the two approaches are compared quantitatively and qualitatively in Section 4 using an electrical vehicle measured in a hemi-anechoic chamber.

## 2 Source quantification and transfer path analysis

This section deals with the measurement procedure to quantify the many sources present in the system and to measure the transfer path in the form of noise transfer functions (NTF). The noise transfer functions are related to a broadband noise source measured with microphones with units  $\text{Pa}/(\text{m}^3 \text{s}^{-1})$ . The goal of the indoor PBN is to accurately reproduce an exterior PBN test from far-field recorded time signals. A schematic of the setup is shown in Figure 1.

The first step of the indoor PBN test is to identify and separate the noise components using the acoustic source quantification (ASQ) technique [9]. The sources are assumed to be airborne, radiating outwards from the powertrain, gearbox, exhaust, tailpipe, front and rear tires, or any other component. By assuming that each noise-producing component can be represented as a superimposed set of monopole sources, the operational acoustic loads can be identified from independent component measurements by an inverse procedure such as the ASQ. Note that the car body effect (reflection, diffraction) are implicitly included in the model since they are present in the measured data.

Two types of ASQ are commonly employed, namely linear phase-based pressure inversion method and power-based energetic approach [20, 21]. While the linear approach accounts for the phase information of the source, the energetic approach treats all components as uncorrelated sources. The latter allows to write the power at the receivers in terms of that of the sources, as

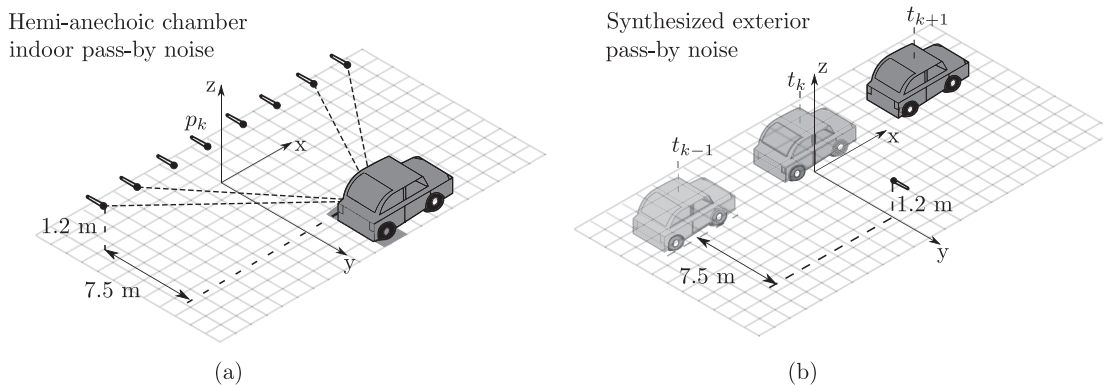
$$|p_i|^2 = |G_{im}|^2 |q_m|^2, \quad (1)$$

where  $q_m(\omega)$  ( $m = 1, \dots, N_q$ ) represents the volume velocity of the sources (in  $\text{m}^3 \text{s}^{-1}$ ),  $p_i(\omega)$  ( $i = 1, \dots, N_{\text{ind}}$ ) is the sound pressure at the near-field indicator microphones (in Pa) and  $G_{im}(\omega)$  is the transfer function between source point  $m$  and observation point  $i$ . These are obtained by measuring the relation between the indicator microphones and a known volume velocity omnidirectional source emitting a broadband noise at the location of the equivalent monopole sources. Figure 2 shows an example for a single source and four indicator microphones. Since  $N_{\text{ind}} \geq N_q$ , equation (1) is solved by inverting  $|G_{mi}|^2$  in a least-squares sense with real positive constraint on the solution [21] for any given angular frequency  $\omega$ . Note that for the problem of multiple sources, the matrix  $G_{mi}$  is a full matrix with cross-terms between all paths and indicators.

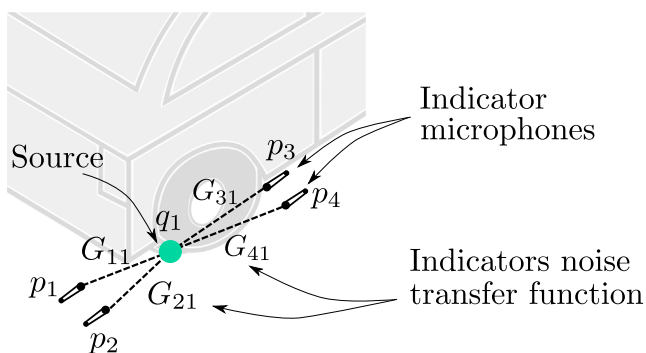
Once the sources are quantified, the propagated sound pressure at the far-field microphones (also referred to as target microphones)  $p_k(\omega)$  ( $k = 1, \dots, N_p$ ) can be obtained energetically as

$$|p_k|^2 = |G_{km}|^2 |q_m|^2, \quad (2)$$

where  $G_{km} \in \mathbb{C}^{N_p \times N_q}$  is the measured transfer function between the  $m$ th source and the  $k$ th far-field microphone. These transfer functions are obtained following a similar procedure as the locally measured transfer functions using the broadband noise volume velocity and far-field microphones, and intrinsically include interactions of the radiated field with the vehicle such as reflections or diffraction. Thus, the set of transfer functions between each source and target microphones represents changes in propagation distance, directivity, and angle of incidence as the vehicle passes by. The far-field microphone array configuration is defined by the desired trajectory of the moving source and can be with any arbitrary shape. Traditionally, in indoor PBN applications, the far-field microphones are position in a line configuration as shown



**Figure 1.** Schematic of indoor pass-by noise sound synthesis where (a) is the measurement setup in a hemi-anechoic chamber and (b) is the synthesized exterior PBN.  $k$  spans from 1 to  $N_p$  far-field microphones.



**Figure 2.** Acoustic source quantification for a source  $q_1$  using four indicator microphones.

in Figure 1. Note that the obtained far-field microphone signals are the sum of the contribution of the  $M$  sources. Alternatively, the pressure field can be obtained for each  $m$ th source and  $k$ th far-field microphone separately using  $|p_{km}|^2 = |G_{km}|^2 |q_m|^2$ , with no implied summation in  $m$ .

Note that the NTF could be replaced by wave-base simulations (finite and boundary element methods) which can include locally and non-locally reactive ground model.

### 3 Moving source sound synthesis

This section describes the two auralization techniques for synthesizing a moving source from the knowledge of the stationary transfer functions obtained in the previous section. The general scene consists of a moving source and a fixed listener able to rotate their head. Figure 3 shows a schematic of the two proposed approaches.

Both approaches work in an online-offline manner. The offline part consists of acquiring the near-field source signals and the stationary noise transfer functions. In the time-varying IIR technique, shown in Figure 3a, the offline processing extends to the computation of the IIR filter parameters (i.e. IIR design). Once the filters are designed, they are recursively applied in the time-domain (i.e. time-varying implementation). This procedure is repeated to implement the head related transfer functions (HRTF)

$H_{km}(\omega)$  corresponding to each far-field microphone position. The IIR representation of the HRTFs has been previously researched [22], and it is an acceptable representation for auralization purposes. Nonetheless, the use of finite impulse response (FIR) is also allowed in the framework with an increase in computation cost in the implementation. Finally, the binaural output signal is obtained from the monaural synthesized signal.

The second technique, shown in Figure 3b involves the representation of the propagated source time signal into spherical harmonics. The entire processing chain is performed in real time and consists of a multi-channel granular synthesis followed by a spherical harmonic encoding, propagation, and a binaural decoding procedure. This method easily allows for interaction with a video rendering tool which communicates with the audio engine to provide input parameters such as the source position, throttle speed and head rotation. The details of each approach are given next.

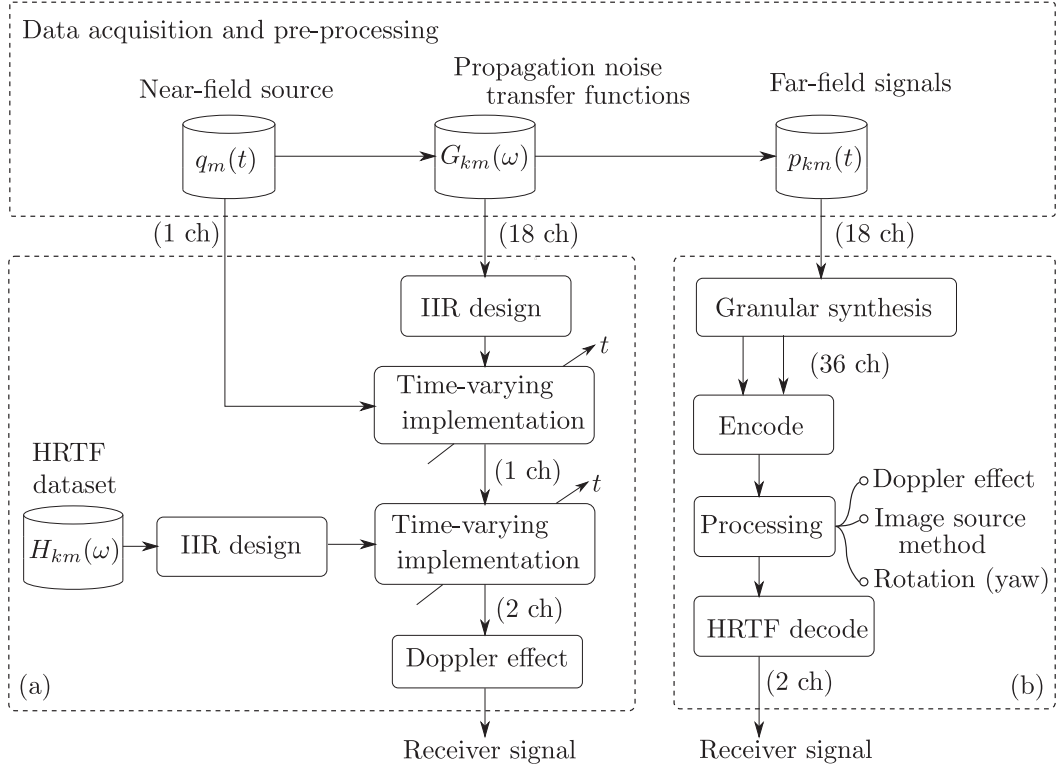
#### 3.1 Time-varying IIR filters (TVIIR) approach

The Time-varying IIR filters (TVIIR) method consists of the following steps: decompose the noise transfer functions (NTFs) into minimum and excess phase, approximate the magnitude of the NTFs into IIR filters, interpolate the filter coefficients according to the far-field microphone positions and implement the filters in time-domain for each source.

To make the notation clearer, the formulation is restricted to a single source (i.e.  $m = 1$ ), and the subscript is dropped as  $G_{km} \equiv G_k$ . The decomposition of the NTFs is given by [22]

$$G_k(\omega) = |G^{(\min)}(\omega)| |G^{(\text{eph})}(\omega)| e^{j\phi^{(\min)}} e^{j\phi^{(\text{eph})}}, \quad (3)$$

where  $\phi^{(\min)}$  denotes the minimum phase,  $\phi^{(\text{eph})}$  denotes the excess phase,  $|G^{(\min)}(\omega)|$  is the minimum phase magnitude and  $|G^{(\text{eph})}(\omega)| = 1$  is the all-pass magnitude. Note that both phases (i.e. minimum-phase and excess phase) are frequency-dependent quantities. While the minimum phase is neglected, the excess phase models the time of arrival and is used in the Doppler effect implemented later in this section.



**Figure 3.** Overview of the two proposed sound synthesis frameworks for a single near-field source; (a) time-varying IIR filter and (b) spherical harmonics.

The magnitude of the minimum phase system is employed in the IIR filter design with two additional pre-processing steps, namely smoothing and warping. The frequency-dependent smoothing consists of a convolution of the NTF with an averaging Hann window whose length is defined as  $\mathcal{Q}$ , which represents the ratio of bandwidth over the center frequency. The warping has the effect of resampling the NTF on a warped frequency scale by defining the following bilinear conformal map for the unit delay  $z^{-1}$  in the  $Z$ -domain [23, 24]

$$z^{-1} \rightarrow \bar{z}^{-1} = \frac{z^{-1} - \varrho}{1 - \varrho z^{-1}}, \quad (4)$$

where  $\varrho$  is the warping coefficient. The warping function has the effect of oversampling the magnitude spectrum at low frequencies and undersampling at high frequencies, thereby preserving the perceived spectral features of the sound. This is done by choosing a value of  $\varrho$  that guarantees a constant density of spectral lines across the frequency bands relevant for hearing (e.g. Bark scale and equally rectangular bands).

According to equation (3), since the all-pass magnitude is unitary, the IIR filters can be designed considering only the minimum phase magnitude. Here, IIR filters are represented as

$$G_k(z) = \frac{b_{0k} + b_{1k}z^{-1} + \dots + b_{N_b k}z^{-N_b}}{a_{0k} + a_{1k}z^{-1} + \dots + a_{N_a k}z^{-N_a}}, \quad (5)$$

where the  $b_{ik}$ ,  $i = 1, 2, \dots, N_b$  and  $a_{ik}$ ,  $i = 1, 2, \dots, N_a$  are the filter coefficients for the  $k$ th far-field microphone,  $N_b$  is the feed-forward filter order and  $N_a$  is the feedback filter order. The coefficients of the IIR filter are estimated using the modified Yule–Walker method – an autoregressive moving-average (ARMA) technique for high-resolution spectral estimation of linear time-invariant systems [25].

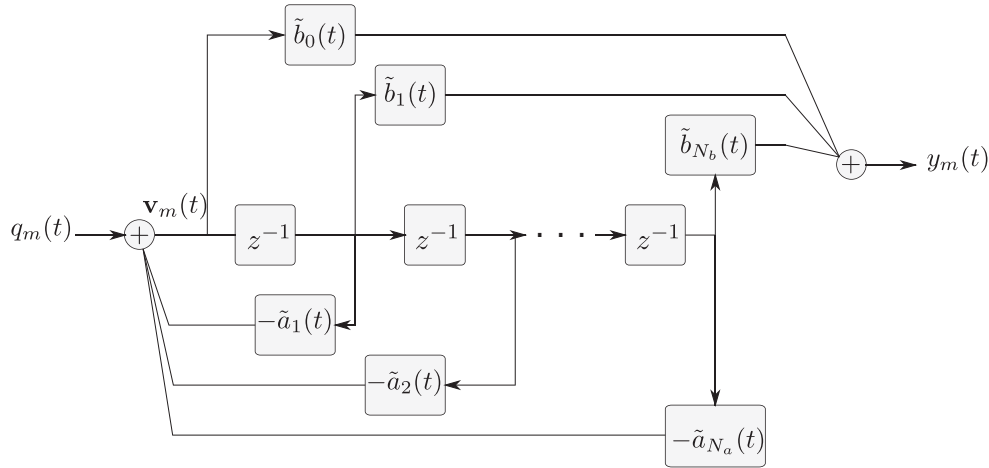
Once the IIR filter is designed for each far-field microphone, a database of  $b_{ik}$  and  $a_{ik}$  coefficients can be constructed as follows

$$b_{ik} = \begin{pmatrix} b_{i1} & \dots & b_{ik} & \dots & b_{iN_p} \\ \vdots & & \vdots & & \vdots \\ b_{N_b 1} & \dots & b_{N_b k} & \dots & b_{N_b N_p} \end{pmatrix}. \quad (6)$$

To update the IIR coefficients at the source audio sample rate, an interpolation strategy is employed between two adjacent far-field microphones. The simplest solution is the linear interpolation given by

$$\tilde{b}_i(t) = \frac{x(t) - x_k}{x_{k+1} - x_k} (b_{i,k+1} - b_{ik}) + b_{ik}, \quad (7)$$

where  $\tilde{b}_i(t)$  and  $x(t)$  are the value of the coefficient and position at the instant corresponding to the  $t$ th sample,  $x_k$  and  $x_{k+1}$  are the two closest far-field microphone positions to the desired position  $x(t)$  within the target trajectory. The interpolation of equation (7) is similar for the  $\tilde{a}_i(t)$  coefficients which is omitted for brevity.



**Figure 4.** Direct-form II with time-varying coefficients block diagram.

The IIR filter with time-varying coefficients can be implemented in the time-domain by solving the direct-form II where the output  $y(t)$  at time  $t$  is computed from the present input  $q(t)$  and the past output samples as described by the following state-variable expression [26]

$$\begin{cases} \mathbf{v}_m(t+1) = \mathbf{F}(t)\mathbf{v}_m(t) + \mathbf{w}q_m(t) \\ y_m(t) = \mathbf{g}^T \mathbf{v}_m(t) + b_0 q_m(t) \end{cases},$$

$$\mathbf{F}(t) = \begin{pmatrix} a_1(t) & a_2(t) & \cdots & a_{N_a-1}(t) & a_{N_a}(t) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \quad (8)$$

where  $\mathbf{v}_m$  represent the state variables of the filter initialized by  $\mathbf{v}_m(0) = 0$ , vector  $w = [1 \ 0 \ \cdots \ 0]^T$ , and vector

$$\mathbf{g} = \begin{Bmatrix} b_1(t) - b_0(t)a_1(t) \\ b_2(t) - b_0(t)a_2(t) \\ \vdots \\ b_{N_b}(t) - b_0(t)a_{N_b}(t) \end{Bmatrix} \quad (9)$$

contains the feed-forward coefficients. In the direct-form II, the delay line is shared between the all-pole and all-zero sections as shown in Figure 4, halving the number of delays compared to the direct-form I.

The approach is valid for  $M$  sources by invoking the superposition principle and assuming that the sources are uncorrelated. As indicated in Figure 3, similarly as the IIR filters derived from NTFs, other filters can be implemented. For instance, a natural improvement to the sound synthesis is to add both the ear canal filtering and the head and torso diffraction through the use of HRTFs which transform a monaural signal into a binaural one.

So far, the sound synthesis considers only the attenuation from the NTF magnitude. Recasting equation (3), the additional piece of information not included in the solution is the propagation time delay from the phase

information resulting in a frequency shift that simulates the Doppler effect. Knowing that the minimum phase and magnitude are uniquely related by the inverse Hilbert transform  $\phi^{(\min)} = \mathbb{H}^{-1}\{\ln(-|G^{(\min)}(\omega)|)\}$ , the time delay  $\tau$  can be inferred from the slope of the unwrapped all-pass excess phase  $\phi^{(\text{epb})}$ . The combination of all time delays between each source and far-field position yields the time delay function.

The time delay is a function of the source position. Indeed, different Doppler effects can be achieved by different source speeds in the direction of the receiver. In the PBN case, the source position and speed are obtained from the tachometer which tracks the revolution per minute (RPM) during measurements. The procedure has to be repeated for each source as the propagation distance changes depending on the source location. However, if the distances between sources are small compared to the propagation distance, an averaged time delay function can be used.

Finally, since the length of the averaged time delay is not an integer multiple of the time increment, interpolation is required of for instance a bandlimited interpolation of the form [27]

$$\tilde{y}_m(\tau) = \sum_{n=-\infty}^{\infty} y_m(t_n) \frac{\sin[\pi f_s(\tau - t_n)]}{\pi f_s(\tau - t_n)}, \quad (10)$$

where  $f_s$  is the sampling rate,  $t_n$  is the discrete version of  $t$  and  $\tilde{y}_m(\tau)$  is the delayed version of the signal  $y_m(t)$ . The Doppler effect applies as a time-varying time delay onto equation (10) in the form  $\tau = \tau(t)$ , following the rate of motion from the source's motion.

### 3.2 Spherical harmonics (SH) approach

The second approach is based on the representation of a sound field into spherical harmonics (SH). The main motivation for utilizing such approach is the ability to synthesize sound propagation past the measurement position, thus eliminating the room size constraint. The SH representation also enables treating boundary conditions using the image source method, which is a valid assumption when

dealing with compact sources and the far-field. Additionally, the spherical harmonics enables a dynamic rotation of the sound field which, combined with head-related transfer functions, allows the listener to perform head adjustments, which is known to improve sound localization [28].

Since the method relies on the spherical harmonics representation it is convenient to arrange the far-field microphones in a circular or spherical array configuration around the object. According to Figure 3, the process is given by the encoder, processing, and decoder. Several libraries have been created to encode, decode and manipulate sound [29, 30]. The procedure here presented follows the SH channel numbering notation [31].

From the definition of the discrete spherical harmonic transform [32] and by considering the sound field as a superposition of plane waves, the discrete SH coefficients can be computed as [33]

$$\phi = \mathbf{Y}\mathbf{p}, \quad (11)$$

where  $\mathbf{p} \equiv p_k$ ,  $\mathbf{Y} = [Y_0, 0(\boldsymbol{\Omega}) \ Y_{1,-1}(\boldsymbol{\Omega}) \ Y_{1,0}(\boldsymbol{\Omega}) \ \dots \ Y_{N,M}(\boldsymbol{\Omega})]^T$  is the matrix of spherical harmonics coefficients,  $N$  and  $M$  are the spherical harmonics' order and degree, respectively,  $\boldsymbol{\Omega} = [\Omega_1 \ \Omega_2 \ \dots \ \Omega_{N_p}]$  are the directivity of the far-field microphones, with  $\boldsymbol{\Omega} = (\theta, \varphi)$  as the short notation for the elevation,  $\theta \in [-\pi/2, \pi/2]$  and azimuth,  $\varphi \in [-\pi, \pi]$ , respectively, and  $\phi = [\phi_{0,0}(\omega) \ \phi_{1,-1}(\omega) \ \dots \ \phi_{N,M}(\omega)]^T$  are the SH coefficients. The encoding procedure consists in estimating  $\phi(\omega)$  by solving equation (11) in a least-squares sense. For the solution to be unambiguous the inverse problem requires that  $L \geq (N+1)^2$  for a 3D case and  $L \geq 2N+1$  for 2D case [32]. Note that the direction of the waves is not accounted for in equation (11) since the radial function is neglected. Nevertheless, the sound field is assumed to be outgoing in the case of a surround microphone array or ingoing in the case of an ambisonics microphone array. The ambisonics microphone array is a rigid compact sphere which can capture the incident sound field. Alternatively, the full radiated sound field (i.e. the directivity pattern) can be captured using a surround microphone array and, hence, the configuration employed in this work.

Both array configurations suffer from spatial aliasing due to the limited spatial sampling on the sphere. To minimize spatial aliasing errors, both the SH expansion order and the microphone sampling distribution need to be carefully considered. It is known that such a representation can be applied to sound fields containing frequencies up to  $f_u < Nc/(2\pi R)$  [32]. Note that the radius of the microphone array influence the accurate frequency range. For instance, surround arrays have a large radius and the frequency range where the spatial aliasing does not occur is limited to low frequencies. The main consequence of neglecting the radial function is the lack of a correction to the SH coefficients that accounts for the microphone array construction (e.g. rigid, open) and radius. Since the array is open, and has a large radius, the effect of the radial function is small and is here neglected. However, an in-depth investigation of this assumption requires further investigation.

Finally, the decoder operation consists in estimating the response at the listener's ears and is given by [34]

$$\hat{y}_m = \mathbf{w}^H \phi, \quad (12)$$

where  $(\cdot)^H$  is the Hermitian conjugate and  $\mathbf{w}$  is a rendering filter.

For completeness, a least-squares minimization technique is performed to find the rendering filter such that it perceptually approximates the solution to the target signal,  $y_m(\boldsymbol{\Omega}) = p(\boldsymbol{\Omega})H(\boldsymbol{\Omega})$  as

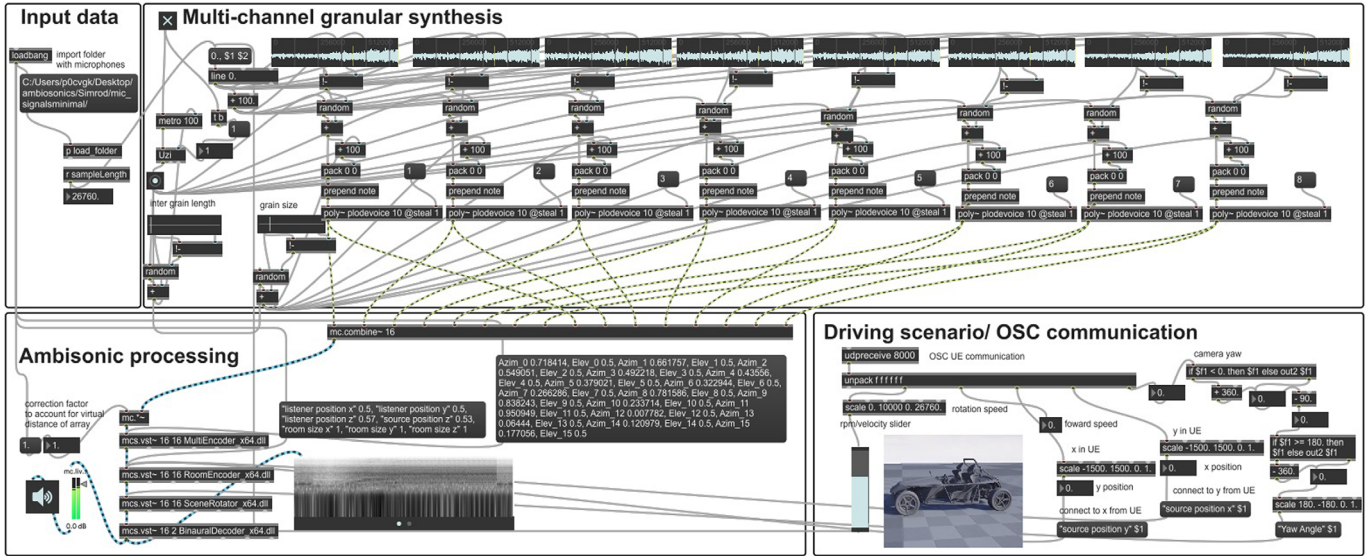
$$\min_{\mathbf{w} \in \mathcal{K}} \sum_{\boldsymbol{\Omega} \in \mathcal{M}} |\mathbf{w}^H \mathbf{Y}(\boldsymbol{\Omega}) - H(\boldsymbol{\Omega})|^2, \quad (13)$$

where  $\mathcal{K}$  is the domain in which  $\mathbf{w}$  is optimized,  $\mathcal{M}$  is the dense set of directions, and  $H(\boldsymbol{\Omega})$  denotes a HRTF at any discrete direction  $\boldsymbol{\Omega}$ . The rendering filter solution reduces to an algebraic expression in the least square sense [34].

Note that the encoding procedure is frequency-independent [35] and the decoding procedure is frequency-dependent. Hence, the spherical harmonics signal needs to be transformed from time domain to time-frequency domain using a fast Fourier transform (FFT). The resulting signal in the time domain is obtained by performing an inverse FFT procedure. The operations are performed in real-time using a block-based processing scheme.

To move the source, sound scene manipulation through transformations of the spherical harmonics coefficients can be performed. The transformations can be frequency-independent such as the rotation of the scene, mirroring across planes, warping, compression, decompression, and amplitude manipulation [35], or frequency-dependent effects such as geometrical distancing effect, reverberation, and diffuseness [36].

The present work uses Virtual Studio Technology (VST) plugins [37] hosted in Cycling '74 procedural language MAX (version 8.1.6), as shown in Figure 5. The multi-channel granular synthesis input is the measured pressure field data with its related tachometer trace (rpm vs. time), allowing virtual accelerations along a given path. The encoding operation is done using the *MultiEncoder* (version 0.6.1), the binaural decoding operation is done using the *BinauralDecoder* (version 0.6) and the translation of the moving source is done using the *RoomEncoder* (version 1.3.1) from the IEM plugin suite [38]. The encoding operation introduced an order-dependent gain in the ambisonic signal. A multi-channel all-pass attenuation filtering of  $-19.4$  dB is applied to the input signal to equalize the SH coefficients, preserving the loudness levels and compensating the gain introduced in the encoding operation. The graphical rendering of the vehicle is implemented using the Unreal Engine [39] to provide visual feedback to the audio scene, as shown in Figure 5. The communication between MAX and Unreal Engine is done through the open sound control (OSC) protocol. For the comparison between the two approaches discussed next, the graphical rendering and multi-channel granular synthesis are not required therefore, they are omitted.



**Figure 5.** Max/MSP patch with multi-channel granular synthesis, spherical harmonics processing, OSC protocol communication with the vehicle visual rendering in Unreal Engine.

In summary, the main contributions of this section are the development of a demonstrator tool with the developed multichannel granular synthesizer, the combination of the measurements with above-mentioned plugins for a novel application (i.e. outdoor pass-by noise sound synthesis) in addition to the OSC implementation with the visual rendering tool.

## 4 Results: indoor pass-by noise

In this section, results from the two sound syntheses are compared using an indoor pass-by noise measurement data set. The assessment is performed quantitatively, and also subjectively by means of a listening test. The synthesized audio scene consists of a vehicle accelerating along a straight path and with a listener positioned 4.7 m away similar to the PBN setup shown in Figure 1b. For the quantitative results, the source-receiver distance in the equivalent outdoor PBN scene is set to reach its minimum value when the vehicle reaches 50 km/h. For the subjective evaluation, two constant moving source speeds are considered.

### 4.1 Measurement setup

The purpose of the measurement is to obtain the source loads' time-signals and the far-field propagation noise transfer functions. An electric vehicle is placed on a chassis dynamometer in a hemi-anechoic chamber as shown in Figure 6a. The transfer functions and source quantification are obtained following the methodology presented in Section 2.

Besides the total contribution, the components considered are: the gearbox, rear left tire (*tireRL*) and rear right tires (*tireRR*). Each tire is represented by two equivalent monopolar sources, instrumented with four microphones each, and the gearbox is represented by a single

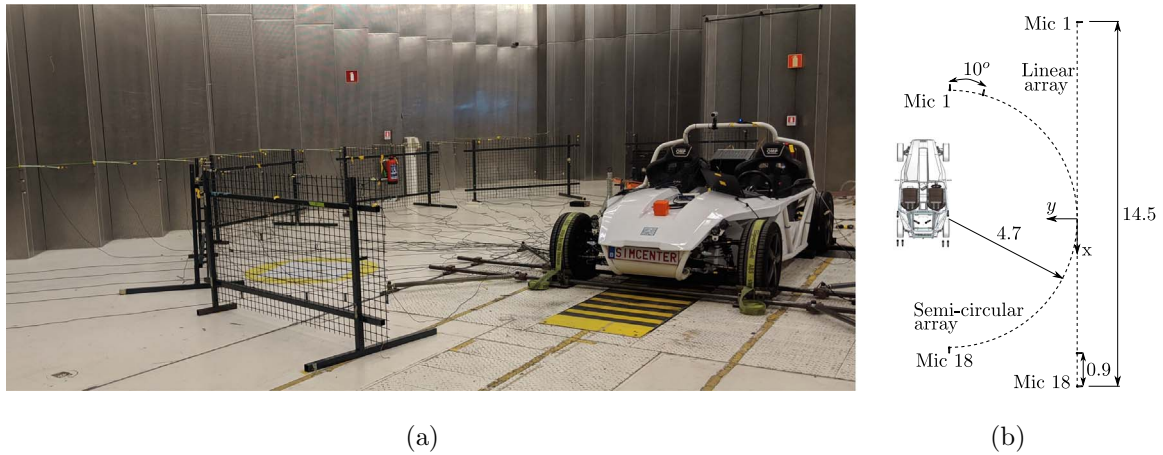
monopolar source, instrumented with two microphones. The separation of the source is an optional step, and it is here done to highlight the acoustic source quantification technique. The advantages of separating the sources are the ability to assess each component individually for troubleshooting and to combine different components in post-processing for the evaluation of not available physical configurations.

In addition to these 10 near-field microphones, two far-field microphone arrays are installed, with linear and semi-circular shape respectively and with a total of 18 microphones each, as shown in Figure 6b. All 46 channels are acquired simultaneously. In the case of the linear array, microphone 1 is positioned at  $(-8.2, 0)$  m and microphone 18 is at  $(6.2, 0)$  m, giving a total array size of 14.4 m. All microphones used are  $1/4''$  externally polarised integrated circuit piezoelectric microphones (GRAS 40-PH) placed at a height of  $1 \pm 0.05$  m. Note that the distance of the far-field microphone does not match the required distance in the ISO 362-3 [3] due to the room size limitations.

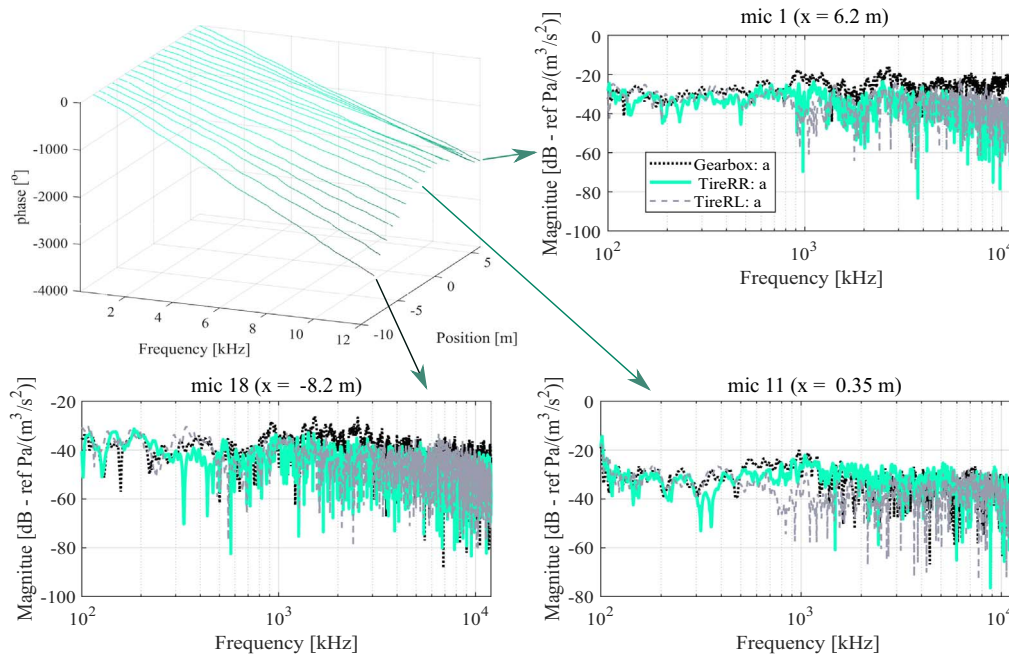
### 4.2 Sound synthesis from time-varying IIR filters

Figure 7 shows the NTFs obtained between the equivalent source and the far-field linear microphone array using an omnidirectional sound source located at the gearbox, left and right tires. The NTFs have a sampling rate of 25.6 kHz. The magnitude of the NTFs display a comb filter-like behavior, induced by interference between the direct acoustic paths and the ground reflections. In upper left corner of Figure 7, the variation of time delay can be observed on unwrapped phase slopes at each microphone position.

Figure 8 shows a comparison of the magnitude of the measured NTF against the IIR filter for different sets of design parameters. The NTF is arbitrarily chosen, and it corresponds to a source signal arriving from the gearbox



**Figure 6.** (a) Photograph of the measurement setup with near-field (indicators) microphones, a far-field semi-circular array, and a far-field linear array; (b) schematic with the setup main dimensions in m.



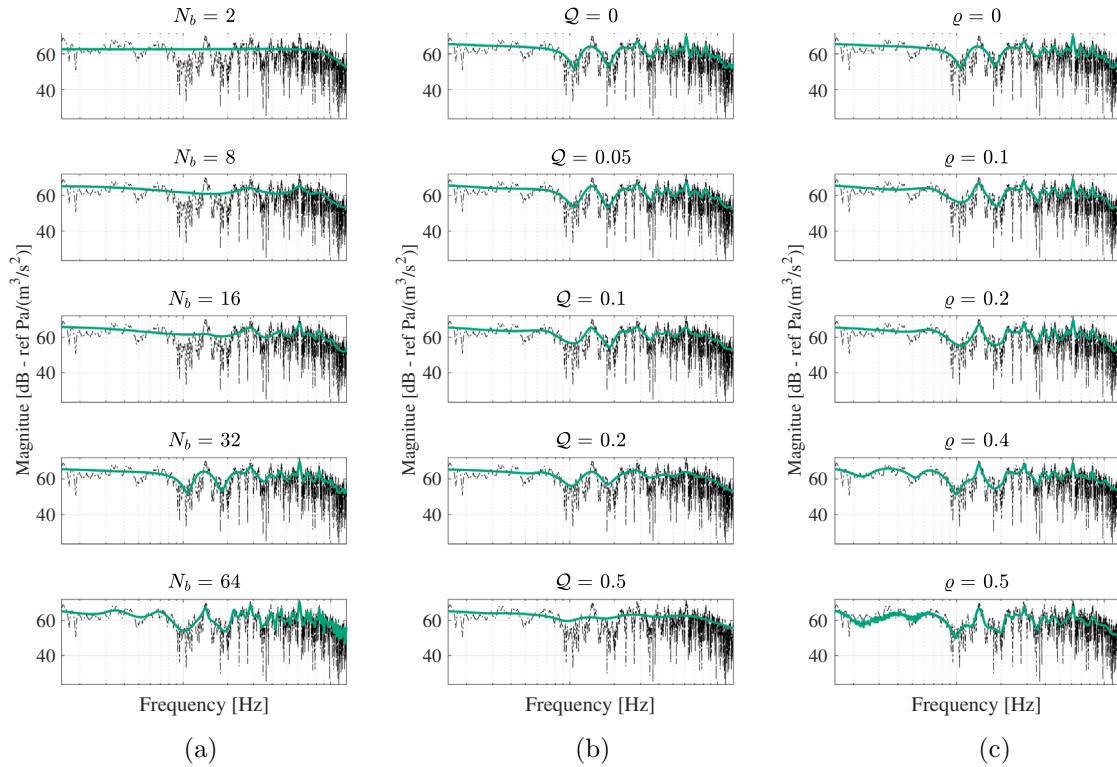
**Figure 7.** Magnitude and phase of the noise transfer functions for three discretized position for the gearbox, and left and right tires; the gearbox phase for each position is shown as surface plot in the upper left corner.

location and receiver located at the far-field microphone 11 (refer to Figure 6b). It can be seen in Figure 8a that the filter design quality is mostly induced by the filter order. The higher the filter order, the finer the details captured by the filter, especially at higher frequencies. The main drawbacks of using IIR filters are that instabilities appear with increasing filter order and that these filters are prone to errors during the time-varying implementation when using direct-form techniques. Both smoothing,  $Q$ , and warping,  $q$ , show smaller effects on the outcome than the feed-forward filter order  $N_b$  and can be used as fine-tuning parameters. An increase in the warping coefficient improves slightly the lower frequencies fit of the filter. The smoothing is

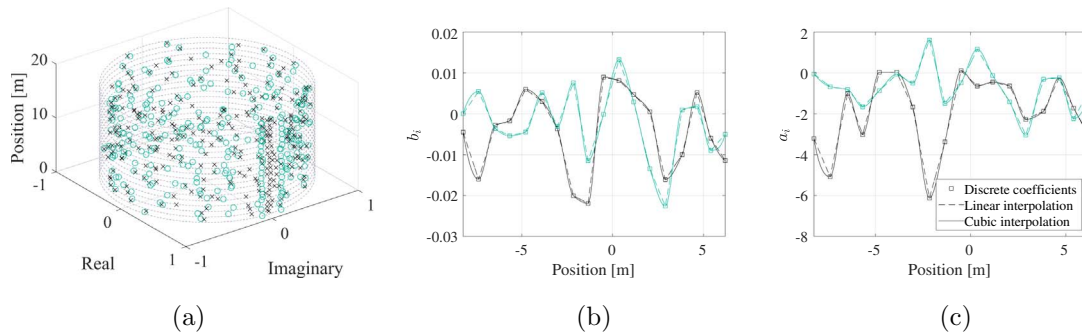
required when applying the warping because the spectrum is undersampled at higher frequencies [22]. For the remaining analysis of this subsection, the NTF's IIR filters are designed using the following parameters:  $Q = 0.02$  and  $q = 0.1$ .

The IIR filter coefficients  $A$  and  $B$  are computed for each microphone position and interpolated in space along the position axis ( $x$ -axis). In this example, the stability of each filter is guaranteed since all the poles are located within the  $Z$ -domain unit circle as seen in Figure 9a. Figures 9b and 9c show two arbitrarily chosen IIR filter coefficients (8th and 16th order) across the microphone positions as well as two interpolation strategies. It can be





**Figure 8.** Propagation noise transfer functions (black dashed line) and IIR filter (green solid line) varying (a) the filter order  $N_b$  from 2 to 64 with  $Q = 0$  and  $\rho = 0$ , (b) smoothing  $Q$  from 0 to 0.5 with  $N_b = 32$  and  $\rho = 0$  and (c) warping  $\rho$  from 0 to 0.5 with  $N_b = 32$  and  $Q = 0$  for microphone 11.

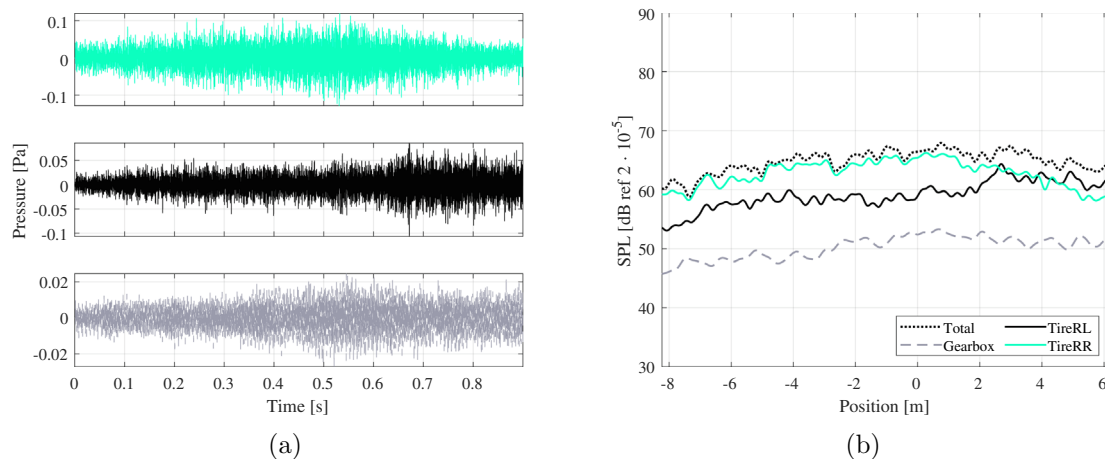


**Figure 9.** Filter properties for the representation of the transfer function between the gearbox and the far-field linear array. (a) Poles (green  $\circ$ ) and zeros (black  $\times$ ) in the  $Z$ -domain, (b)  $b$  and (c)  $a$  coefficients of the IIR filter along the position axis with the design parameters  $N_b = 32$ ,  $Q = 0.01$  and  $\rho = 0.05$  for two arbitrarily chosen orders (8th and 16th) and the linear and cubic interpolation schemes.

noticed that the cubic interpolation has a slightly smoother transition between coefficients. However, the implementation using the cubic interpolation suffers from high sensitivity to coefficient variation due to the recursive nature of the IIR filters. This occurs in both direct-form I and II implementations, which could be attributed to disturbances in the future values of the internal state variables and transients in the output [26]. To solve this issue, the order of the filter can be reduced or one can use linear

interpolation instead, which does not demonstrate the same high sensitivity behavior. Alternatively, higher-order IIR filters could be converted into a cascade of bi-quadratic (2nd order) stable IIR filters in an attempt to increase numerical robustness. In this example, this alternative is not further explored and a linear interpolation is employed.

Figure 10 shows the resulting time signals for each component and the sound pressure level with the total noise level. The implementation is performed using the direct-



**Figure 10.** (a) Monaural pressure time signal without the Doppler effect for the (top) rear right tire (tireRR), (center) the rear left tire (tireRL) and for the (bottom) gearbox; (b) sound pressure level for each component and the total SPL contribution.

form II scheme (refer to Eq. (8) and Fig. 4). The signals are presented without the Doppler effect and the HRTFs. The maximum sound pressure level obtained is below 70 dB and the rear right tire noise shows the highest sound pressure level, as expected for an electric vehicle at 50 km/h. In the far-end position ( $x = 5$  m), the rear tire noise level becomes closer with the left side and slightly surpassing the right side.

### 4.3 Sound synthesis from spherical harmonics

In the SH implementation, the propagated time signals at the far-field microphones are first obtained through convolution between the ASQ-estimated sources with the measured NTFs on the semi-circular array (see Fig. 3). For this particular case, before the encoding operation, the propagated time-signals are mirrored to form a full 2D circular domain with a total of 36 input channels which allows for a reconstruction up to 17th order. This relies on the implied assumption that the vehicle and its radiated sound are symmetric as observed from the far-field array. Additionally, as the measurements already include propagation attenuation, none is added to the encoded signal. Furthermore, the aliasing – free region is bounded by an upper frequency limit. Such frequency limit is nevertheless difficult to accurately estimate with the current setup due to the unknown directivity and spatially-extended nature of the source of interest [40]. Also note that the proposed setup only allows for reliable synthesis in the horizontal plane containing the microphones, which is suitable for the PBN application.

The spherical harmonics coefficients from the encoded source are translated by relying on the real-time image source propagation plugin *RoomEncoder*. The moving path follows a linear trajectory going from  $(-8.2, 0)$  m to  $(6.2, 0)$  at a height of 0.5 m, similar to the measurement setup. The receiver is positioned at  $(0, 4.7)$  m away from the source trajectory and at a height of 1.0 m.

### 4.4 Comparison of sound synthesis outputs for a pre-defined scene

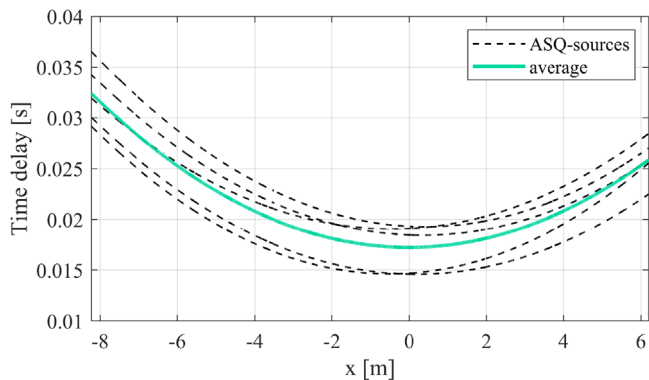
This subsection compares the two sound synthesis approaches for a pre-defined scene. The Doppler effect and the HRTFs are added to the TVIIR output where the required time delay is obtained following the procedure of Section 3.1. Figure 11 shows the time delay curves for each ASQ source and an averaged time delay. Since the vehicle trajectory is parallel to the line array, the Doppler effect is implemented by fitting the measured time delay to a quadratic polynomial function.

The HRTFs are the same as the one implemented in the SH approach [41], which are designed as 8th order IIR filters. Note that the final time delay does not account for the one introduced by the HRTFs processing which is here assumed negligible for the considered propagation distance. The 32nd order TVIIR is compared to a 5th order spherical harmonics order.

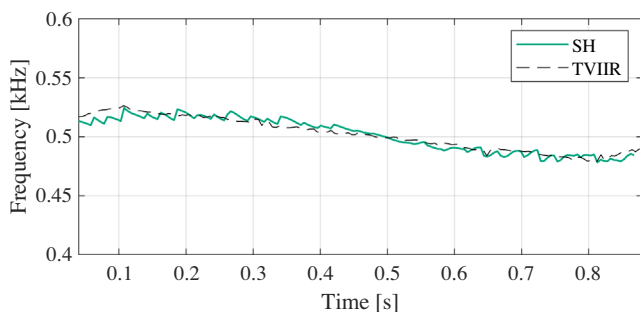
To check for undesirable artifacts, the two approaches are first compared using a simple harmonic source, emitting a sine wave at 500 Hz as shown in Figure 12. In this example, the harmonic source signal is replicated in a circular configuration be equivalent to the original problem.

Figure 12 shows that both methods yield a comparable Doppler shift. The shift is almost linear due to the short time segment imposed by the measurement room and the proximity between source and receiver where the vehicle appears to be moving in constant velocity. Slightly differences are observed. The instantaneous frequency shows a level of variability in the SH approach, which can be attributed to latency in the SH processing chain and due to buffer size in the block processing operation imposed by the sound card. However, it is worth noting that the audio synthesis does not suffer from clicks, artifacts or audible degradation. Audio samples are provided as supplementary files [42].

Figure 13a shows the receiver binaural time signals and Figure 13b shows the sound pressure levels synthesized with



**Figure 11.** Time delay obtained from excess phase  $\phi^{(\text{eph})}$  for each ASQ-source and averaged time delay.



**Figure 12.** Instantaneous frequency at receiver location from a single source emitting a 500 Hz sine wave and derived from the time-varying IIR (TVIIR) and spherical harmonics (SH) techniques.

both implementations from static ASQ-estimated source signals.

It can be observed that both approaches are in good agreement in terms of sound pressure level. In this case, the TVIIR is considered as the true reference since the method has been validated against an energetic pass-by noise sound synthesis approach [18]. The average sound level difference across the 14 m trajectory is of 1.4 dB and 0.9 dB for the left and right ear signals, respectively. Note that the initial delay observed in Figure 13a arises from the initial propagation delay. A small discrepancy is seen around the  $-2$  m and 4 m positions in Figure 13b which is not perceptible in the audio. The synthesis of click-free moving sources using the two techniques is the main outcome of this paper and a comparison against a real case is left out for future investigation.

Figure 14 shows the spectrograms of the right ear signals synthesized by both approaches. Contrary to the sound pressure level, the spectrograms are noticeably different. While the Doppler shift is similar, the spectrum content of both approaches diverges. In the TVIIR approach, the frequency content is concentrated at low frequencies and in the SH approach, there is a larger distribution across the frequency range, especially at higher frequencies. This

difference can be attributed to the simplifications inherent to the two approaches. Moreover, improvements in the matching of the two approaches can be achieved at the expense of higher IIR filter order and higher SH order.

#### 4.5 Online listening test

To evaluate the proposed sound synthesis approaches, a subjective evaluation is performed by means of an online listening test [43]. The objective is to evaluate the realism of the moving source and to check for any noticeable perceptual differences between the two approaches. The listening test was performed with a total of 20 participants. The majority of them had an engineering background and previous experience with listening tests.

A reference signal, consisting of a static microphone at a central position in operational condition was played before the jury test. Note that this reference sound is provided only as an example for the vehicle at a static position. The moving source aspect is inferred by what the listener understands as a moving source, i.e. his/her prior knowledge. Additionally, there was no control of the sound level and type of headphones the participants use. However, the participants were asked to not adjust their headphones during the test.

Twelve synthesized signals are investigated among the two approaches, the different orders, and the cinematic conditions. The orders for each approach are selected to reflect different levels of accuracy in the spherical harmonic and the filter design. For the SH approach, the sound synthesis is performed using a 1st, 3rd and 5th-order spherical harmonics. For the TVIIR approach, the sound synthesis is performed using 8th, 16th and 32nd order filters. Additionally, the sounds are presented in both 25 km/h and 50 km/h speeds. The speeds are selected to reflect the usual speeds found in an urban environment.

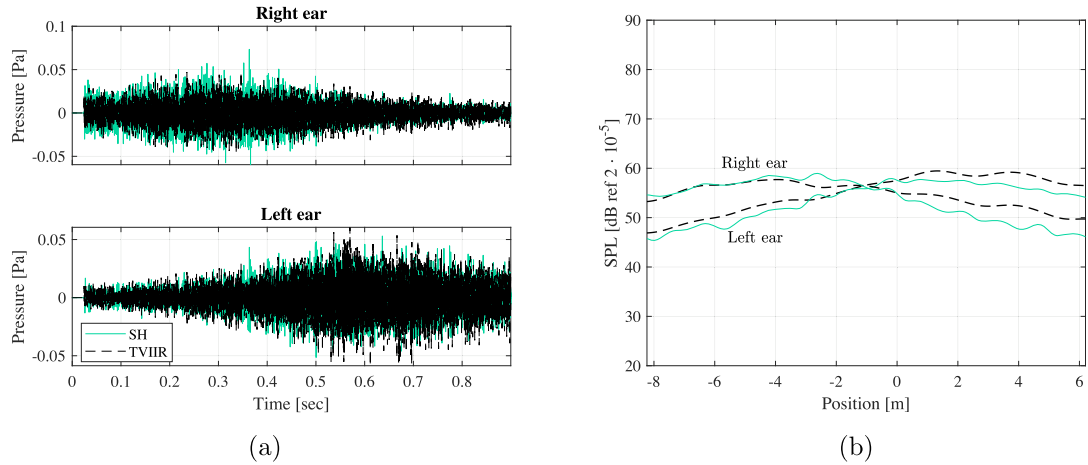
The online listening test consists of two sections. The first section evaluates the perceived speed using a continuous scale in the form of a slider, from 12.5 km/h to 75 km/h with a 12.5 km/h increment. In this section, only the 32nd order filter for the TVIIR and the 5th order spherical harmonics for the SH approach are used.

The second section proposes a pairwise comparison. The participants had to choose which one of the two replayed sounds is the most realistic. The pairs of sounds were presented in an arbitrarily chosen order. To limit the time of the test, not all possible combinations of sounds are presented. Thus, a total of 16 questions are selected.

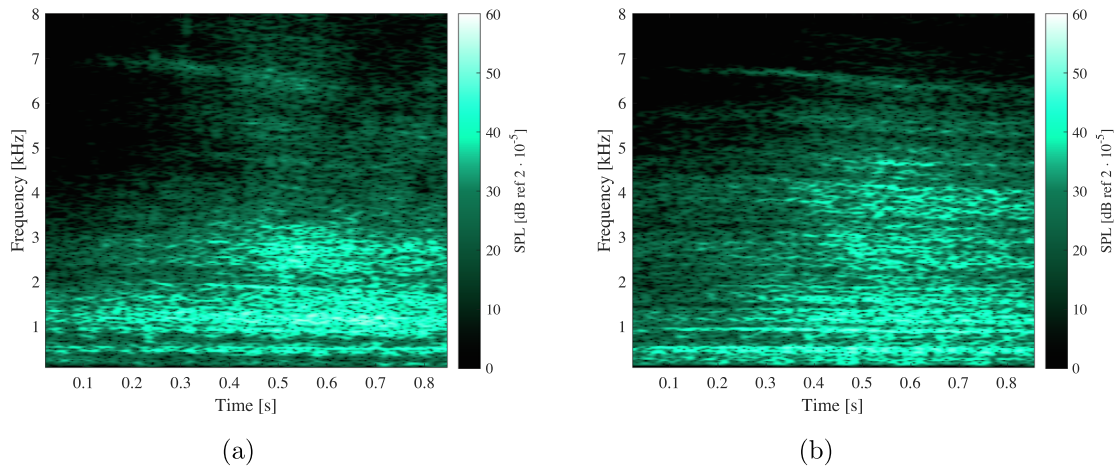
To evaluate the participant preference, the merit score (MS) is here used [44]

$$\text{MS}(y_i) = \frac{1}{N_s - 1} \sum_{j \neq i}^{N_s} P(y_i | y_j), \quad (14)$$

where  $N_s$  is the number of sounds available for the comparison and  $P(y_i | y_j)$  is the probability of a sound  $y_i$  being preferred over sound  $y_j$ . The MS describes the average of the preference of a certain sound  $y_i$  compared to the other sounds  $y_j$ .



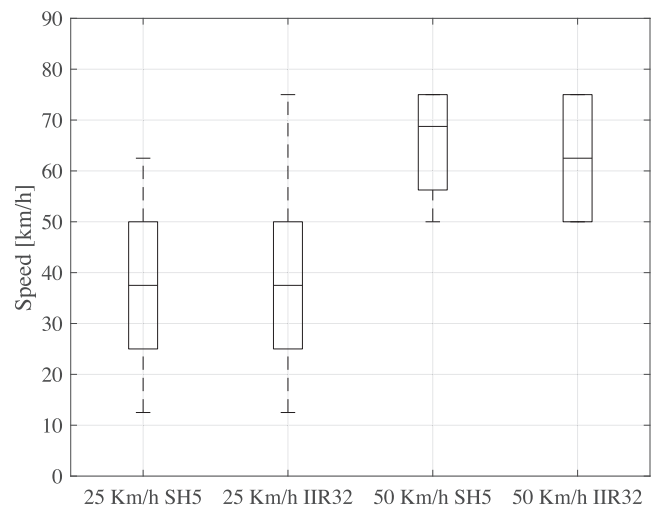
**Figure 13.** (a) Time signals and (b) sound pressure levels against the source position for both time-varying implementations.



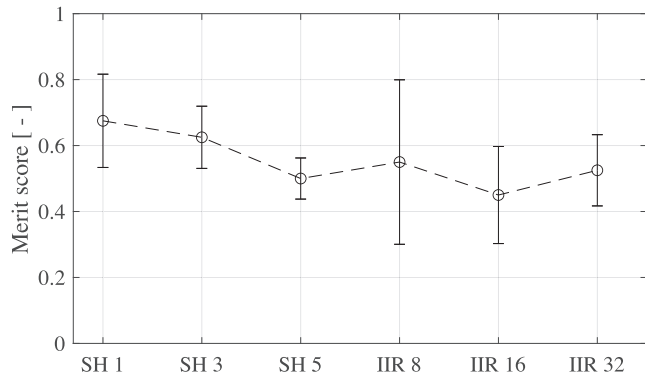
**Figure 14.** Spectrograms of the right ear output signal derived from the (a) time-varying IIR and (b) Spherical Harmonics approaches.

Figure 15 shows the box plot of the perceived speed with median, 25th, and 75th percentiles. The present jury test is performed in the absence of a reference sound with a specified vehicle speed, and therefore it is not expected that the participants identify the exact vehicle speed. Nevertheless, the two speeds are perceived in the correct order by the participants. The reason for an overestimation rather than an underestimation is admittedly the short total duration of the signals, 1.6 s for the 25 km/h scene and 0.9 s for the 50 km/h scene, which can be perceptually interpreted as the total exposure of the listener to the vehicle, thereby conveying a faster-moving scene. In terms of uncertainty, the perceived vehicle speed shows a larger spread across participants for the low speed than for the high speed. In addition, a larger spread is observed for the TVIIR approach than for the SH approach.

Figure 16 shows the merit score of the synthesized signals at 25 km/h. It can be observed that the 1st order SH has the highest merit score, followed by the 3rd order SH.



**Figure 15.** Apparent perceived speed versus actual speed for the two proposed approaches.



**Figure 16.** Merit score for the synthesized moving source signals at 25 km/h; the greater the merit score, the higher the preference; error bars represent the standard deviation.

In terms of This indicates that the SH approach is perceived as more realistic than the TVIIR. As observed in Figure 14, this can be attributed to the presence of the higher frequency content, which can provide more details to the audio scene. However, the observed merit scores are very close to each other indicating that both sound synthesis approaches can produce a similar outcome which might be induced by the very short samples. Therefore, a clear preference for a certain approach or order cannot be concluded from the presented jury test.

## 5 Conclusion

In this paper, two frameworks for the sound synthesis of a moving source using measurements in static condition and controlled environment were presented. The time-varying IIR filtering approach consists of the design and implementation of IIR filters and the spherical harmonics approach consists of representing the incident sound field into spherical harmonics. Both methods were implemented in an online-offline manner, relying on post-processing a set of measured transfer functions and manipulated in real-time for sound synthesis and predictions. The framework aims to provide accurate sound pressure levels in the far-field that comply with standards. Moreover, it allows for additional sources and filters (e.g. sound barriers) to create different moving source scenarios. Quantitative and qualitative results were shown for a pre-defined scene using an indoor pass-by noise test on an electrical vehicle.

Both methods have inherent simplifications. While the time-varying IIR filtering approach simplifies the spectral content by a certain polynomial coefficient order, the SH approach simplifies the signals by decomposing them into a basis of spherical harmonics. Nevertheless, despite both methods relying on a different strategy, they preserve the total amplitude of the original signal. Indeed, the results showed that both methods are capable of accurately and efficiently synthesize a moving source from propagation noise transfer functions recorded using a far-field microphone array. While the resulting sound pressure levels from both approaches were closely matching, the resulting

spectrograms displayed some differences, attributed to the inherent simplifications. Therefore, both techniques can be viewed as complementary. The time-varying IIR filtering approach allows for the accurate analysis of the sound pressure level from transfer path measurements and for a component-base troubleshooting. The SH approach allows for the spatialization of the audio for an immersive audio experience and is well suited to combine with acoustic source quantification techniques that goes beyond monopole sources such as in [45].

The jury test preliminarily validates the two approaches from a perceptual standpoint. In the evaluation, the resulting merit scores were very close across all tested samples which indicated that both sound synthesis approaches can indeed be similarly realistic. These results were affected by the total duration of the synthesized sound samples which are constrained by the dimensions of the room. Nevertheless, an increase in the total duration of the synthesized sound is feasible. This can be achieved in the SH approach by increasing the propagated distance using simulation (e.g. image source method) and in the TVIIR approach by including additional propagation noise transfer functions or by using extrapolation techniques [46].

## Conflict of interest

The authors declare no conflict of interest.

## Acknowledgments

We gratefully acknowledge the European Commission for its support of the Marie Skłodowska Curie program through the H2020 ETN PBNv2 project (GA 721615). The authors would also like to thank Mr. Fabio Bianciardi from Siemens Digital Industries Software for the support with Simcenter Testlab processing, Prof. Angelo Farina from the University of Parma for his valuable inputs and comments on the spherical harmonics processing and finally, the participants of the online jury test.

## References

1. ISO 362-1: Measurement of noise emitted by accelerating road vehicles – engineering method – part 1: M and N categories. 2015. Accessed 2018-10-12T12:52:21Z.
2. UN, ECE: Addendum 50: Regulation No. 51 Revision 3. Uniform provisions concerning the approval of motor vehicles having at least four wheels with regard to their sound emissions, 2016, p. 69.
3. ISO 362-3: Measurement of noise emitted by accelerating road vehicles — engineering method — Part 3: Indoor testing M and N categories. Technical Report. 2016.
4. K. Janssens, F. Bianciardi, L. Britte, P. Van de Ponsele: Pass-by noise engineering: A review of different transfer path analysis techniques, ISMA. 2014, p. 18.
5. P. Corbeels: Using component test bench measurements to predict pass-by noise contributions for trucks virtually, in: Aachen Acoustics Colloquium – AAC, 27–29 November 2017, Aachen. 2017.

6. T. Asakura, T. Miyajima, S. Sakamoto: Prediction method for sound from passing vehicle transmitted through building façade. *Applied Acoustics* 74 (2013) 758–769.
7. M. Hornikx, R. Waxler, J. Forssén: The extended fourier pseudospectral time-domain method for atmospheric sound propagation. *Journal of the Acoustical Society of America* 128 (2010) 1632–1646.
8. F. Georgiou, M. Hornikx, A. Kohlrausch: Auralization of a car pass-by using impulse responses computed with a wave-based method. *Acta Acustica united with Acustica* 105 (2019) 381–391.
9. D. Berckmans: Tools for the synthesis of traffic noise sources (technieken voor de synthese van verkeersgeluid). Ph.D. thesis, Katholieke Universiteit Leuven, 2010.
10. S. Guidati, R. Sottek, K. Genuit: Simulated pass-by in small rooms using noise synthesis technology. 2004.
11. E. Bongini, S. Molla, P.E. Gautier, D. Habault, P.O. Mattéi, F. Poisson: Synthesis of noise of operating vehicles: Development within SILENCE of a tool with listening features. In: B. Schulte-Werning, D. Thompson, P.E. Gautier, C. Hanson, B. Hemsworth, J. Nelson, T. Maeda, P. Vos (Eds.), *Noise and vibration mitigation for rail transportation systems*, Vol. 99, Springer, Berlin, Heidelberg, 2008, pp. 320–326.
12. J. Forssén, T. Kaczmarek: Auralization of traffic noise within the LISTEN project – preliminary results for passenger car pass-by, in: J. Kang (Ed.), *Euronoise 2009: Action on noise in Europe*, Institute of Acoustics, Edinburgh. 2009, p. 11.
13. A. Fiebig, R. Sottek, E. Kuczmarski: Auralization of road traffic noise and its value for environmental noise assessment, AIA-DAGA, Merano. 2013.
14. J. Maillard, J. Jagla: Auralization of non-stationary traffic noise using sample based synthesis – comparison with pass-by recordings, in: *Proceedings of the AIA-DAGA Conference on Acoustics*, Merano, Italy. 2013, p. 13.
15. E. Salomons, D. van Maercke, J. Defrance, F. de Roo: The harmonoise sound propagation model. *Acta Acustica united with Acustica* 97 (2011) 62–74.
16. R. Pieren, T. Butler, K. Heutschi: Auralization of accelerating passenger cars using spectral modeling synthesis. *Applied Sciences* 6 (2015) 5.
17. F. Yang: Traffic flow auralisation based on single vehicle pass-by noise synthesis, in: *Proceedings of the 23rd International Congress on Acoustics*, 9–13 September 2019, Aachen. 2019.
18. M. Alkmim, F. Bianciardi, G. Vandernoot, L. De Ryck, J. Cuenca, K. Janssens: Pass-by noise synthesis from transfer path analysis using IIR filters, in *Vibration Engineering for a Sustainable Future*, Springer. 2019.
19. M. Alkmim, G. Vandernoot, L. De Ryck, J. Cuenca, K. Janssens: Virtual pass-by noise sound synthesis from transfer path analysis data, in: *Audio Engineering Society (AES) International Conference on Automotive Audio*, 8–10 June 2022, Detroit. 2022, p. 7.
20. K. Janssens, P. Aarnoutse, P. Gajdatsy, L. Britte, F. Deblauwe, H. Van der Auweraer: Time-domain source contribution analysis method for in-room pass-by noise, in: *SAE Technical Paper 2011-01-1609*. 2011.
21. P. Van de Ponsele, K. Janssens, L. De Ryck: Source – transfer – receiver modeling approaches – a historical review of methods, in: *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 19–22 August 2012, New York. 2012, p. 11.
22. J.M. Jot, V. Larcher, O. Warusfel: Digital signal processing issues in the context of binaural and transaural stereophony, in: *Audio Engineering Society Convention 98*, Audio Engineering Society, 1995.
23. A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U.K. Laine, J. Huopaniemi: Frequency-warped signal processing for audio applications. *Audio Engineering Society Convention 108*, Audio Engineering Society, 2000.
24. A. Makur, S. Mitra: Warped discrete-Fourier transform: Theory and applications. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 48 (2001) 1086–1093.
25. B. Friedlander, B. Porat: The modified Yule-Walker method of ARMA spectral estimation. *IEEE Transactions on Aerospace and Electronic Systems AES-20* (1984) 158–173.
26. V.P. Valimaki: Discrete-time modeling of acoustic tubes using fractional delay filters, Ph.D. thesis, 1998.
27. J.O. Smith III: Digital audio resampling, <https://ccrma.stanford.edu/~jos/resample/>, Center for Computer Research in Music and Acoustics (CCRMA), Stanford University. 2020/9/17 (accessed 2023/7/6).
28. J. Blauert: *Spatial hearing: the psychophysics of human sound localization*, MIT Press, 1997.
29. L. McCormack, A. Politis: SPARTA & COMPASS: Real-time implementations of linear and parametric spatial audio reproduction and processing methods, in: *2019 AES International Conference on Immersive and Interactive Audio*, 27–29 March 2019, York. Audio Engineering Society, 2019.
30. T. Carpentier, M. Noisternig, O. Warusfel: Twenty years of IRCAM spat: Looking back, looking forward, *41st International Computer Music Conference (ICMC)*, Sep 2015, Denton, TX, United States. 2015, pp. 270–277.
31. M. Kronlachner, F. Zotter: Spatial transformations for the alteration of ambisonic recordings, in: *Proceedings of the 2nd International Conference on Spatial Audio*, 21–23 February 2014, Erlangen. 2014, p. 73.
32. B. Rafaely: *Fundamentals of Spherical Array Processing*. Springer Topics in Signal Processing, Springer-Verlag, Berlin Heidelberg, 2015.
33. J. Daniel, S. Moreau, R. Nicol: Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging, in: *Audio Engineering Society Convention 114*. Audio Engineering Society, Mar 1, 2003.
34. C. Schorkhuber, M. Zaunschirm, R. Holdrich: Binaural rendering of ambisonic signals via magnitude least squares. *Fortschritter der Akustik (DAGA)*, Munich. 2018.
35. A. Politis: *Microphone array processing for parametric spatial audio techniques*. Ph.D. thesis, Aalto University. 2016.
36. F. Zotter, M. Frank: *Ambisonics: a practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*, Springer Nature, 2019.
37. Steinberg: *Virtual studio technology (VST)*. 2021. <https://new.steinberg.net/vst-instruments/>, accessed 2021-08-17.
38. IEM: *IEM Audioplugins/IEMPluginSuite*, GitLab. 2022. <https://git.iem.at/audioplugins/IEMPluginSuite>, accessed 2022-08-12T14:53:54Z.
39. Epic Games: *Unreal engine*, 2021. <https://www.unrealengine.com>, accessed 2021-08-17.
40. D. Deboy, F. Zotter: Acoustic center and orientation analysis of sound-radiation recorded with a surrounding spherical microphone array, in: *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, Vol. 21, 6–7 May 2010, Paris. 2010.
41. B. Bernschütz: *A spherical far field HRIR/HRTF compilation of the neumann KU100*, DAGA Fortschritte der Akustik, Meran, Italy. 2013.
42. M. Alkmim, G. Vandernoot, J. Cuenca, K. Janssens, W. Desmet, L. De Ryck: *Audio samples: Real-time sound synthesis of pass-by noise: comparison of spherical harmonics and time-varying filters*, KU Leuven RDR, 2023.

43. M. Alkmim: Online listening test. 2021. <https://mansour.alk.gitlab.io/pbnv2jurytest/>. accessed 2021-07-08T13:37:08Z.
44. E. Parizet, N. Hamzaoui, G. Sabatie: Comparison of some listening test methods : A case study. Acta Acustica united with Acustica 19 (2005) 356–364.
45. C. Puhle, V. Becker, A. Jahnke, F. Knappe: Estimation of partial sound sources with non-spherical directivity for analysis of pass-by noise in hemi-anechoic indoor test benches, 9–10 June 2022, Berlin. 2022.
46. M. Alkmim, J. Cuenca, L. De Ryck, N. Kournoutos, A. Papaioannou, J. Cheer, K. Janssens, W. Desmet: A semi-circular microphone array configuration for indoor pass-by noise sound synthesis, in: 49th International Congress and Exposition on Noise Control Engineering (Inter. Noise), 23–26 August 2020, Seoul. 2020, p. 7.

**Cite this article as:** Alkmim M. Vandernoot G. Cuenca J. Janssens K, et al. 2023. Real-time sound synthesis of pass-by noise: comparison of spherical harmonics and time-varying filters. Acta Acustica, 7, 37.