Low Frequency Simulations for Ambisonics Auralization of a Car Sound System

Daniel Pinardi
Dept. of Engineering and Architecture
University of Parma
Parma, Italy
daniel.pinardi@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/I3DA48870.2021.9610959

***Abstract*—In this paper, a technique is described for obtaining the High Order Ambisonics (HOA) Impulse Responses (IRs) of an automotive infotainment system, relying on Finite Elements Method (FEM) simulations performed in COMSOL Multiphysics. The resulting HOA IRs are employed for auralizing the car sound system, either inside an Ambisonics listening room with a loudspeaker rig or with binaural rendering on a Head Mounted Display (HMD), benefiting from head-tracking and personalized Head Related Transfer Functions (HRTFs). This allows performing subjective tests before the prototype is built and preserving the auditory experience with a degree of realism unattainable with the static binaural approach. Measurements performed in a prototype vehicle with a spherical microphone array are compared to FEM simulations. A good agreement between numerical and experimental methods have been demonstrated.**

***Keywords*—Ambisonics, auralization, automotive, car sound system, Finite Elements Method, low frequency, simulation**

I. Introduction

Subjective tests are widely employed in the automotive industry for developing, tuning, and evaluating car sound systems. The standard approach relies on listening sessions and binaural measurements performed on a prototype vehicle to score the sound system performance. Binaural recordings can also be used for head-locked listening sessions performed outside the vehicle. Recently, the Virtual Reality (VR) technology started to diffuse also in the automotive field, allowing for enhanced and immersive listening sessions through the auralization technique, making use of spatial audio and head tracking for preserving the auditory experience and naturalness of perception. This innovative approach requires a microphone array to record the spatial information of the sound field and the availability of a prototype vehicle.

The auralization method was developed in the past relying mainly on binaural approach [1]: a dummy head records a stereo signal at the listening position and the playback is performed through headphones or stereo-dipole systems [2], [3]. In 2001, a paper was published describing a noise and vibration simulator [4]: a virtual environment capable to reproduce faithfully the experience of being inside a vehicle.

Despite the success of the traditional binaural method, several drawbacks are inherent in this approach. A dummy head relies on its Head Related Transfer Functions (HRTFs), which do not fit for every possible listener. In addition, headphones playback of a static stereo signal does not benefit of head tracking, resulting in a poor localization of noise sources. Individualization of HRTFs for binaural rendering with head tracking was discussed by the authors in [5].

A microphone array could provide significative improvements. Such systems preserve the directional information of the sound field by sampling the sound pressure in many positions of the space. The set of signals recorded by the capsules (usually known as A-format) is encoded into a spatial audio format, commonly Ambisonics [6], [7] or Spatial PCM Sampling (SPS) [8]. In this way, the directional information is preserved and can be further analyzed, e.g., through color maps [9], [10] or reproduced, either binaurally or with loudspeakers rigs [11], [12]. Despite the first usages of the Ambisonics method for auralizing car audio systems dating back to the nineties, the old binaural approach is still used [13], [14]. Nevertheless, the measurement of a car sound system with microphone arrays [15] requires the availability of a vehicle. On the opposite, the current trend is to abandon the experimental approach in favor of numerical simulations, for predicting the sound system performance of a new car model before prototypes are built.

In this work, a method is presented for auralizing a car sound system in the low frequency range, i.e. $20\ Hz - 1\ kHz$, employing Finite Elements Simulations (FEM) and Ambisonics spatial audio. The geometry of a spherical microphone array with 32 capsules is introduced in a COMSOL Multiphysics model of a car cockpit, allowing the wave fronts emitted by the loudspeakers to diffract over the rigid surface of the microphone array. The complex values of sound pressure are evaluated at the capsules position and then converted into Ambisonics Impulse Responses. Reference measurements performed with a spherical microphone array inside the existing car are employed for validating the results.

II. Definitions and Theory

This chapter provides definitions and theoretical basis for Ambisonics audio, and the encoding operation required to convert the signals into spatial audio.

A. Ambisonics Spatial Audio

The Ambisonics technique was developed in the seventies to describe the complete acoustical spatial information of the sound field in the measurement point. The Ambisonics format is obtained by converting the pressure signals recorded by the capsules of a dense microphone array into a set of virtual microphones having directivity patterns described by Spherical Harmonics (SH), mathematical functions of various orders, with orthonormal properties. Explicit formulation of the SH functions up to order five is reported in [16]. The Ambisonics format employed in this paper follows the standard AmbiX [17]: SN3D normalization and ACN channel numbering.

Microphone arrays featuring four capsules can encode First Order Ambisonics (FOA): four virtual microphones having directivity patterns described by SH of order 0 and 1, respectively pure pressure (omnidirectional spherical pattern) and pure particle velocity (“figure of eight” patterns aligned with the three Cartesian axes). For improving the limited spatial resolution of FOA, the approach was later extended to High Order Ambisonics (HOA), including more channels corresponding to virtual microphones having directivity patterns described by SH functions of higher orders. In this work, comparative measurements were done employing an Eigenmike32™ (EM), a HOA microphone array launched in 2010 by Gary Elko [18] and featuring 32 capsules arranged over a sphere of $84mm$ diameter. The signals recorded by the EM were encoded into Ambisonics 3rd order, which includes 16 SH.

The authors are aware that the size of the employed array is quite small if compared with the wavelengths of the considered frequency range. The choice allowed for comparing fairly the numerical simulations with experimental measurements with the aim of validating the methodology. An extension of the approach for improving the auralization through simulations is currently under development, by employing a larger microphone array model equipped with dozens of capsules to encode Ambisonics format including orders higher than three. A larger dimension allows extending the Ambisonics orders toward lower frequencies, while the increment in the number of capsules preserves their surface density, hence the minimum distance between them, allowing to extend Ambisonics orders toward higher frequencies.

B. *The Encoding*

For converting the Impulse Responses at the capsules of the microphone array (A-format) into Ambisonics (B-format), a beamforming operation is required. In this work, beamforming was performed relying on a linear processing: a matrix of Finite Impulse Response (FIR) filters synthesizes the required virtual microphones. The filter coefficients are obtained, in frequency domain, employing the regularized Kirkeby inversion [19]:

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

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.

The coefficients of the target directivity matrix A are those of the SH functions, usually defined as follows [20]:

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

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 [20].

The frequency dependent regularization parameter β_k [21], shown in Fig. 1, applies a constraint to the Kirkeby inversion, preserving the *SNR* at low and high frequency, where the capsules are respectively too close and too far for the beamforming. In particular, β_k assumes a small value at intermediate frequencies and a much larger value at very low and very high frequencies.

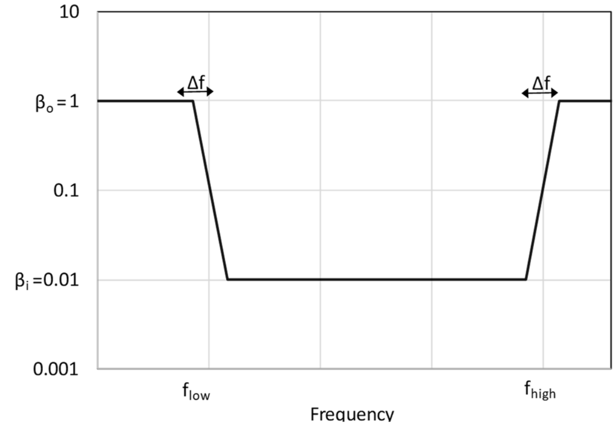

Fig. 1. Frequency dependent regularization parameter β_k.

The parameters of the Kirkeby inversion for the EM were defined as follow: $\beta_i = 0.01$ for the in-band regularization parameter, $\beta_o = 1$ for the out-band regularization parameter, $\Delta f = 0.3$ octaves for the transition bandwidth, $f_{low} = 20\ Hz$ for the lower transition frequency, $f_{high} = 14\ kHz$ for the higher transition frequency.

III. Microphone Array Response

The spatial response of the array (C matrix) provides the complex information of sound pressure at the capsules for many DoA of the sound wave and it is required to calculate the filtering matrix H of (1). There are three possible approaches for obtaining the C matrix:

- Experimental solution, measuring the microphone array in an anechoic chamber and imposing sound waves to arrive from many directions [22], [23] either rotating the microphone array or employing a moving loudspeaker.
- Theoretical solution, solving the analytical equations that define the interaction between sound waves and the geometry of the microphone array [9], [10], [24].
- Numerical solution, simulating the diffraction of sound waves against the surface of the array with Finite Element Method (FEM) or Boundary Element Method (BEM).

The theoretical solution of plane waves diffracted by a rigid sphere (i.e., the surface of a spherical array) represents in practice a sound source positioned in free field at infinite distance from the array. This condition inside the car is not met, since sound sources (i.e., loudspeakers) and reflective surfaces (i.e., panels and windows) are very close to the array. Therefore, the curvature of the wave fronts is significant when they hit the surface of the array, making the near field effect worth considering. For this reason, the experimental and numerical approaches were employed for obtaining the spatial array responses required to calculate the beamforming filters.

A. *Experimental Characterization of the Microphone Array*

The EM was measured in an anechoic room [22], [23] with a two-axis turntable and a loudspeaker positioned at 1-meter distance. The schematic of the experimental setup is shown in Fig. 2. The turntable is controlled by the PC via Ethernet. The test signal is sent from the PC to the interface board via Firewire. This system allows playing the test signal and synchronously recording all the 32 channels of the EM. The sound source is a Genelec studio monitor type 8351A.

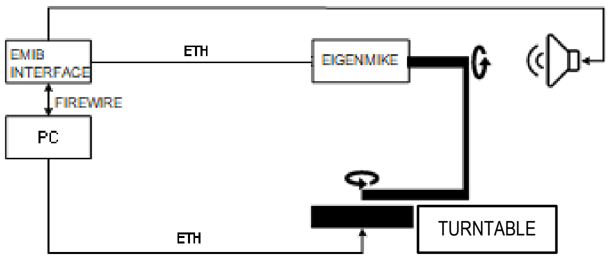

Fig. 2. Experimental setup for characterizing the microphone array.

The test signal is an Exponential Sine Sweep (ESS) [25], pre-equalized for flatting the spectrum of the sound source in the range $50\,Hz - 18\,kHz$ within $\pm 0.5\,dB$. The measurement was repeated for $D = 240$ DoA, arranged over a T-design distribution [26], [27] of order $T = 21$. After each measurement, the PC sends to turntable the new measurement direction and then the ESS is played and recorded.

IRs were calculated by convolving the recorded signals with the inverse sweep [25]. By applying the Fast Fourier Transform (FFT) of $K = 8192$ points on the IRs, the $C_{[d,m,k]}$ matrix of (1) is obtained, with dimensions $[240; 32; 8192]$. Finally, the solution of (1) provides the filtering matrix H for encoding the impulse responses of the sound system into Ambisonics, which can be then predicted or measured inside the prototype vehicle.

B. Numerical Characterization of the Microphone Array

As described in the introduction of this work, a model of a microphone array was introduced in the model of the car. This "virtual" array must be characterized similarly to the real one, for obtaining an encoding matrix to be employed with FEM simulations of the sound system. The two encoding matrices are not interchangeable: the measured one in fact corrects also phase and magnitude mismatches between the capsules, whilst numerical model considers an ideal response of the microphones.

To get comparable results, the geometry of the model is almost identical to the Eigenmike32™: a sphere of $84mm$ diameter with 32 capsules virtually placed in the same positions (Fig. 3).

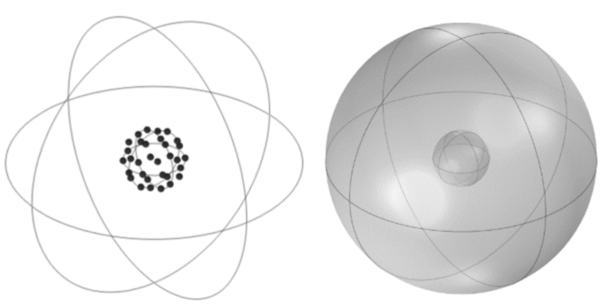
Fig. 3. COMSOL model of the virtual spherical microphone array.

The simulation was solved in COMSOL Multiphysics in the frequency range $20\,Hz - 4\,kHz$, with a resolution of $5.86\,Hz$. First, the method was validated by comparing the condition of Plane Wave Radiation (PWR) against the theoretical solution. The implementation of the theoretical solution of plane waves diffracted by a rigid sphere was found in the form of MATLAB code in [28]. The frequency responses of numerical and theoretical solutions are compared in Fig. 4. One can note that the two results are identical, ensuring that the FEM approach is reliable for calculating the spatial response of a microphone array.

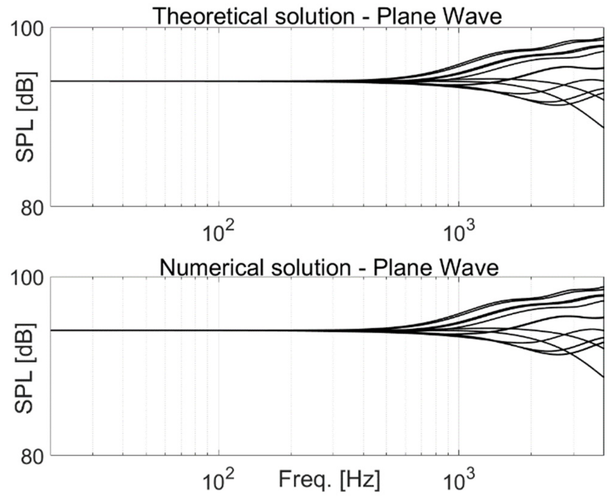

Fig. 4. Theoretical and numerical solutions, comparison of the frequency responses up to 4 kHz with plane waves.

Then, the PWR condition was substituted with Spherical Wave Radiation (SWR), thus positioning a virtual point source at the same distance of experimental setup, $1\,m$. In Fig. 5 the frequency responses for PWR and SWR are compared. The geometrical divergence of the spherical wave fronts is now considered, and it is possible to note a relevant difference, particularly below $1\,kHz$.

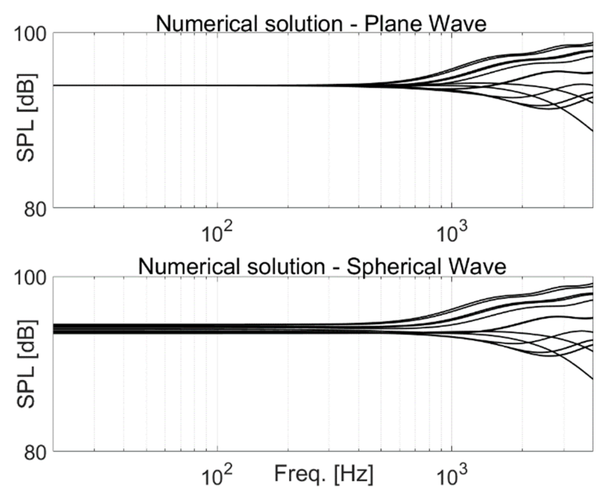

Fig. 5. Numerical solution, comparison of the frequency responses up to 4 kHz between plane and spherical waves.

Finally, the simulation with SWR was compared with the experimental measurement of the EM performed in the anechoic room at $1\,m$ distance (Fig. 6). For simplifying the visualization, the following capsules have been compared: 1, 9, 19, 21 and 29. One can note a good matching between the spherical waves model and the experimental measurement.

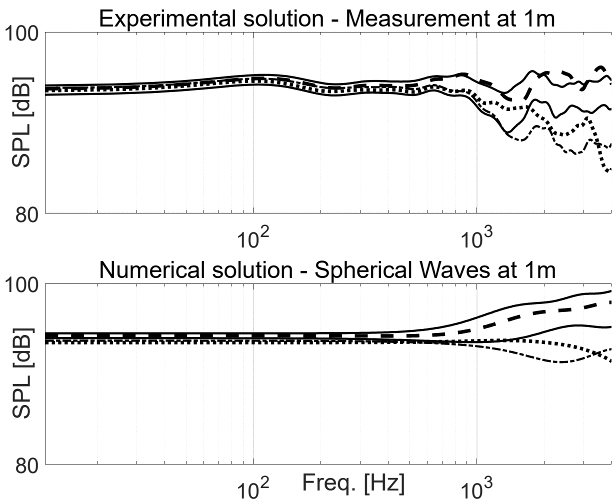

Fig. 6. Comparison of the frequency responses up to 4 kHz between anechoic measurement and spherical wave radiation simulation, distance 1 m.

Finally, the simulation was calculated for the same T-design distribution already employed for the anechoic measurement, with $D = 240$ DoA. Since the solution is performed in frequency domain, the complex matrix C is directly obtained.

IV. Measurement of the Car Sound System

The target car was a Hyundai large sedan, equipped with a top-level sound system featuring 17 loudspeakers driven by a 12-channels amplifier. Reference impulse responses were measured by placing the EM in each of the four seats and playing a sequence of 12 independent exponential sine sweeps of 30 s each one, feeding one after the other the 12 channels of the sound system. The measurement setup is shown in Fig. 7. To feed the sources bypassing the head unit and the DSP of the car, which apply the factory equalization, a breakout box was used. The breakout box is fed with the signals coming from two QSC CX-168, ultra-flat, low noise, class AB amplifiers, connected to an Antelope ORION-32 soundcard. The 32 signals coming from the EM are recorded synchronously by the same soundcard through the Eigenmike Interface Board (EMIB), which passes the audio stream over MADI protocol.

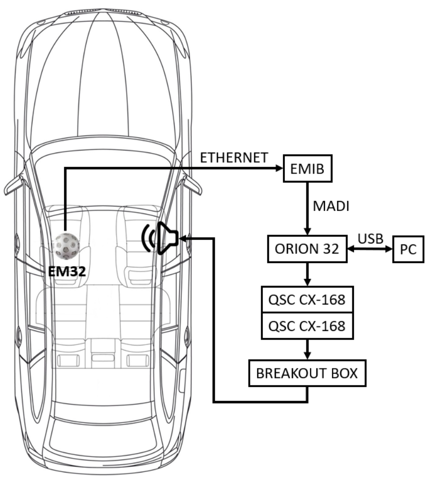

Fig. 7. Car sound system measurement setup.

The recorded ESS are convolved with the inverse sweeps to get the IRs. In this way, a Multiple-Input Multiple-Output (MIMO) matrix of impulse responses is obtained, describing the transfer functions between the 12 inputs (independent channels) and the 32 outputs (capsules of the microphone array). This matrix, called SPK2MIC, is convolved with the beamforming matrix H of (1), which can be referred as MIC2SH, being Ambisonics the output format. For the EM and Ambisonics 3rd order, the matrix H has dimension 32x16. The convolution between SPK2MIC and MIC2SH matrices produces the auralization matrix SPK2SH, of dimension 12x16. Four SPK2SH matrices have been measured, one for each seat.

To perform the auralization, the software employed for the binaural reproduction performs a further convolution between the Ambisonics signals and the SH2Binaural matrix, which in our case must have dimension 16x2. If the reproduction system is equipped with head tracking, the counter rotation of the sound field is applied in the SH domain before the binaural rendering. A more exhaustive explanation of the whole process can be found in [5].

V. Simulation of the Car Sound System

Three different types of loudspeakers were introduced in the simulations: one subwoofer (positioned in the trunk), four woofers (one for each door) and seven midranges (one for each door, one for the front central and two for the rear surrounds). The data of the loudspeakers were provided in the form of displacement as a function of the frequency, measured in the center of the cone with a Klippel Analyzer, a standard measurement tool in automotive industry. The data were imported into COMSOL Multiphysics, and the displacement was assigned to the surfaces of the loudspeakers, assuming a pistonic behavior in the low frequency range.

The frequency-dependent values of the complex acoustical impedance for each surface were provided by Hyundai Motor Company and cannot be disclosed.

In Fig. 8 the FEM model of the car is shown, with the spherical microphone array previously described positioned in front of each seat, in the same position of the EM during the measurement of the real car.

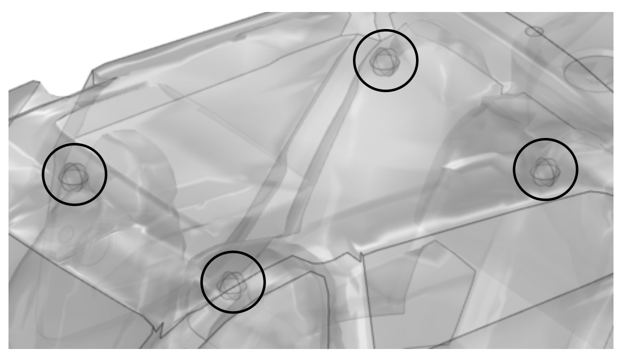

Fig. 8. FEM model of the car with a virtual spherical array in front of each seat

Twelve simulations were performed, one for each loudspeaker, in the frequency range $20\ Hz - 1\ kHz$, with a resolution of $5.86\ Hz$. The complex values of sound pressure were evaluated at each frequency step in the 32 capsules of each virtual array, by means of four datasets of type "Cut Point 3D". To define such datasets, it is required to provide the absolute position of the capsules in Cartesian coordinates, which can be calculated by knowing the position of the center of the array and the directions (azimuth, elevation) of each capsule.

Four SPK2MIC matrices were obtained by applying the Inverse Fast Fourier Transform (IFFT). Finally, the auralization matrices SPK2SH are computed, by convolving each SPK2MIC matrix with the filtering matrix MIC2SH obtained from numerical simulation (see Section III.B).

VI. Analysis of Results

A. Reverberation Times and Spectra

For comparing the results of the numerical simulations with the experimental measurements, the first channel of the Ambisonics format was used: it synthesizes a virtual omnidirectional microphone, corresponding to SH number 0 with ACN numbering.

First, the reverberation times in octave bands were calculated and compared (Fig. 9). To reduce the number of plots and to increase the reliability of the results, the four positions FL, FR, RL, RR were averaged. The reverberation time was averaged also among all the sound sources, as it is a property of the environment. One can note a good matching at all octave bands, with a difference never exceeding *50 ms*.

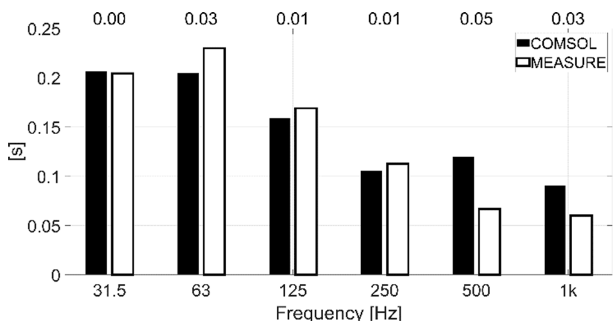

Fig. 9. Reverberation time in octave bands, all sources averaged.

Then, octave band spectra of SPL have been calculated separately for each source: surround-L (Fig. 10), surround-R (Fig. 11), woofer-RL (Fig. 12), woofer-RR (Fig. 13), woofer-FL (Fig. 14), woofer-FR (Fig. 15), subwoofer (Fig. 16), central (Fig. 17), midrange-FL (Fig. 18), midrange-FR (Fig. 19), midrange-RL (Fig. 20) and midrange-RR (Fig. 21). In the frequency range of utilization of each source, the deviation never exceeds 6 dB. In detail, midranges are supposed to be used from 300 Hz and above, woofers in the range 100 Hz – 300 Hz and the subwoofer in the range 20 Hz – 100 Hz.

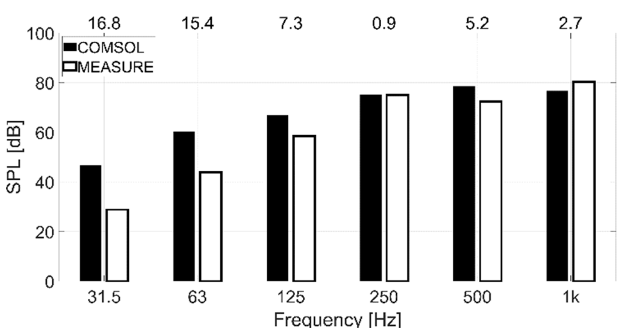

Fig. 10. Octave band spectrum, surround midrange left.

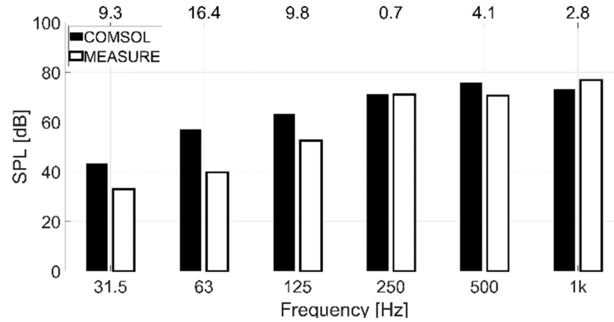

Fig. 11. Octave band spectrum, surround midrange right.

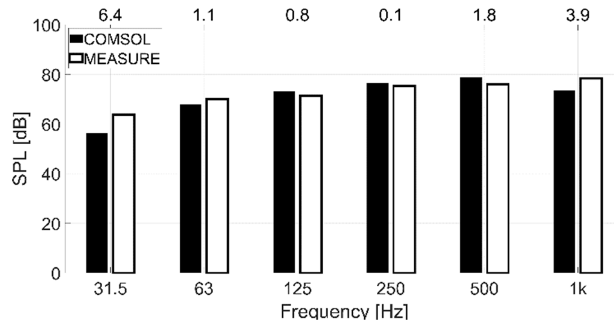

Fig. 12. Octave band spectrum, woofer rear left.

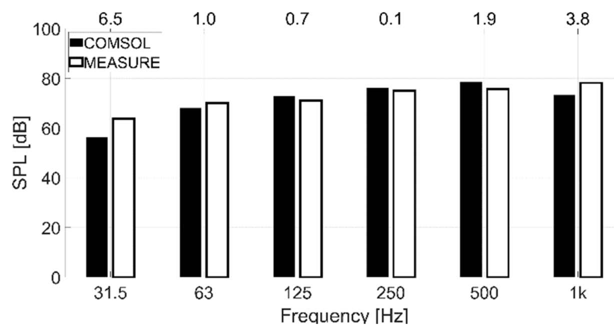

Fig. 13. Octave band spectrum, woofer rear right.

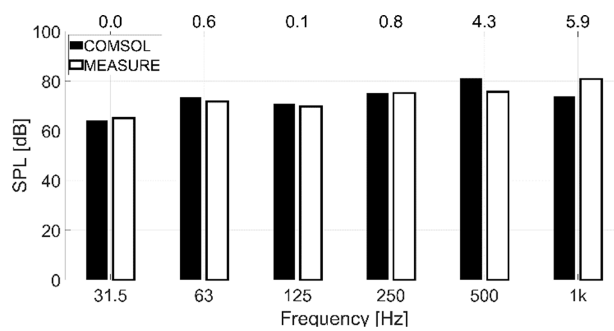

Fig. 14. Octave band spectrum, woofer front left.

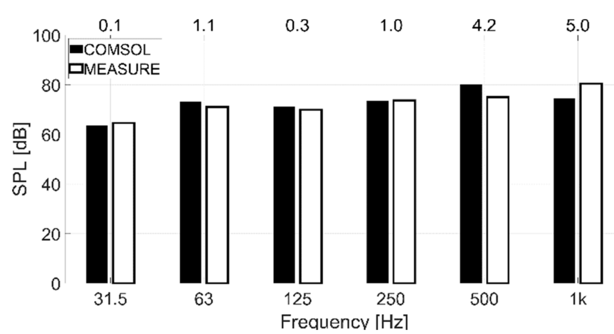

Fig. 15. Octave band spectrum, woofer front right.

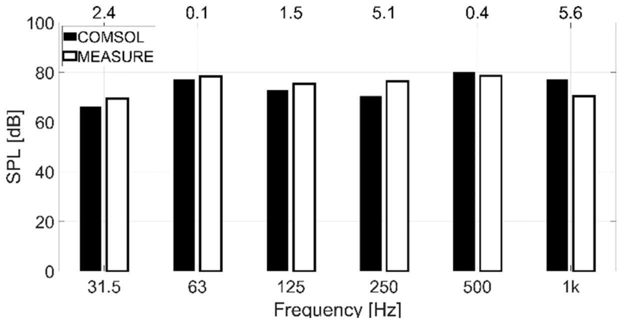

Fig. 16. Octave band spectrum, subwoofer.

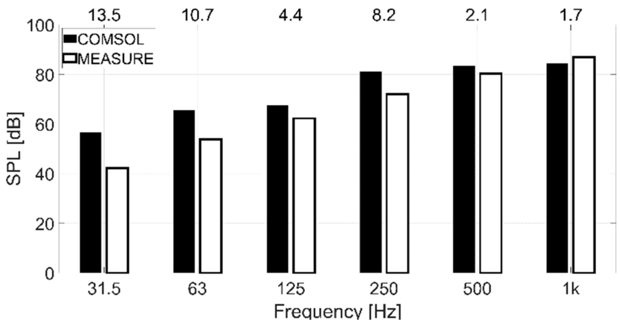

Fig. 17. Octave band spectrum, central midrange.

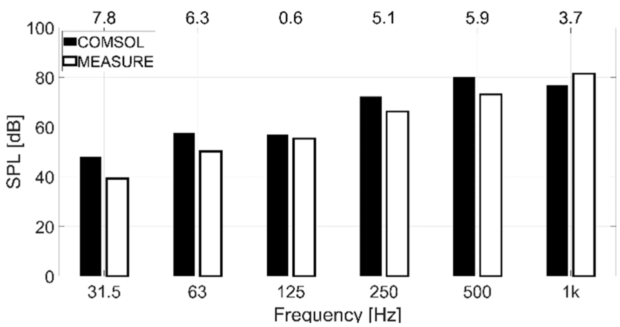

Fig. 18. Octave band spectrum, midrange front left.

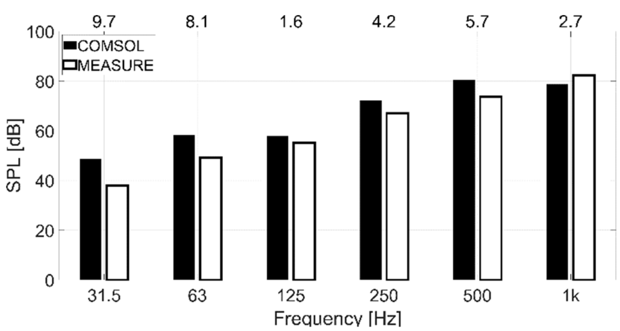

Fig. 19. Octave band spectrum, midrange front right.

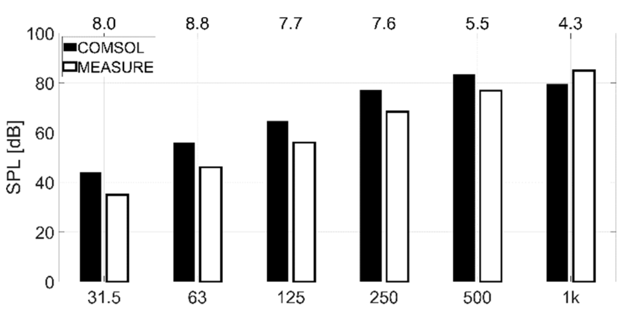

Fig. 20. Octave band spectrum, midrange rear left.

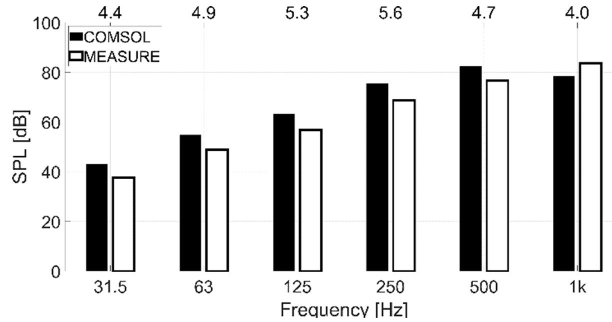

Fig. 21. Octave band spectrum, midrange rear right.

B. *Sound Sources Visualization*

Finally, a spatial analysis of the results was performed for assessing the correct localization of the loudspeakers, relying on a Plane Wave Decomposition beamformer [28] for calculating color maps of the SPL, evaluated on a monospaced grid having a resolution of 1 degree. The color maps are superimposed to an equirectangular image of the interior of the car. These pictures have been obtained with a compact, dual lens panoramic camera positioned in the same location of the microphone array during experimental measurement. The same images can be used also inside the VR environment for providing a visual feedback spatially coherent with the auralized sound.

The color maps are shown for the driver position (FL) and three sources, comparing the numerical and experimental results: woofer FR in Fig. 22, midrange FL in Fig. 23, and subwoofer in Fig. 24. One can note that the localization is accurate and consistent among the simulations and measurements.

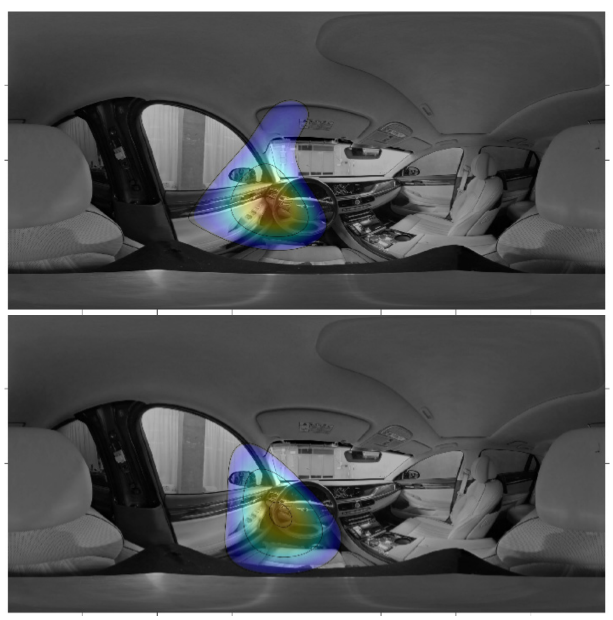

Fig. 22. SPL color map of woofer FL, seat FL, numerical result (above) and experimental result (below).

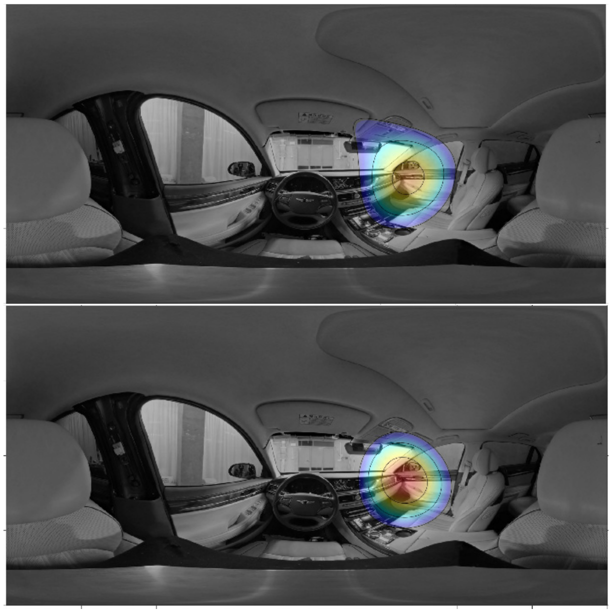

Fig. 23. SPL color map of midrange FR, seat FL, numerical result (above) and experimental result (below)

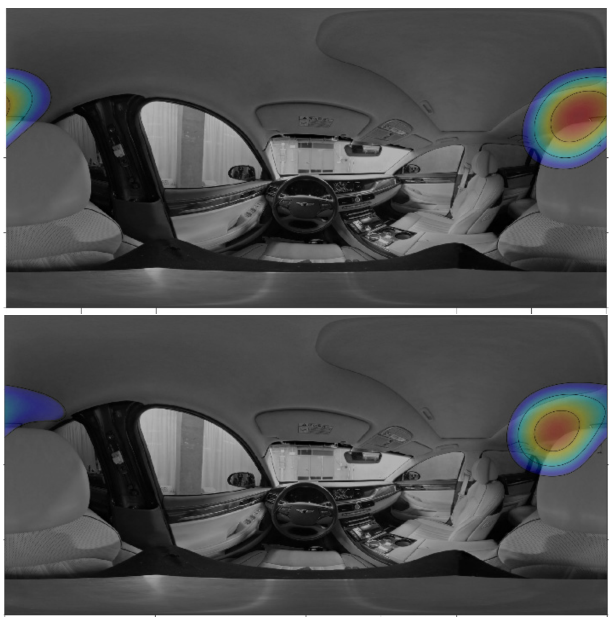

Fig. 24. SPL color map of subwoofer, seat FL, numerical result (above) and experimental result (below).

VII. AURALIZATION

Both experimental and numerical results can be employed for real-time auralization, allowing a listener to evaluate subjectively their similarity. The auralization includes a visual display of the equirectangular background picture taken from each listening positions. The image is displayed on a HMD by using the open-source program *pan360* by Andre Chen [29]. It has been modified by implementing the attitude data broadcasting in quaternion format via OSC protocol. These data are received by the IEM Scene-rotator VST plugin, which counter-rotates the HOA signals compensating for the head rotation. The convolution between the FIR filter matrices SPK2SH and SH2BIN is performed by another VST plug-in, such as *mcfx-convolver* [30]. The SH2BIN matrix is obtained by measuring the individualized HRTFs in an Ambisonics listening room [5], equipped with 16 loudspeakers and by employing a special pair of in-ear MEMS microphones, named "Exofield" [31]. All the processing can be performed real-time in a Digital Audio Workstation (DAW).

Informal comparison tests were conducted with six listeners by using the ABX method, presenting four music pieces, and repeating the ABX test 10 times for each piece. All the tests were conducted at the driver seat position. The results of the ABX tests are shown in TABLE 1.

TABLE 1

Number of correct answers to ABX test				
Subject	**Piece A**	**Piece B**	**Piece C**	**Piece D**
1	9	10	8	6
2	5	6	5	4
3	7	6	8	5
4	8	7	5	6
5	9	8	8	6
6	7	9	7	5

Significant results with a sequence of 10 choices require at least nine correct answers, which were obtained only for four of the 24 independent tests. In these cases, the numerical and the experimental results are clearly distinguishable. In all the other tests, the number of correct answers is less than 9/10 and, in many cases, approaches the perfectly indeterminate case of 5/10, which means that the listener was unable to discriminate between the numerical and the experimental filters.

The authors will publish a further step of the current research including listening tests inside the car, all four seat positions and more advanced analysis schemes, such as MUSHRA [32].

VIII. CONCLUSIONS

The authors proposed a solution for auralizing the car sound systems in the low frequency range based on FEM simulations combined with a virtual microphone array. The described methodology was successfully employed for auralizing the sound system of a large sedan, featuring three types of loudspeakers and twelve channels.

Numerical simulations performed in COMSOL Multiphysics were employed for predicting the acoustic response of the car cabin. The key point of the described technique consists in the introduction of a virtual spherical microphone array in the model of the car. In this way, the diffraction of the wave fronts against the rigid surface of the array was considered, allowing to get the complex values of sound pressure at the capsules. Numerical results were converted into High Order Ambisonics format with a matrix of FIR filters, previously calculated by means of a FEM simulation of the array. Such simulation was performed with a point source located at short distance, thus radiating spherical waves. This allowed the consideration the near field effect, which is fundamental for the correct localization of the auralized sound sources, and it is the second important contribution of this work.

Experimental measurements of the sound system were performed with an Eigenmike32™ placed over each of the four seats. These measurements were converted into Ambisonics format by means of a different filtering matrix, calculated with an anechoic measurement of the array at short distance. This allowed the two approaches to be compared.

Numerical and experimental results were compared in terms of reverberation times, octave band spectra, SPL color maps and listening tests. The reverberation times are well matched at all octave bands, with a maximum deviation of $50\,ms$. The SPL spectra showed a maximum deviation of $6\,dB$ in the octave bands of utilization of each source. Finally, SPL color maps were calculated employing a plane wave decomposition beamformer. The visualization of the sound sources demonstrated to be accurate and consistent between the two methods. Informal listening tests, conducted with the ABX method, support this claim. In most cases the auralizations based on simulations and measurements were indistinguishable from each other.

Acknowledgment

D. Pinardi and A. Farina would like to thank Hyundai Motor Company for technical support and for having provided a prototype car.

References

[1] K. Genuit, "Investigation and simulation of vehicle noise using the binaural measurement," in *SAE Noise and Vibration Conference, SAE 870959*, 1987.

[2] 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.

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

[4] G. Goetchius, C. Ketelhut, B. Smallwood and C. Eaton, "Subjective Evaluation of NVH CAE Model Predictions Using an Operator-in-the-Loop Driving Simulator," in *SAE Noise and Vibration Conference, SAE 2001-01-1590*, 2001.

[5] 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.

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

[7] M. A. Gerzon, "Ambisonics. Part two: Studio techniques," *Studio Sound,* vol. 17, no. 10, p. 60, 1975.

[8] A. Farina, M. Binelli, A. Capra, E. Armelloni, S. Campanini and A. Amendola, "Recording, Simulation and Reproduction of Spatial Soundfields by Spatial PCM Sampling (SPS)," *International Seminar on Virtual Acoustics,* 2011.

[9] 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.

[10] 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.

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

[12] F. Fazi, P. Nelson, J. Christensen and J. Seo, "Surround System Based on Three-Dimensional Sound Field Reconstruction," *125th AES Convention,* 2008.

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

[14] 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.

[15] A. Farina, D. Pinardi, M. Binelli, M. Ebri and L. Ebri, "Virtual reality for subjective assessment of sound quality in cars," *144th AES Convention,* 2018.

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

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

[18] 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.

[19] 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.

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

[21] 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.

[22] 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.

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

[24] Institute of Electronic Music and Acoustics, "IEM Plug-in Suite," [Online]. Available: https://plugins.iem.at/download/.

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

[26] R. H. Hardin and N. J. A. Sloane, "McLaren's Improved Snub Cube and Other New Spherical Designs in Three Dimensions," *Discrete and Computational Geometry,* vol. 15, pp. 429-441, 1996.

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

[28] 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.

[29] A. Chen, "Vive Cinema," [Online]. Available: https://github.com/openbigdatagroup/vivecinema.

[30] M. Kronlachner, "Multichannel Audio Plug-in Suite," [Online]. Available: http://www.matthiaskronlachner.com/?p=1910.

[31] 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 .

[32] I.-R. BS.1534-3, *Method for the subjective assessment of intermediate quality levels of coding systems,* ITU: International Telecommunication Union, Radiocommunication Assembly, 2015.

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. degree 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 sensors arrays, simulations and auralization, applied to automotive field and underwater acoustics.

Angelo 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.

Angelo is author of more than 250 scientific papers and three widely employed software packages (Ramsete, Aurora plugins, DISIA).

Jong-Suh Park received his BS and MS degree in Astronomy at Seoul National University, Korea. After receiving the MS 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 the area of 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 in order to build digital development environments for future mobility.