Geometrical Acoustics Simulations for Ambisonics Auralization of a Car Sound System at High Frequency

Daniel Pinardi
Dept. of Engineering and Architecture
University of Parma
Parma, Italy
daniel.pinardi@unipr.it

Kseniia Riabova
Dept. of Engineering and Architecture
University of Parma
Parma, Italy
kseniia.riabova@unipr.it

Marco Binelli
Dept. of Engineering and Architecture
University of Parma
Parma, Italy
marco.binelli@unipr.it

Angelo Farina
Dept. of Engineering and Architecture
University of Parma
Parma, Italy
angelo.farina@unipr.it

Jong-Suh Park
Automotive R&D Division
Hyundai Motor Company
Seoul, South Korea
jspark@hyundai.com

2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA) | 978-1-6654-0998-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/I3DA48870.2021.9610977

***Abstract*—Nowadays, numerical simulations are the key tool for car audio systems and interiors design, since they allow estimating the acoustic response, thus the listening experience. This requires simulating the transfer path between each loudspeaker of the sound system and the listening position, which may be described, for instance, through a spherical harmonic expansion of the sound field; hence, permitting to auralize the sound system through real-time binaural decoding of the Ambisonics format over a head-tracked Head Mounted Display (HMD). In this paper, two numerical methods were compared for calculating the spatial impulse responses in High Order Ambisonics (HOA) format: ray tracing and pyramid tracing, respectively implemented in Comsol Multiphysics and Ramsete. The latter is equipped with a built-in post-processor capable to compute multichannel impulse responses in various formats including Ambisonics up to fifth order. An external post-processing tool was developed to calculate the Ambisonics impulse responses of Comsol Multiphysics simulations. First, the two methods were cross validated with two test cases: free field propagation and a room model. Results were compared in terms of time of arrival (ToA) of the direct sound, reverberation times, octave band spectra and accuracy of the source localization. Afterwards, the sound system of an existing high-end car was simulated. The results were compared with a set of experimental measurements performed inside the cockpit of the real car, employing a spherical microphone array.**

Keywords— Ambisonics, auralization, automotive, car sound system, high frequency, pyramid tracing, ray tracing, simulation

I. Introduction

The auralization technique is widely employed to design and evaluate noise and sound inside cars. In the past, this method was relying mainly on binaural approach: a dummy head was placed at the listening position for the recording and the playback was performed through headphones or stereo-dipole systems [1], [2]. [3], [4], [5]. However, binaural recordings suffer from head-locked reproduction, making them suboptimal for a complete Virtual Reality (VR) experience.

To improve the auralization, Ambisonics spatial audio may be employed: it allows head-tracked playback with three degrees of freedom of the head rotations. The Ambisonics theory [6] consists in encoding the signals recorded by a microphone array into different output channels, corresponding to virtual microphones having directivity patterns described by the Spherical Harmonics (SH) functions [7]. The original approach, named First Order Ambisonics (FOA), made use of four SH. With the aim of improving the spatial resolution, the method was later extended to High Order Ambisonics (HOA), having a number of channels N that increases with the square of the Ambisonics degree n:

$$N = (n+1)^2 \tag{1}$$

The comparative measurements described in this work were done with the Eigenmike32™ (EM), a HOA microphone array [8] featuring 32 capsules arranged over a sphere of $84\ mm$ diameter. Car sound system auralizations based on measurement with Ambisonics method and compact microphone arrays can be found in [9], [10]. However, the current trend is to abandon the experimental approach in favor of numerical simulations, so that the sound system performance can be evaluated even before prototypes of the car are built.

In this work, a method for the auralization of a car sound system employing third order Ambisonics obtained through numerical simulations is presented. Simulations are performed at medium and high frequencies with two different geometrical acoustics methods: Ray Tracing and Pyramid Tracing. First, the two solutions are compared for two reference test cases, free field propagation and indoor propagation, for validating several quantities: time of arrival of the direct sound, directivity and correct localization of the sound sources, the reverberation times and octave band spectrum. Then, the complete simulation of a car sound system installed in a large sedan is performed with both methods. Results are validated with reference measurements taken with the EM inside the car.

The paper is organized as follow. In section II, theory recalls regarding geometrical acoustics simulations and Ambisonics spatial audio are provided. In section III and section IV, the ray tracing and pyramid tracing processing are described. In section V and in section VI, the two methods are employed for the test cases and for simulating the car sound system, respectively. Section VII summarizes the conclusions.

II. Definitions and Theory

A. Geometrical Acoustic Methods

In the low-frequency region, the acoustic field inside an enclosed space is dominated by separate modes, whereas at higher frequencies the sound field becomes diffuse, and it is dominated by stochastic outcome of overlapping modes. In a room, low and high frequency regions are commonly divided by the Schroeder frequency f_{Sc} [11]:

$$f_{Sc} = 2000\sqrt{\frac{T_{60}}{V}} \tag{2}$$

where T_{60} is the reverberation time and V is the room volume. For small rooms, like a car cockpit, f_{Sc} falls around $300 - 400\ Hz$ [12]. In the analyzed case, $f_{Sc} = 335\ Hz$ considering a reverberation time of 0.1 s (see Fig. 14) and a car volume of $3.55\ m^3$. Generally, wave-based simulations, which solve the wave propagation or Helmholtz equation with Finite Elements Method (FEM) or Boundary Elements Method (BEM), are employed in the low frequency region. As frequency increases, these simulations become computationally heavier and less stable, since the number of elements depends on the wavelength.

On the opposite, ray tracing and pyramid tracing are geometrical acoustics methods applicable at higher frequencies. The sound wave is partitioned using rays or beams propagating along straight paths and carrying an initial amount of energy. The energy is dissipated, reflected, or scattered whenever a ray or a beam interacts with a boundary. Sound sources are characterized by the power and the directivity in frequency bands; boundary conditions are imposed as frequency-dependent reflection and scattering coefficients.

Ray tracing (RT) implies that a sound source emits an arbitrary number of sound rays. When these rays hit a receiver of a finite volume, their Time of Arrival (ToA), Direction-of-Arrival (DoA) and energy in frequency bands are stored. Instead, when using pyramid tracing sound travels bounded in geometrically diverging pyramids and the receiver is a point. This reduces the number of beams to be launched and, consequently, the required computation time. The software employed in this work are Comsol Multiphysics for Ray Tracing [13] and Ramsete for pyramid tracing [14].

B. Ambisonics encoding

The DoA of each ray, or pyramid, hitting a receiver are required to convert the simulations results into Ambisonics impulse responses. In both cases, this information consists in the three normalized direction vectors in Cartesian coordinates, n_x, n_y, and n_z. Thus, each ray/pyramid arrival detected by the receiver is encoded into SH by relying on the theoretical equations in Cartesian coordinates [15].

When using a microphone array instead, the conversion of the raw signals captured by the capsules into Ambisonics format can be performed in two different ways, linear processing [16] or parametric processing [17], [18], [19], [20], [21]. In this work, the first approach was used, thus beamforming is computed with a matrix of Finite Impulse Response (FIR) filters, calculated in frequency domain with the regularized Kirkeby inversion [22]:

$$H_{m,v,k} = \left[C^*_{m,d,k} \cdot C_{m,d,k} + \beta_k \cdot I\right]^{-1} \cdot \left[C^*_{m,d,k} \cdot A_{d,v} \cdot e^{-j\pi k}\right] \quad (3)$$

where $m = [1, \dots, M]$ are the capsules; $v = [1, \dots, V]$ are the virtual microphones; k is the frequency index; $d = [1, \dots, D]$ are the DoA of the sound waves; the matrix C is the complex response of each capsule m for each direction d; the matrix A defines the frequency independent amplitude of the target directivity patterns; $e^{-j\pi k}$ introduces a latency that ensures filters causality; $\cdot$ is the dot product; I is the identity matrix; $[\,]^*$ denotes the conjugate transpose; $[\,]^{-1}$ denotes the pseudo-inverse; β is a frequency-dependent regularization parameter [23].

The matrix C of (3), also known as spatial array response, can be obtained either with an experimental, theoretical, or numerical approach. The latter, solves numerically the interaction between sound waves and the geometry of the array in a simulation, usually of type FEM or BEM. The theoretical method relies on the analytical solution of the equations describing the diffraction of sound waves over the surface of a microphone array [24], [25]. Finally, the experimental approach used in this work, consists in measuring the microphone array from many DoA arranged in a spherical design geometry [26] inside an anechoic room, rotating the microphone array in front of a loudspeaker with a two-axis turntable. A full explanation of the method can be found in [27].

The Ambisonics format employed in this work is compliant with the current standard *AmbiX* [28]. The coefficients of the target directivity matrix A are those of the SH functions, usually defined as follows [29]:

$$A_{d,v} = \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-v)!}{(n+v)!}}\, P_n^v(\cos\theta) e^{iv\varphi} \quad (4)$$

where (θ, φ) are the angles of each direction d, respectively elevation and azimuth; n is the degree of the SH, an integer value ≥ 0; v is the order of the SH, comprised in the range $[-n \leq v \leq +n]$; P_n^v are the associated Legendre polynomials [29].

III. Ray Tracing Post-Processing

Comsol built-in post-processor does not provide Ambisonics results. Therefore, an external Matlab software has been developed for computing the Impulse Responses (IRs) in Ambisonics format up to fifth order.

Each ray detected by the receiver generates a Dirac Delta function (ideal impulse), which is filtered, scaled, and delayed. An octave bands filter bank was used, as shown in Fig. 1, with standard center frequencies defined by:

$$f_c = 1000 \cdot 2^o \quad (5)$$

where o are the integer numbers from -5 to 4. The filters are phase-matched at the crossing frequencies to be reconstructive, thus providing a flat spectrum ($\pm 1.5\ dB$ in the range $30\ Hz - 20\ kHz$), shown in Fig. 2. This is the same filter bank implemented in Ramsete.

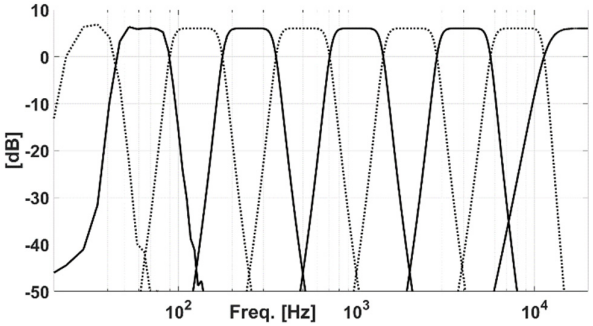

Fig. 1. Spectrum of the filter bank, separated octaves.

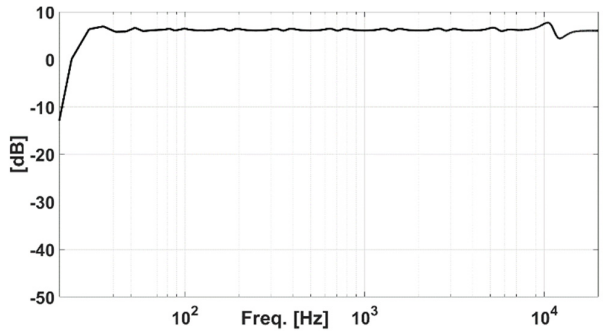

Fig. 2. Spectrum of the filter bank, summed octaves.

Then, each filtered impulse is delayed by the time of arrival of the ray at the receiver; the sound pressure p [Pa] is obtained by summing at each arrival the ten contributions, scaled as follow:

$$p = \sum_{f_c} \sqrt{\frac{Q(f_c) \cdot d_R \cdot c \cdot \rho}{V_R}}, \tag{6}$$

where $Q(f_c)$ is the residual power carried by the ray at each center frequency; d_R is the length of the segment (travel path) covered by the ray inside the receiver; c is the speed of sound; ρ is the air density; V_R is the volume of the spherical receiver; f_c are the octave bands center frequencies. Finally, the rays are encoded into SH with the previously described technique.

IV. Pyramid Tracing Post-Processing

In Ramsete, the arrival of a pyramid over a point receiver can be treated in two different ways:

- Discrete arrival, storing the ToA, DoA, sound intensity (ten octave bands) and history of previous reflections.
- Contribution to the diffuse field, ignoring the DoA but adding the intensity (ten octave bands) to the cumulated intensity storage for a given interval of time of arrivals.

All the energy carried by each pyramid is discrete and, at each reflection, a fraction of this energy is converted to diffuse energy, according to the value of the scattering coefficient, whose default value is 0.1. After a predefined maximum number of reflections, all the energy is considered diffuse. Consequently, directional information and precise ToA are only available for the initial direct sound and early reflections, whilst for the late energy arrivals everything is mixed in the diffuse field. Fig. 3 shows the typical result of a Ramsete simulation: pink lines represent the discrete energy arrivals and black lines the diffuse energy.

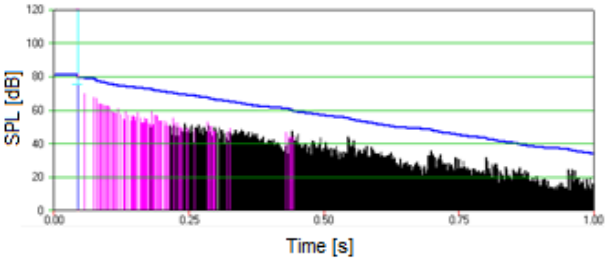

Fig. 3. Hybrid energetic impulse response in Ramsete.

The energetic quantity being traced in Ramsete is Sound Intensity I $[W/m^2]$, hence the sound pressure p results:

$$p = \sqrt{I \cdot \rho \cdot c} \tag{7}$$

where I is the sound intensity carried by a pyramid that arrives at the receiver, ρ is the air density and c is the speed of sound.

The diffuse energy is treated by generating a number of discrete arrivals, randomizing the DoA for each time interval. Each arrival occurs with a randomized ToA within the corresponding time interval. The number of discrete, randomized-direction energy arrivals being employed is usually between 1 and 5 for each ms. Increasing this number, the resulting impulse response becomes smoother and more accurate, at the price of a longer processing. The random generation of the DoA makes the reconstructed sound field perfectly diffuse. This condition is verified only when Sabine's conditions for a perfect diffuse field are met. This effect can be attenuated by increasing the number of reflections after which the discrete processing is abandoned, and all the energy is considered diffuse.

The post-processor of Ramsete, named Ramsete View, was already capable to export results as Ambisonics Impulse Responses up to order five (36 channels).

V. Cross-Validation of Ray Tracing and Pyramid Tracing Methods

A. Free field propagation case

A free field propagation analysis was computed to validate the correct management of source power and directivity. First, an omnidirectional source was simulated, with power at each octave band defined in TABLE I. The source is in the origin and the receiver is located along the x-axis at 1 meter distance, hence $d_{SR} = 1\ m$.

TABLE I

Omnidirectional Noise Source Power in Octave Bands		
Freq. [Hz]	**Lw [dB re 1 pW]**	**w [Watt]**
31.5	92	0.0016
63	94	0.0025
125	96	0.0040
250	98	0.0063
500	100	0.0100
1000	102	0.0158
2000	104	0.0251
4000	106	0.0398
8000	108	0.0631
16000	110	0.1000

In Ramsete, the number of pyramids N_p launched is defined as follow:

$$N_p = 8 \cdot 2^s, \tag{8}$$

where s is an integer number ≥ 0. If $s = 0$, the unit sphere is divided into its eight octants. A value of $s = 10$ was used, hence $N_p = 8192$, and the same number of rays, N_r, were launched in Comsol RT. In both cases, an impulse response of $10\ ms$ with a resolution of $1\ ms$ was computed, since it is only required that the direct sound reaches the receiver, covering the distance of $1\ m$. As it is possible to see in Fig. 4, the two spectra are identical. The same time of arrival is calculated with both methods: the peak of the IR occurs after 140 samples, which is exactly $2.9\ ms$ as expected.

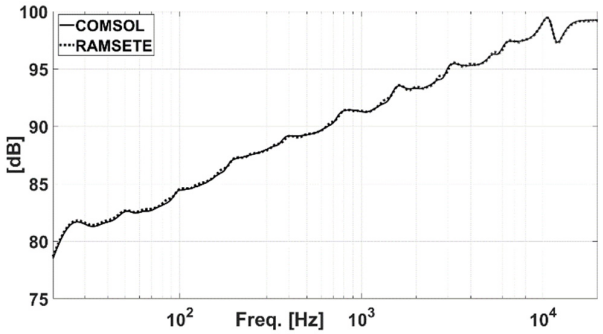

Fig. 4. Comparison of the spectra between Comsol and Ramsete for an omnidirectional noise source in free field.

Then, a flat noise source was simulated, having a power level of 110 dB at all octave bands and the directivity of a fourth order cardioid, shown in Fig. 5. The source is in the origin and three receivers are located at 1 m distance, one on-axis and the other two located horizontally off-axis, at 45° and 90° respectively. All the other parameters of the simulations are the same as in the previous case.

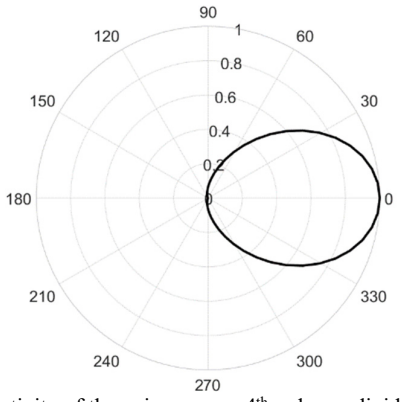

Fig. 5. Directivity of the noise source, 4th order cardioid.

The results are shown in Fig. 6. One can note that the solutions are always identical; time of arrivals are correctly calculated too.

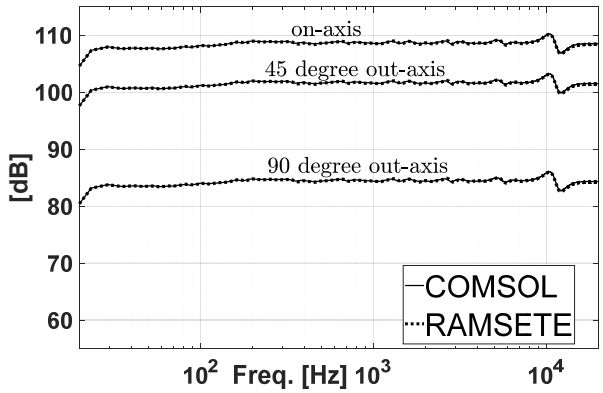

Fig. 6. Comparison of the spectra between Comsol and Ramsete for a flat directive noise source in free field.

B. Indoor propagation case

The room model shown in Fig. 7 was employed for checking the correct management of absorption coefficients and source localization. The geometry was designed in a 3D CAD software and imported in Comsol and Ramsete. A trapezoidal design was chosen to minimize the possibility that rays start bouncing back and forth between the walls (flutter echo): for this reason, only the floor and the ceiling are parallel. The distance between source and receiver is $d_{SR} = 2.29\ m$, and the volume is $35.4\ m^3$.

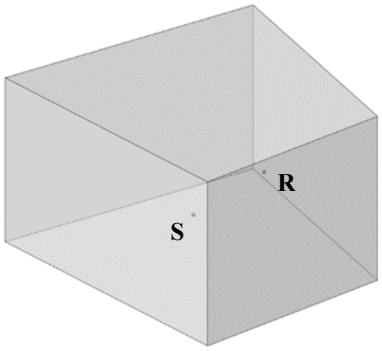

Fig. 7. Room model with a source (S) and a receiver (R).

Two materials were defined, and their absorption coefficients are shown in TABLE II. Scattering coefficient is the same for both software, and equal to 0.1 at all frequencies and materials.

TABLE II

Absorption Coefficients - Room Model		
Freq. [Hz]	**Floor**	**Walls and Ceiling**
31.5	0.02	0.06
63	0.04	0.12
125	0.08	0.12
250	0.1	0.15
500	0.12	0.2
1000	0.14	0.18
2000	0.16	0.3
4000	0.18	0.4
8000	0.2	0.5
16000	0.22	0.8

In Comsol, the wall condition was defined as "Mixed Diffuse and Specular Reflection", with the same absorption coefficients for diffuse and specular reflection. A human voice source was used in this case. An impulse response 1 s long was simulated with both software, with a resolution of 1 m, launching $N_p = N_r = 8192$ pyramids and rays, respectively. The result is shown in Fig. 8, and the matching between the two methods is almost perfect above 500 Hz, while there is a difference up to 10 dB towards lower frequencies. However, other numerical methods are usually employed below 1 kHz, such as FEM simulations [30].

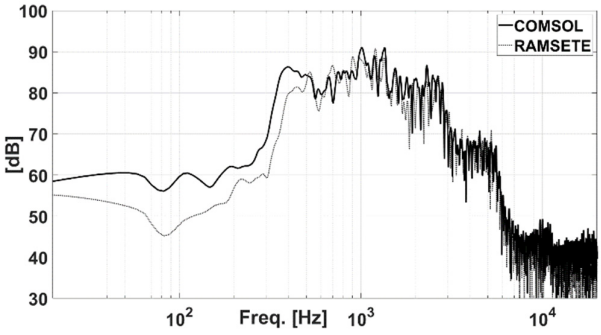

Fig. 8. Comparison of the spectra between Comsol and Ramsete for a directive noise source in a room.

The reverberation times RT30 were computed for each octave band, showing a good matching above 1 kHz and a general tendency to converge towards high frequencies, as one can note in Fig. 9.

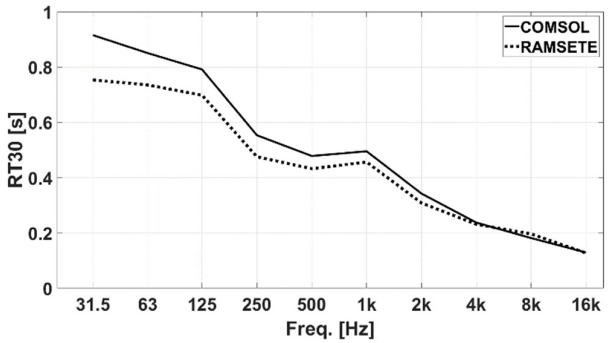

Fig. 9. Comparison of reverberation time between Comsol and Ramsete for a directive noise source in a room.

Finally, the capability of the two methods to correctly localize the position of the sound source was evaluated. For this task, it was opted to calculate the full band (20 Hz – 20 kHz), equirectangular color maps of Sound Pressure Level (SPL), evaluated on a monospaced grid having a resolution of 1 degree, and relying on a Plane Wave Decomposition beam-former [31]. The IRs were exported in Ambisonics 3rd Order (16 channels) and processed; results are shown in Fig. 10 and Fig. 11 for Comsol and Ramsete, respectively. In both cases, the direct sound and two reflections (wall and floor) are detected. The accuracy of the source localization is almost perfect: the deviation is 3° on the azimuth and 1° for the elevation. The position of the sound source, $\vartheta = 153.4°$ and $\varphi = 12.6°$, is marked with an asterisk.

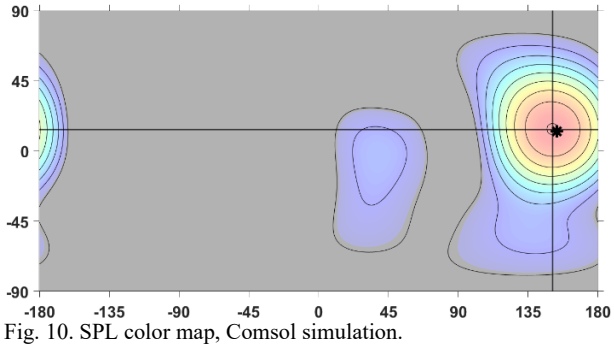

Fig. 10. SPL color map, Comsol simulation.

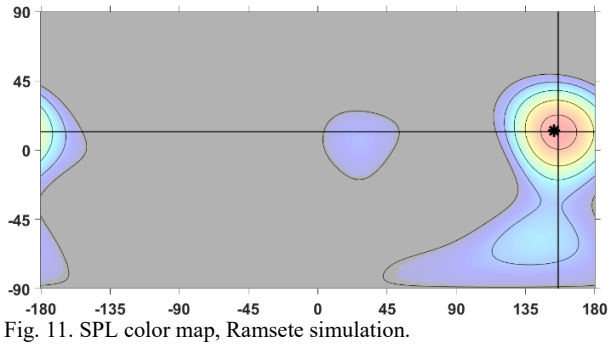

Fig. 11. SPL color map, Ramsete simulation.

VI. Auralization of a Car Sound System

A. Geometrical Acoustics Simulations

Three types of loudspeakers were employed in the simulations: tweeter, midrange and central. Firstly, their directivity was measured. Each device under test (DUT) was mounted on a wooden baffle installed over a turntable controlled via Ethernet and measured in an anechoic room with at 1 W the Exponential Sine Sweep (ESS) technique [32]. A Bruel&Kjaer omnidirectional microphone type 4189 was positioned on-axis respect to the loudspeaker at 1 m distance. The schematic of the measurement system is shown in Fig. 12.

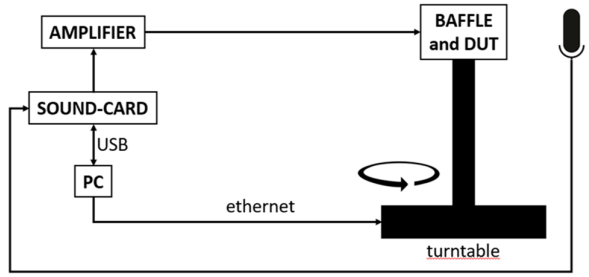

Fig. 12. Loudspeaker directivity measurement system.

Once IRs are obtained, the directivity and power of each source is computed, applying an averaging in octave bands, and silencing the rear radiation, since the loudspeakers radiate only on one side when mounted on the car. Power and directivity of the sources in octave bands constitute the input of the simulations.

Then, a simulation of the car cabin was performed for each sound source, for a total number of seven simulations, since the sound system features seven independent channels: one central, two surrounds (tweeter) and a pair of midrange and tweeter (mid+tw) in each of the four doors. Four receivers were positioned in the car model at the same time, allowing to obtain the Ambisonics IRs at the four seats for each source. The spherical receiver employed in Comsol has a radius $r_R = 0.1$ m. The simulations have been calculated for 0.5 s to include the whole reverberation, with a resolution of 1 ms, and launching $N_p = 16384$ pyramids in Ramsete and $N_r = 16384$ rays in Comsol Multiphysics.

The surface condition was set to "mixed diffuse and specular reflection" in Comsol, whilst in Ramsete the number of deterministic reflections (discrete arrivals) was set to five. Seven different materials were defined: seats, windows, hatshelf, roof, dashboard and central tunnel, doors, and car trim. The absorption coefficient of the windows was set to 0.01, whilst the values for the other surfaces of the car cockpit cannot be disclosed. The method employed for obtaining the absorption coefficients will be discusses in a future paper. The scattering coefficient is 0.1 at all frequencies for all materials.

The mesh was calculated including at least six elements per wavelength, hence the maximum element size λ is provided by:

$$\lambda_{max} = \frac{c}{(6 \cdot f_{\max})} \tag{9}$$

where $c = 343\ m/s$ is the speed of sound; $f_{\max} = 1\ kHz$ is the maximum simulated frequency. Therefore, we get $\lambda_{max} = 57.2\ mm$.

The post-processing of these simulations produced four auralization matrices, one for each seat, called SPK2SH and having dimensions 7x16, where 7 is the number of independent channels and 16 is the number of SH for the Ambisonics third order. The auralization can be performed in a VR environment, by using an HMD equipped with a head-tracker. The binaural rendering is obtained by convolving in real-time the SPK2SH matrix with a SH2Binaural matrix, which has a number of inputs depending on the Ambisonics order (1) and two outputs (the ears). The SH2BIN matrix is obtained by measuring the individualized HRTFs in an Ambisonics listening room [33], equipped with 16 loudspeakers located at 1.5 m from the listeners and by employing a special pair of in-ear MEMS microphones, named "Exofield" [34]. A more exhaustive explanation can be found in [35].

B. Reference measurement of the car sound system

The sound system was measured in the four seats with the EM microphone array, playing sequentially an ESS through each of the seven independent channels. By means of a breakout box, the test signal was delivered directly to each loudspeaker employing two ultra-flat, low noise, class AB amplifiers (QSC CX-168), connected to a 32 channels soundcard (Antelope ORION-32). The measurement setup is shown in Fig. 13.

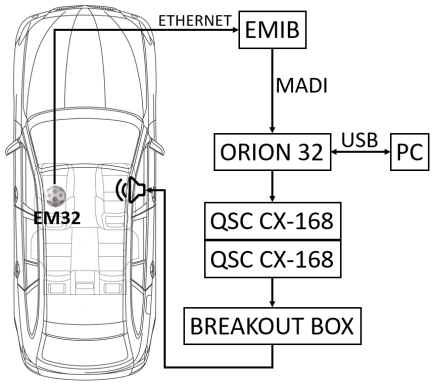

Fig. 13. Car sound system measurement setup.

A Multiple-Input Multiple-Output (MIMO) matrix of impulse responses was measured for each seat: it describes the transfer functions between the inputs (seven independent channels) and the outputs (32 capsules of the microphone array). This matrix, namely SPK2MIC, is convolved with the beamforming matrix H from (3), namely MIC2SH, having dimension 32x16. This operation produces the auralization matrix SPK2SH, having the same dimension of the simulated one, hence 7x16. The subsequent processing for auralization is the same previously described.

C. Analysis of the results

First, the reverberation times and the octave band spectra were calculated and compared for the two numerical simulations, i.e., Comsol and Ramsete, and the experimental measurements. For this analysis, only the first channel of the Ambisonics format was used, since it synthesizes a virtual omnidirectional microphone.

The reverberation time, shown in Fig. 14, was averaged among all the sources and the seats to increase the reliability. One can note that both numerical solutions are accurate already above $250\ Hz$, with a deviation of $\pm 0.05\ s$ respect to the experimental measurements. Instead, as expected, they are not accurate at very low frequencies, resulting in an underestimation of the reverberation time. As already mentioned, other solutions should be employed in that frequency range.

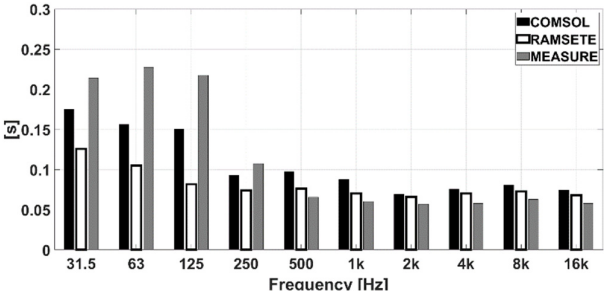

Fig. 14. Octave band reverberation times.

Then, octave band spectra were calculated. The results are presented separately for the seven independent sources, but averaged among the four seats: surround-left (Fig. 15), surround-right (Fig. 16), central (Fig. 17), mid+tw front-left (Fig. 18), mid+tw front-right (Fig. 19), mid+tw rear-left (Fig. 20), and mid+tw rear-right (Fig. 21). One can note a good matching between the two numerical solutions and the experimental measurement, with deviations of $\pm 1\ dB$ above 500 Hz.

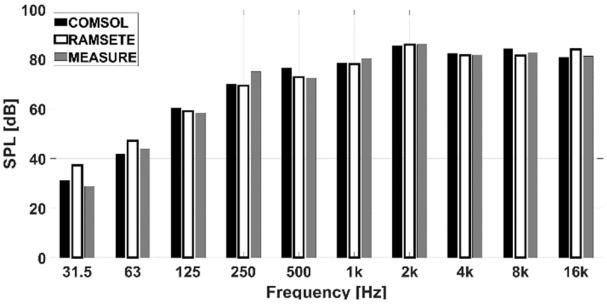

Fig. 15. Octave band spectrum, surround left.

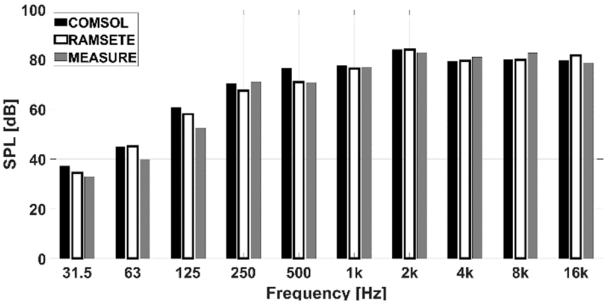

Fig. 16. Octave band spectrum, surround right.

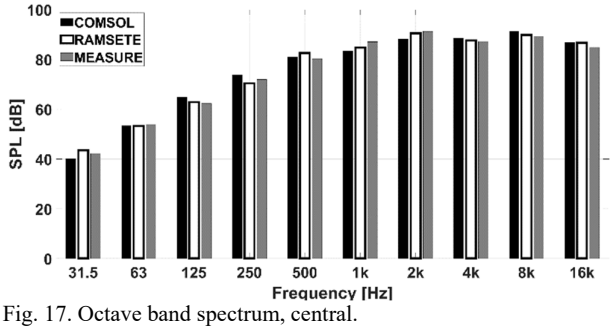

Fig. 17. Octave band spectrum, central.

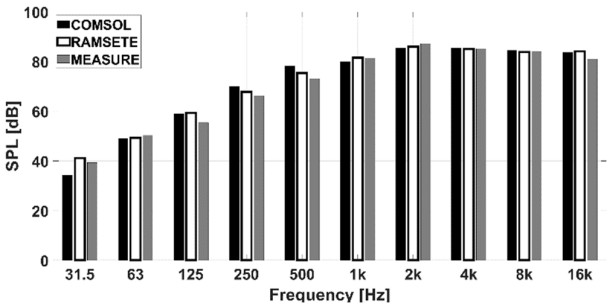

Fig. 18. Octave band spectrum, mid+tw front left.

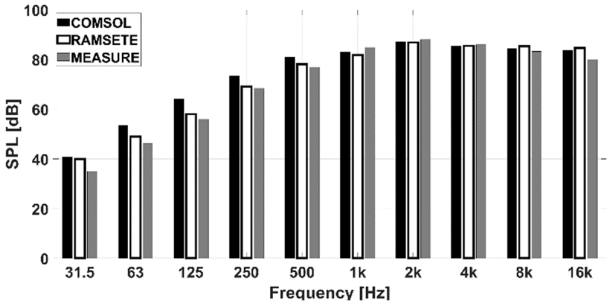

Fig. 19. Octave band spectrum, mid+tw front right.

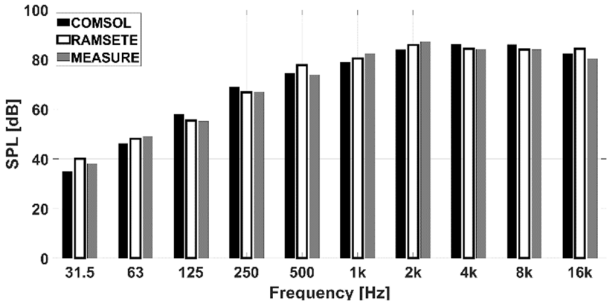

Fig. 20. Octave band spectrum, mid+tw rear left.

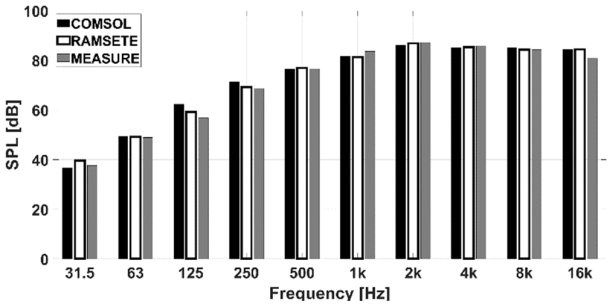

Fig. 21. Octave band spectrum, mid+tw rear right.

Finally, a spatial analysis was performed by calculating the color maps of SPL in octave bands for each source and position, and for the three methods, ray tracing, pyramid tracing, and experimental measurement. The background of the color map is an equirectangular image taken with a dual lens panoramic camera mounted in the same position of the microphone array at each recording point. As an example, the SPL color maps are shown in Fig. 22, Fig. 23, and Fig. 24 for the central loudspeaker at front-left seat, in the octave band centered at 2 kHz. One can note that the visualization of the source is accurate and consistent among the three methods.

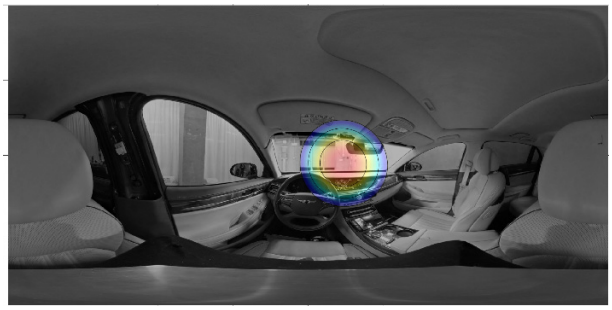
Fig. 22. SPL color map of central loudspeaker, recording position FL, 2 kHz octave band, numerical simulation (Comsol).

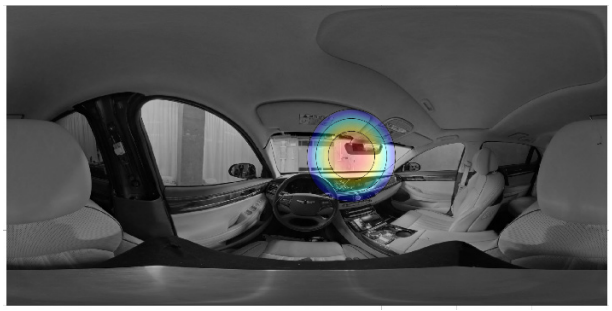
Fig. 23. SPL color map of central loudspeaker, recording position FL, 2 kHz octave band, numerical simulation (Ramsete).

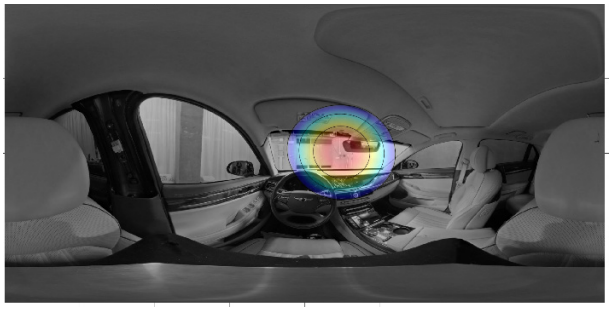
Fig. 24. SPL color map of central loudspeaker, recording position FL, 2 kHz octave band, experimental measurement.

VII. Conclusions

Two numerical methods for predicting the acoustic Impulse Response of an environment employing High Order Ambisonics spatial audio were compared: ray tracing, implemented in Comsol Multiphysics, and pyramid tracing, implemented in Ramsete. The built-in post-processor of Comsol Multiphysics revealed unable to calculate the Ambisonics Impulse Responses, therefore an external post-processor was developed in Matlab.

First, the two methods were compared in the free field case and with a room model, allowing to validate a variety of information: time of arrival of the direct sound, directivity, power and positioning of the sound sources, reverberation time and spectrum of the impulse responses.

Then, the complete sound system of a car was simulated. Experimental measurements were also performed as reference data, employing a spherical microphone array positioned in each of the four seats. At first, reverberation times and octave bands spectra of the sound sources were compared among the three methods to validate the results, showing that both numerical solutions are accurate already above 250 Hz, which is consistent with the Schroeder frequency calculated for this problem. Equirectangular color maps of the spatial distribution of Sound Pressure Level were calculated with a

Plane Wave Decomposition beam-former for assessing the correct visualization of the sound sources, which is accurate and consistent among the three methods.

In conclusion, it must be mentioned the main difference still present between the two numerical methods: Ramsete revealed to be approximately 50 times faster than Comsol Multiphysics in terms of calculation time.

REFERENCES

[1] E. Granier, M. Kleiner, B.-I. Dalenbäck and P. Svensson, "Experimental Auralization of Car Audio Installations," *Journal of Audio Engineering Society,* vol. 44, no. 10, pp. 835-849, 1996.

[2] A. Farina and E. Ugolotti, "Subjective comparison of different car audio systems by the auralization technique," in *103rd AES Convention*, New York, 1997.

[3] A. Pyzik and A. Pietrzyk, "High frequency modelling of a car audio system," in *144th AES Convention*, Milano, 2018.

[4] F. Malbos, M. Bogdanski and M. Strauss, "Virtual Reality Experience for the Optimization of a Car Audio System," in *AES International Conference on Automotive Audio*, Neuburg an der Donau, 2019.

[5] F. Malbos, M. Bogdanski and M. Strauss, "Loudspeaker Simulations in a Car Cabin," in *Proc. of the 2015 COMSOL Conference*, Grenoble, 2015.

[6] M. A. Gerzon, "Periphony: With-height Sound Reproduction," *Journal of the Audio Engineering Society,* vol. 21, pp. 2-10, 1973.

[7] N. M. Ferres, "An Elementary Treatise on Spherical Harmonics and Subjects Connected with them," *London: Macmillan and Co.,* 1877.

[8] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *Proc. on IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* vol. 2, pp. 1781 - 1784, 2002.

[9] A. Farina and E. Ugolotti, "Subjective comparison between Stereo Dipole and 3D Ambisonics surround systems for automotive applications," in *16th AES Conference*, Rovaniemi, 1999.

[10] S. Tervo, J. Pätynen, N. Kaplanis, M. Lydolf, S. Bech and T. Lokki, "Spatial Analysis and Synthesis of Car Audio System and Car Cabin Acoustics with a Compact Microphone Array," *Journal of the Audio Engineering Society,* vol. 63, no. 11, pp. 914-925, 2015.

[11] H. Kuttruff, Room Acoustics 4th edition, London and New York: Taylor & Francis, 2001.

[12] M. Kleiner and J. Tichy, Acoustics of Small Rooms, Taylor & Francis, 2014.

[13] COMSOL, Acoustics Module User's Guide.

[14] A. Farina, "RAMSETE - a new Pyramid Tracer for medium and large scale acoustic problems," in *Proc. of Euro-Noise Conference*, Lyon, 1995.

[15] A. Farina, "Explicit formulas for High Order Ambisonics," August 2017. [Online]. Available: http://www.angelofarina.it/Aurora/HOA_explicit_formulas.htm.

[16] A. Farina, S. Campanini, L. Chiesi, A. Amendola and L. Ebri, "Spatial Sound Recording With Dense Microphone Arrays," *55th AES Conference,* August 2014.

[17] S. Delikaris-Manias and V. Pulkki, "Parametric Spatial Filter Utilizing Dual Beamformer and SNR-Based Smoothing," in *AES 55th International Conference: Spatial Audio*, 2014.

[18] V. Pulkki, "Directional audio coding in spatial sound reproduction and stereo upmixing," in *Proc. 28th AES International Conference*, Pitea, 2006.

[19] V. Pulkki, "Spatial Sound Reproduction with Directinal Audio Coding," *Journal of the AES,* vol. 55, no. 6, pp. 503-516, 2007.

[20] S. Berge and N. Barrett, "High Angular Resolution Planewave Expansion (HARPEX)," in *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, 2010.

[21] V. Pulkki, A. Politis, M. V. Laitinen, J. Vilkamo and J. Ahonen, "First-order directional audio coding (DirAC)," in *Parametric Time-Frequency Domain Spatial Audio*, 2017, pp. 89-138.

[22] O. Kirkeby, F. Orduna, P. A. Nelson and H. Hamada, "Inverse filtering in sound reproduction," *Measurement and Control,* vol. 26, no. 9, pp. 261 - 266, November 1993.

[23] H. Tokuno, O. Kirkeby, P. A. Nelson and H. Hamada, "Inverse filter of sound reproduction systems using regularization," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences,* Vols. E80-A, no. 5, pp. 809 - 820, 1997.

[24] L. McCormack, S. Delikaris-Manias, A. Farina, D. Pinardi and V. Pulkki, "Real-time conversion of sensor array signals into spherical harmonic signals with applications to spatially localised sub-band sound-field analysis," *144th AES Convention,* 2018.

[25] L. McCormack, S. Delikaris-Manias, A. Politis, D. Pavlidi, A. Farina, D. Pinardi and V. Pulkki, "Applications of spatially localized active-intensity vectors for sound-field visualization," *Journal of the Audio Engineering Society,* vol. 67, no. 11, pp. 840-854, 2019.

[26] D. Pinardi, "Spherical t-Designs for Characterizing the Spatial Response of Microphone Arrays," in *I3DA - International Conference on Immersive and 3D Audio*, Bologna, 2021.

[27] A. Farina, A. Capra, L. Chiesi and L. Scopece, "A Spherical Microphone Array for Synthesizing Virtual Directive Microphones in Live Broadcasting and in Post Production," *40th International Conference: Spatial Audio: Sense the Sound of Space,* 2010.

[28] C. Nachbar, F. Zotter and E. Deleflie, "Ambix - A suggested Ambisonics format," *Ambisonics Simposium,* 2011.

[29] E. G. Williams, Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography, Academic Press, 1999.

[30] D. Pinardi, A. Farina and J.-S. Park, "Low Frequency Simulations for Ambisonics Auralization of a Car Sound System," in *I3DA - International Conference on Immersive and 3D Audio*, Bologna, 2021.

[31] A. Politis, "Acoustical Spherical Array Processing Library," Department of Signal Processing and Acoustics, Aalto University, Finland, 2016. [Online]. Available: http://research.spa.aalto.fi/projects/spharrayproc-lib/spharrayproc.html#59.

[32] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," *108th AES Convention,* 2000.

[33] M. Binelli, D. Pinardi, T. Nili and A. Farina, "Individualized HRTF for playing VR videos with Ambisonics spatial audio on HMDs," *AES International Conference on Audio for Virtual and Augmented Reality,* 2018.

[34] JVC-Kenwood Corporation, "Exofield® Headphone Technology Replicates the Acoustic Space of a Room," 2018. [Online]. Available: http://pro.jvc.com/pro/pr/2018/ces/JVC_Exofield.html.

[35] M. Binelli, D. Pinardi, T. Nili and A. Farina, "Individualized HRTF for playing VR videos with Ambisonics spatial audio on HMDs," *AES International Conference on Audio for Virtual and Augmented Reality,* 2018.

D. Pinardi received the M.S. (Cum Laude) degree in Mechanical Engineering from University of Parma, Italy, in July 2016, with a thesis on loudspeaker modelling. In March 2020, he got the Ph.D. in industrial engineering with a thesis on the design of microphone, hydrophone and camera arrays for spatial audio recording. He is a research assistant of Prof. Angelo Farina from 2016, mainly specialized in spatial audio. Topics of his interest are design of sensor arrays, acoustics simulations and auralization, applied to the automotive field and underwater acoustics.

K. Riabova earned the M.S. degree in mechanical engineering from Kharkiv Polytechnic Institute, Ukraine, in June 2011 with a thesis on analytical solutions of contact problems for composite shells. She worked at RWTH Aachen University, Germany, in 2011 and at McGill University, Canada, in 2015 as a research intern. In 2017 she got a PhD in Industrial Engineering at the University of Parma, Italy, with a thesis on non-destructive testing for damage detection via vibration measurements.

She pursued her career in Parma as a Postdoc in the team of prof. Angelo Farina working on in-car sound field and microphone arrays simulations. Kseniia contributed to numerous research projects with her work on vibration analysis, numerical modelling with application to vibrations and acoustics, advanced material modelling, acoustic-structure, and fluid-structure interaction. She is the author of 10+ scientific publications.

M. Binelli received his B.S and M.S. degrees in Electronic Engineering from University of Parma, Italy, in 2003 and 2006, respectively. He received his Ph.D. from University of Parma in 2010, with a thesis on analysis and equalization technique in automotive acoustics.

Since 2010, he is employed as assistant researcher at University of Parma. His research topics are equalization, psychoacoustics, active noise control, microphones and loudspeakers arrays, VR and spatial audio.

A. Farina got his M.S. degree in civil engineering in December 1982 at the University of Bologna, Italy, with a thesis on the acoustics and vibrations inside a tractor cab. In 1987, he got a Ph.D. in Technical Physics at the University of Bologna with at thesis on experimental assessment of concert hall acoustics. He was a full-time researcher since 1st November 1986 at the University of Bologna and since 1st March 1992 at the University of Parma. He became Associate Professor on 1st November 1998, and he is Full Professor of Environmental Applied Physics since 1st May 2005 at the University of Parma, where he has the chair of Applied Acoustics and Technical Physics.

During his academic career, Angelo worked in several fields of Applied Acoustics, including noise and vibration, concert hall acoustics, simulation software and advanced measurement systems. In the last 10 years, he focused mostly on applications involving massive microphone and loudspeaker arrays. In 2008, Angelo Farina was awarded with the AES fellowship for his pioneering work on electroacoustic measurements based on exponential sine sweeps. He is author of more than 250 scientific papers and three widely employed software packages (Ramsete, Aurora plugins, DISIA).

J. S. Park received his B.S. and M.S. degrees in Astronomy at Seoul National University, Korea. After receiving the M.S. degree in Aeronautics and Astronautics at Stanford University in 1999, he got a Ph.D. in Mechanical Engineering at Georgia Institute of Technology in 2004, with a thesis on the prediction of vibrational stability in manufacturing processes.

After joining Hyundai Motor Group in 2004, he has been involved in various vehicle projects as a senior research engineer of the noise and vibration CAE team, focusing on the development of new simulation technologies especially in multiphysics such as fuel sloshing, horns, and loudspeakers. He is expanding his field of interest from the conventional technologies for vehicle noise reduction and sound quality enhancement to the convergence of diverse methodologies such as immersive sound, multisensory perception, and virtual reality to build digital development environments for future mobility.