Full-Digital Microphone Meta-Arrays for Consumer Electronics

D. Pinardi, Andrea Toscani, M. Binelli, L. Saccenti, A. Farina, and L. Cattani

***Abstract*—Microphone arrays of various sizes and shapes are widely employed in consumer electronics devices such as smart speakers, speakerphones, smart TVs, smartphones, PCs, and headphones. Therefore, the possibility of creating arrays of arrays (i.e., meta-arrays), by easily connecting several smaller units, is particularly interesting: it allows for different shapes, with an increased number of microphones and improved performance, while keeping costs low. In this paper, full-digital microphone meta-arrays are presented: they are constituted by an arrangement of several triangular units, each featuring four digital Micro Electro-Mechanical Systems (MEMS) capsules connected in daisy-chain through the Automotive Audio Bus (A^2B). Two prototypes have been built: a planar meta-array made of seven triangles (28 capsules) and a 3-dimensional meta-array with 14 triangles (56 capsules). Numerical simulations were carried out and compared against experimental measurements performed on the prototypes, showing an excellent matching. Finally, the 3-dimensional meta-array was compared with a spherical microphone array widely considered the state-of-the-art equipment in the last decade, demonstrating its potential in telecommunication applications.**

***Index Terms*—Ambisonics, automotive audio bus (A^2B), beamforming, consumer products, digital MEMS microphones, meta-arrays, microphone arrays, planar arrays, spatial audio, 3D virtual microphone system.**

I. Introduction

MICROPHONE arrays are employed in consumer electronic devices for several applications, such as teleconferencing, speakerphone units, home assistants, PCs, smartphones, and smart TVs. Their use, combined with advanced processing algorithms, allows isolating a specific source (such as a person's voice) or purposefully filter out undesired sounds (such as background noise during videoconference meetings). Indeed, they are used for dereverberation or echo reduction [1], [2], [3], [4] and speech recognition [5]. Other applications include Active Noise Control (ANC) [6], [7], multi-channel audio recording [8], beamforming [9], [10], [11], acoustics simulations [12], and structural health monitoring [13]. Arrays of microphones also allow determining the Direction of Arrival (DoA) of sound [14], [15]. When designing such multichannel systems, the type of capsules and data transmission architecture are critical aspects to consider, as the gain and phase matching of the capsules have a significant influence on the audio quality. The geometry of the array, the number of capsules and their layout, instead, mostly affect the spatial accuracy and the working frequency range of beamforming. As a result, we see a steadily growing demand for systems capable of supporting more and more channels. Examples of massive multichannel microphone arrays can be found in [16], [17], [18].

Regarding the choice of the capsule type, most of today's solutions employ analog microphones, which can deliver high-quality audio signals. On the other hand, they entail bulky wiring and noise immunity problems, particularly in case of long wires connecting the capsules with the Analog-to-Digital (A/D) converters. In addition, analog capsules, pre-amplifiers, and A/D converters contribute to increase the cost of the system and design complexity. Conversely, low-cost digital MEMS microphones are more robust to electrical noise, usually at the price of worse acoustic performances (dynamic range and Signal-to-Noise Ratio). Another important aspect to consider during the development of microphones arrays is the geometry and the disposition of capsules. In literature this topic is widely debated: the array's performance, the spatial sampling and the robustness of the processing method is significantly impacted by the position of the capsules [19], [20], [21], [22], [23]. Furthermore, it is of great importance the prior study related to MEMS arrays based on the fundamental equations for semiconductors, such as the conservation of mass, momentum, and energy, and Maxwell's equations [24], [25].

In this work, the use of triangular array modules to develop full-digital meta-arrays, obtained by connecting several triangular microphone arrays, is investigated and descripted, concluding the work presented in [26]. In this previous work, Pinardi et al. presented a full-digital microphone array composed of MEMS microphones connected through A^2B. The A^2B bus was developed as a high bandwidth, bidirectional digital bus to address issues with audio distribution in automotive applications [27]. The authors also studied the positioning of MEMS capsules to identify the optimal layout in terms of beamforming directivity. Each array has four MEMS microphone capsules arranged in a pseudo-triangular form, and it constitutes a subordinate node in a multiple A^2B network.

A planar array with seven triangular units and 28 capsules has been simulated, prototype, and measured,

Manuscript received 6 December 2022; revised 28 February 2023; accepted 14 April 2023. Date of publication 18 April 2023; date of current version 18 August 2023. This work was supported by ASK Industries S.p.A. *(Corresponding author: Andrea Toscani.)*

D. Pinardi, Andrea Toscani, M. Binelli, L. Saccenti, and A. Farina are with the Department of Engineering and Architecture, University of Parma, 43125 Parma, Italy (e-mail: daniel.pinardi@unipr.it; andrea.toscani@unipr.it; marco.binelli@unipr.it; leonardo.saccenti@unipr.it; angelo.farina@unipr.it).

L. Cattani is with the Research and Development Department, ASK Industries S.p.A., 42124 Reggio Emilia, Italy (e-mail: luca.cattani@askgroup.global).

Digital Object Identifier 10.1109/TCE.2023.3267836

showing excellent agreement between numerical and experimental solutions. Two 3-dimensional meta-arrays have been simulated for assessing the optimal geometry: a regular dodecahedron with 12 triangular units and 48 capsules, and a truncated octahedron with 14 triangular units and 56 capsules, which ultimately demonstrated better performance. A prototype was built, measured, and compared with the Eigenmike-32 [28], a spherical microphone array featuring 32 capsules, and considered the reference system of the last decade for voice and music recordings. Virtual microphones were encoded, employing a Finite Impulse Response (FIR) filter matrix, for combining the signals coming from the microphone capsules, to obtain arbitrary directivity beams. Two different sets of virtual microphones were used in this work: 3D Virtual Microphone System (3DVMS) [29] and Ambisonics [30], allowing comparisons to be made between numerical simulations and experimental measurements.

The paper is arranged as follow: Section II provides the mathematical background of beamforming; Section III illustrates the methods employed for numerical and experimental approaches while in Section IV the principles of A^2B bus and multiple A^2B network architecture are described. Section V presents the practical aspects of simulation, construction, and measurement of meta-arrays. Section VI, eventually, summarizes the conclusions.

II. Signal Processing Background

The conversion of the pressure signals at capsules locations into virtual microphones having arbitrary directivity is performed with a linear processing based on a matrix of FIR filters computed with the regularized Kirkeby inversion [31] algorithm:

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

where $m = [1,\ldots,M]$ is the capsule index, $v = [1,\ldots,V]$ is the virtual microphone index, k is the frequency index, $d = [1,\ldots,D]$ is the index of the DoA of the sound wave; the matrix $\boldsymbol{C}$ is the complex array response, the matrix $\boldsymbol{A}$ defines the frequency independent amplitude of the target directivity patterns, $e^{-j\pi k}$ introduces a latency that ensures filters causality, the dot $(\cdot)$ is the scalar product, $\boldsymbol{I}$ is the identity matrix, $[\ ]^*$ denotes the conjugate transposition, $[\ ]^{-1}$ denotes the pseudo-inversion, β is a frequency-dependent regularization parameter [32].

The set of DoA employed for simulating and measuring the array responses is a spherical t-design geometry [33], which ensures a uniform sampling of the sphere [34]. A spherical design of order $t = 21$, consisting in a total of $D = 240$ directions, was used.

As stated in the introduction, two different sets of virtual microphones were used in this work: 3DVMS and Ambisonics. In the first case, the target function is a set of unidirectional virtual microphones having fourth order cardioid directivity without any side or rear lobes, defined as follow:

$$\boldsymbol{A}_{v,d}(\vartheta, \varphi) = [0.5 + 0.5 \cdot \cos(\vartheta) \cdot \cos(\varphi)]^4 \quad (2)$$

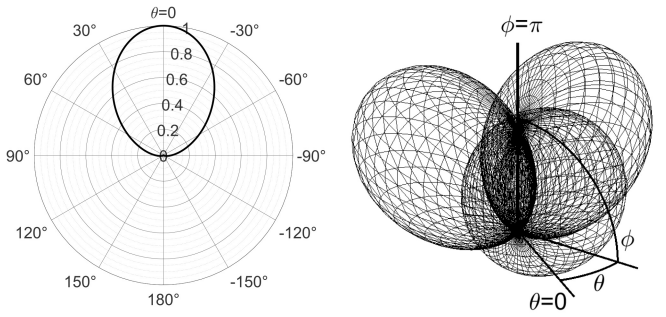

Fig. 1. Horizontal polar pattern of a fourth order cardioid (left). 3-dimensional plot of the 3DVMS virtual microphones aiming at directions ($\vartheta = 0°;\ \varphi = 45°$); ($\vartheta = 120°;\ \varphi = 45°$); ($\vartheta = -120°;\ \varphi = 45°$) (right).

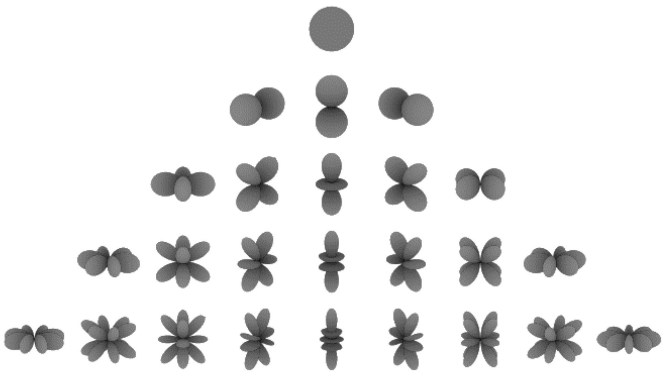

Fig. 2. Ideal three-dimensional directivity of the first $n = 25$ Spherical Harmonics, up to $o = 4$.

where ϑ is the azimuth, in the range $[-\pi;\pi]$ and φ is the elevation, in the range $[-\pi/2;\pi/2]$. For this work, the 3DVMS target function was defined as a set of three virtual microphones aiming at directions: ($\vartheta = 0°;\ \varphi = 45°$); ($\vartheta = 120°;\ \varphi = 45°$); ($\vartheta = -120°;\ \varphi = 45°$). A polar pattern of a fourth order cardioid can be seen in Fig. 1 (left) and a 3-dimensional plot of the 3DVMS virtual microphones can be seen in Fig. 1 (right).

In case of Ambisonics, the directivity of the virtual microphones is defined by more complex mathematical functions named Spherical Harmonics (SH) [35], [36]. SH are organized in orders, and the relation between the Ambisonics order o and the number n of SHs is:

$$n = (1 + o)^2 \quad (3)$$

An explicit formulation of the SH equations up to order 7 can be found in [37], for both spherical and Cartesian coordinates, while in Fig. 2 the first $n = 25$ SHs up to order $o = 4$ can be seen.

In principle, each virtual microphone is obtained by multiplying, in the frequency domain, the response of the array $\boldsymbol{C}_{m,d,k}$ with the beamforming filter $\boldsymbol{H}_{m,v,k}$, as follow:

$$\boldsymbol{V}_{d,k} = \boldsymbol{C}_{m,d,k} \cdot \boldsymbol{H}_{m,v,k} \quad (4)$$

In case of ideal synthesis of the beamforming filters $\boldsymbol{H}$, there is a perfect match between the obtained directivity $\boldsymbol{V}$ and the target function $\boldsymbol{A}$, resulting in $\boldsymbol{V} = \boldsymbol{A}$ at all frequencies and directions. The effective directivity obtained

with the beamforming can be evaluated by employing several metrics as a function of frequency: directivity factor Q and half-power beamwidth BW with 3DVMS format and Spatial Correlation (SC) and Level Difference (LD) with Ambisonics format. The mathematical definitions of these parameters can be found in [38], [39]. Regarding Q factor, the higher its value, the higher the virtual microphone directivity. The parameter BW denotes the beam opening angle between the central axis and the direction at -3 dB, therefore lower is better.

SC denotes the similarity between the directivity of each SH and the corresponding ideal one, disregarding the level matching. It varies in the range [0; 1], and it is equal to 1 in the ideal case. LD denotes the difference in dB between the obtained SH function and the ideal one. It has values in the range $(-\infty;\ 0]$, and it assumes a value of 0 dB in case of perfect level matching. By imposing a minimum acceptability threshold to SC and LD parameters, it is possible to define a frequency range of validity for each Ambisonics order. SC usually defines the maximum frequency, while LD the minimum one. The following limits will be used in this work:

$$SC > 0.95 \tag{5}$$

$$LD > -1\ dB, \tag{6}$$

III. Numerical and Experimental Methods

The section describes the methods employed in this work for numerical simulations and experimental measurements. Whatever the approach, the FIR filter matrix $\boldsymbol{H}$ is obtained by solving (1), which always requires to know the matrix $\boldsymbol{C}$. The purpose of simulations and measurements is precisely to provide the matrix $\boldsymbol{C}$.

Numerical simulations were operated with Finite Elements Method (FEM), in frequency domain. The stimulus is given by an ideal point source radiating spherical waves at 1 m distance (near field), for each direction d of the previously described spherical design grid. The solution was calculated in a spherical air domain having a radius of 0.15 m, while the arrays were modeled as rigid bodies. A 3-dimensional modelling was used, discretized with a tetrahedral mesh, featuring six elements per wavelength, as suggested in [40]. Simulations were solved with a frequency resolution given by:

$$\Delta_f = f_s / n_{FFT} \tag{7}$$

where $f_s = 48\ kHz$ is the standard sampling frequency in current audio applications [41] and $n_{FFT} = 2^{13}$ is the number of frequency bins chosen for the Fast Fourier Transform (FFT). Therefore, a frequency resolution of $\Delta_f \approx 5.86\ Hz$ was employed. For each direction d, the complex pressure is evaluated at the M points corresponding to the position of the capsules, thus providing the matrix $\boldsymbol{C}$ required in (1). Simulations were solved in the frequency range 180 Hz – 5 kHz, where the minimum frequency coincides with the minimum valid frequency of the measurement (see later in this section) and the maximum frequency is determined by a practical computational limit.

Experimental measurements were carried out in the acoustics laboratory of the Department of Engineering and

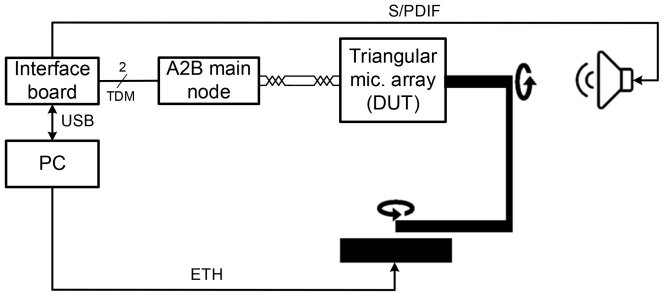

Fig. 3. Experimental measurement setup.

Architecture at the University of Parma (Parma, Italy). Each meta-array was measured with a two-axis turntable and a "coincident" active studio monitor (featuring three loudspeakers mounted on the same axis, which is the main sound propagation direction) positioned on-axis at 1 m, thus maintaining the same condition of numerical simulations. The schematic of the experimental setup is shown in Fig. 3. More details about the measurement method can be found in [42].

The turntable is controlled by the PC via ethernet link. After each measurement, the PC sends to the turntable the new measurement direction, and the two motors of the turntable rotate the device under test (DUT) consequently. The same directions employed for simulations were used. An A^{2}B interface board (see Section IV) connected to the PC via USB was employed for recording the M pressure signals from the DUT while delivering the test signal to the studio monitor through S/PDIF. Hence, the measurement system is full-digital, with synchronized input-output. The test signal was an Exponential Sine Sweep (ESS) [43], pre-equalized for flattening the spectrum of the sound source in the range 50 Hz – 18 kHz within $\pm$0.5 dB. Linear Impulse Responses (IRs) were calculated by convolving the recorded signals with the inverse filter associated to the test signal, namely the inverse sweep [43]. The virtual anechoic room technique [44] was adopted to exclude the environment reverberation; therefore, the linear part of the IRs was cut before the first reflection, thus limiting the lower frequency to 180 Hz, as:

$$f_{min} = f_s / N_{samples} \tag{8}$$

where $f_s = 48\ kHz$ is the sampling frequency and $N = 260$ is the number of samples between the peak of the IR and the first reflection. Eventually, the $\boldsymbol{C}$ matrix is obtained by applying the FFT to the trimmed IRs.

IV. Multiple A^{2}B Networks Architecture

An A^{2}B network is composed by a main node and up to 10 subordinate nodes connected in a daisy-chain topology. The maximum cable length between the main node and the last subordinate node is 40 m, whilst the maximum distance between two adjacent nodes is limited to 15 m. All nodes are connected using an Unshielded Twisted Pair (UTP) cable, and subordinate nodes can be power supplied through the same cable (up to 2.7 W), keeping the wiring as simple as possible. Access to the bus and data flow are managed by dedicated transceivers, removing the need of additional devices that would increase the system cost and design complexity.

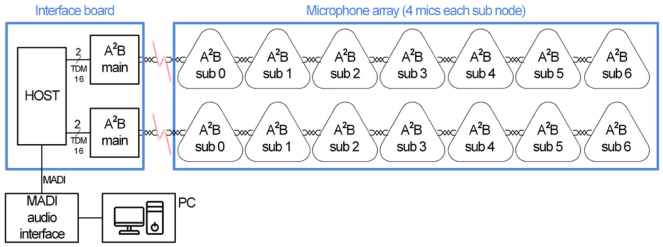

Fig. 4. Architecture of the proposed system.

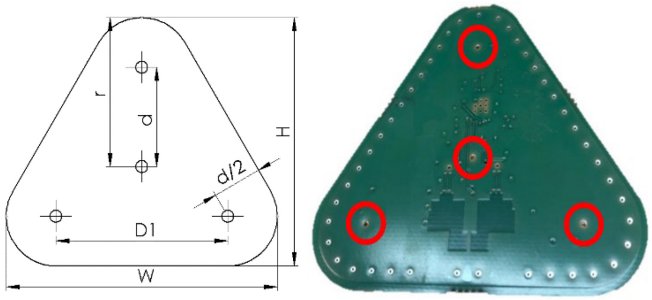

Fig. 5. Triangular unit drawing (left) and prototype of triangular unit (right).

TABLE I
MAIN DIMENSIONS OF THE TRIANGULAR ARRAY UNIT

d	25.0 mm
r	37.5 mm
D1	46.0 mm
H	62.5 mm
W	68.3 mm

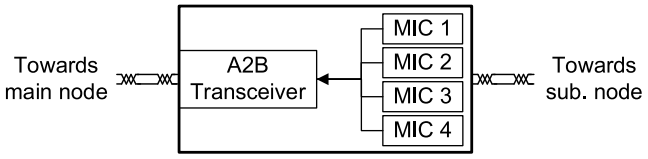

Fig. 6. Block diagram of the triangular microphone array unit.

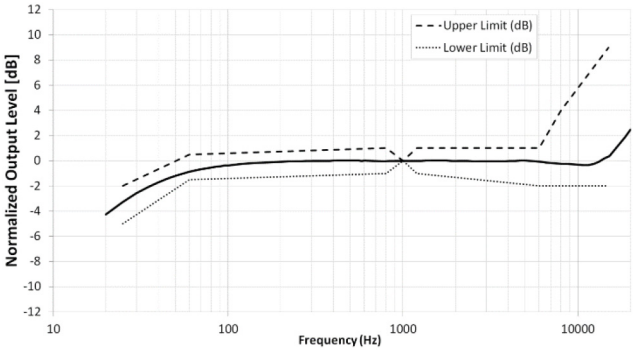

Fig. 7. Free field frequency response of the MEMS capsule, normalized to 1 kHz sensitivity.

A single network has a total bandwidth of 50 Mbit/s, which allows transporting up to 32 audio channels with a sample rate of 44.1 kHz or 48 kHz and sample width of 16 bits, 24 bits, or 32 bits, using a TDM format. In addition to audio channels, Inter Integrated Circuit (I^2C) commands and General Purpose Input/Output (GPIO) signals can be transmitted to other nodes for diagnostic or other purpose. More details about A^2B bus can be found in [45].

A multi A^2B network can be generated by connecting to four A^2B buses, in which all signals are synchronized. The architecture of the proposed system (Fig. 4) is composed by an interface board and two A^2B buses. The interface board includes two main nodes, one for each bus, and a host block. The host block provides the same clock signal to both main nodes, which distribute it to the entire network, thus synchronizing all the subordinate nodes. The layout of the latter is the same as in [26]. The host block also converts TDM signals from A^2B transceivers into AES67 and MADI protocols [46], commonly used for digital audio distribution, without the need of additional devices that would increase the system cost and design complexity. Only a simple 8-bit Reduced Instruction Set Computer (RISC) microcontroller is required on the host block to configure the network at start-up.

Due to bandwidth limitation, a single A^2B network is limited to 28 input channels. Since each triangular module is equipped with four MEMS microphones, a single main node drives up to seven subordinate nodes, as in the planar meta-array configuration. The entire network is instead limited to 14 modules (56 channels), as in the truncated octahedron meta-array configuration.

V. PRACTICAL REALIZATION OF META-ARRAYS

In this section, the practical aspects related to the design, simulation, construction, and measurement of meta-arrays are discussed. The constitutive unit of the meta-arrays is, by definition, an array itself. The unit employed in this work has a triangular geometry and four capsules, one located in the center and three at the vertexes of a triangle. A top view of the drawing (left) and a picture of the triangular unit (right) can be seen in Fig. 5, while the main dimensions are summarized in Table I. According to [47], the maximum beamforming frequency is physically constrained by the ratio between the distances of the capsules and the wavelength, as:

$$d = \frac{1}{4} \cdot \frac{c}{f} \tag{9}$$

where $c = 343$ m/s is the speed of sound, f is frequency, and d is the minimum distance among the capsules. The triangular array was designed to fit voice applications, such as teleconferencing, speech recognition, speakerphone, or ANC systems, which works typically in the frequency range 300 Hz – 3.4 kHz, as referenced in [48], [49]. The resulting minimum distance for $f = 3400$ Hz is about 25 mm. More about the sizing of the triangular unit can be found in [26].

A block diagram of the triangular array is shown in Fig. 6. One can note each unit can be connected in daisy chain with an unshielded twisted pair exploiting the A^2B bus, to form a larger A^2B network by positioning many array boards side-by-side, in two or three dimensions.

The digital MEMS capsule is characterized by an Acoustic Overload Point (AOP) of 130 dB (SPL), SNR of 69 dB(A), dynamic range of 105 dB and an operating voltage range 1.62 – 3.6 V. The free field normalized frequency response is shown in Fig. 7. Electrical parameters and acoustic performance are provided by the manufacturer in [50].

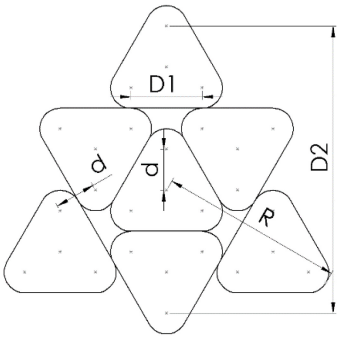

TABLE I
MAIN DIMENSIONS OF PLANAR META-ARRAY

d	25.0 mm
D1	46.0 mm
D2	175.0 mm
R	112.5 mm

Fig. 8. Top view of the planar meta-array drawing (left). Main dimensions of the planar meta-array (right).

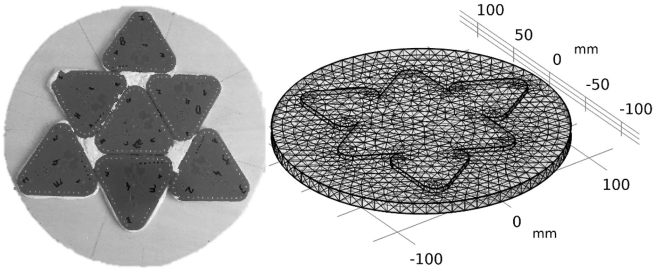

Fig. 9. A prototype of the planar meta-array with 7 triangular arrays and $M = 28$ capsules (left). FEM model of the planar meta-array (right).

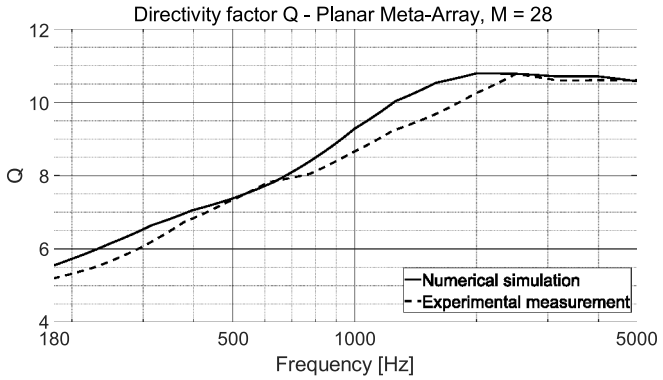

Fig. 10. Directivity factor Q for planar meta-array, numerical model (solid line) and experimental measurement (dash line). NB: higher is better.

At first, a planar meta-array featuring seven triangular arrays and M = 28 capsules was designed. A top view of the drawing is shown in Fig. 8 (left), Table I summarizes the main dimensions (right). The triangular unit was designed so that the minimum distance between the capsules of two contiguous triangular units is the same distance between the capsules in a unit (i.e., d = 25.0 mm). It can be seen the maximum distance between the capsules increased from D1 = 46.0 mm for the single triangular array to D2 = 175.0 mm for the meta-array.

The prototype was built by mounting the triangular units over a circular wooden baffle (Fig. 9 left) having a radius of 150 mm. The numerical FEM model was provided with the rigid baffle too (Fig. 9 right).

A planar array is not optimal for encoding Ambisonics since all capsules lay on the horizontal plane, thus the spatial information on the vertical direction is limited. Therefore, 3DVMS virtual microphones were used. In Fig. 10 and Fig. 11, the parameters Q and BW are shown, respectively.

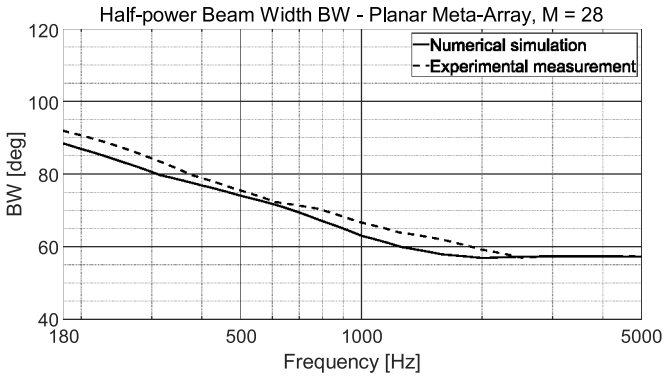

Fig. 11. Half-power Beam Width BW for planar meta-array, numerical model (solid line) and experimental measurement (dash line). NB: lower is better.

The charts are within the minimum valid frequency given by the measurement, that is 180 Hz, and the maximum frequency of the simulation, that is 5 kHz. One can note an excellent agreement between numerical simulation and experimental measurement, despite the numerical simulation, which is perfectly ideal, shows a slightly better performance toward lower frequencies. In the frequency range where the metrics have a flat trend, above 2 kHz for the numerical simulation and 2.5 kHz for the experimental measurement, the target function $\boldsymbol{A}$ is matched, with Q = 10.7 and BW = 57°. At lower frequencies, a progressive decrease of the directivity with respect to the target function is observed, down to a minimum of Q = 5.5; BW = 88° for the numerical simulation and Q = 5; BW = 93° for the experimental measurement, at 180 Hz.

Then, a 3-dimensional meta-array was designed, with the aim of maximizing the spatial performance, given the limit of 14 triangular arrays that can be connected through two A^2B networks. Two geometries were compared before building a prototype, to optimize the triangular arrays positioning in space: a regular dodecahedron [51] and a truncated octahedron [52]. They both show a high value of the Isoperimetric Quotient (IQ), a parameter for assessing the approximation of a sphere by a polyhedron, defined as:

$$IQ = 36\pi\left(\frac{V^2}{S^3}\right) \tag{10}$$

where V is the volume and S is the surface of the polyhedron. By definition the sphere has an IQ = 1, and this is the maximum possible value of the parameter [53]. The values of IQ for the regular dodecahedron and the truncated octahedron are as follows:

$$IQ_{reg.dodecahedron} = 0.7547 \tag{11}$$

$$IQ_{trunc.octahedron} = 0.7770 \tag{12}$$

The dodecahedron is made of 12 faces, and it is widely employed for both acoustic measurements [54] and microphone arrays [55]. The truncated octahedron is formed by 14 faces instead of 12, hence it allows for $M = 56$ in place of $M = 48$. Minimum and maximum distances between the capsules for the two meta-arrays are shown in Table II.

Being the arrays 3-dimensional, Ambisonics format was used as target function. It was encoded up to order $o = 5$,

TABLE II
MIN. AND MAX. DISTANCES BETWEEN MICROPHONE CAPSULES

Array type	Minimum Distance	Maximum distance
Regular dodecahedron	25.0 mm	142.6 mm
Truncated octahedron	25.0 mm	146.0 mm

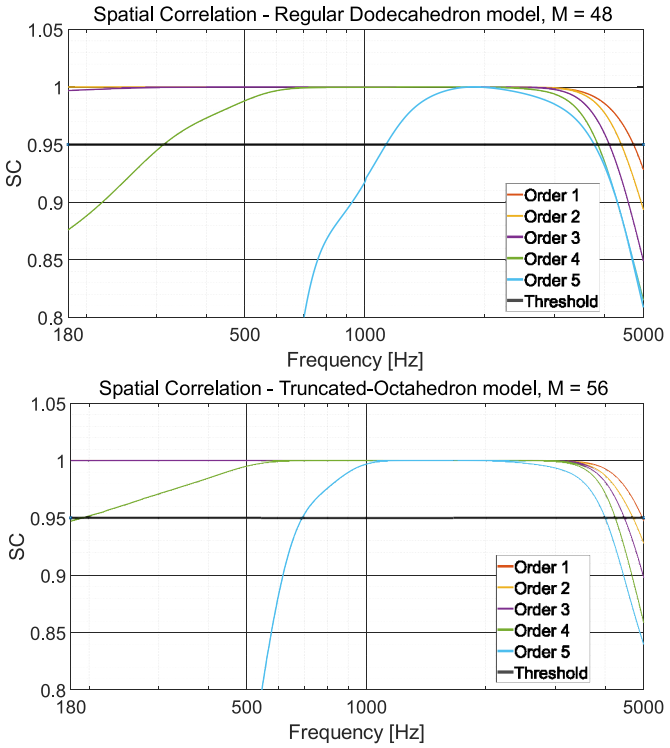

Fig. 12. SC for the regular dodecahedron meta-array (above) and for the truncated octahedron meta-array (below). NB: closer to 1 is better.

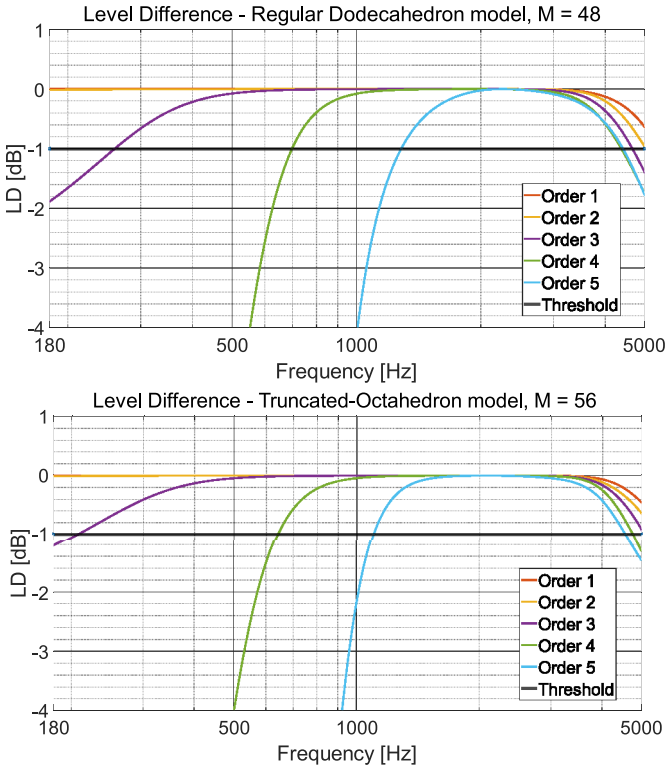

Fig. 13. LD for the regular dodecahedron meta-array (above) and for the truncated octahedron meta-array (below). NB: closer to 0 dB is better.

TABLE III
MIN. AND MAX. FREQUENCIES FOR AMBISONICS ORDERS O = 1 ÷ 5

O	Regular Dodecahedron M = 48 Numerical model	Truncated-Octahedron M = 56 Numerical model	Regular Dodecahedron M = 48 Numerical model	Truncated-Octahedron M = 56 Numerical model
	Minimum frequency		Maximum frequency	
1	< 180 Hz	< 180 Hz	4.7 kHz	4.9 kHz
2	< 180 Hz	< 180 Hz	4.4 kHz	4.7 kHz
3	250 Hz	200 Hz	4.1 kHz	4.5 kHz
4	700 Hz	650 Hz	3.8 kHz	4.2 kHz
5	1.3 kHz	1.1 kHz	3.7 kHz	4.0 kHz

thus resulting in $n = 36$ spherical harmonics, as given by (3). The charts of SC and LD are shown in Fig. 12 and Fig. 13 for the two geometries, within the frequency range 180 Hz – 5 kHz.

By applying the thresholds (5) and (6), the minimum and maximum frequencies for each Ambisonics order were defined (Table III). It can be seen the truncated octahedron meta-array has better performance: the frequency ranges are wider for all orders. Therefore, this geometry was chosen to build a prototype (Fig. 14).

The truncated octahedron meta-array was compared with the Eigenmike-32, a spherical microphone array of 42 mm radius and featuring 32 analogue capsules. Both arrays have been measured with the two-axis turntable (see Section III). *SC* and *LD* metrics are shown in Fig. 15 and Fig. 16 for the two arrays, within the frequency range 180 Hz – 5 kHz. By applying the two thresholds (5) and (6), the minimum and maximum frequencies for each Ambisonics order were defined (Table IV).

One can note the truncated octahedron performed better than the Eigenmike-32 at low frequencies. In addition, having 56 capsules instead of 32, it allows encoding Ambisonics format up to fifth order, while the Eigenmike-32 cannot overcome the fourth. Hence, it provides a better spatial accuracy in the entire vocal band (300 Hz – 3.4 kHz), which is of main interest for telecommunication and ANC applications. It must be said the Eigenmike-32 provides beamforming up to 5 kHz and beyond, thus significantly exceeding the limit of the truncated octahedron. This is due to the smaller distance between the capsules, which in the case of the Eigenmike-32 is 12 mm, i.e., less than half that of the truncated octahedron. Therefore, it remains superior in music applications. Eventually, by comparing Table III and Table IV, an excellent agreement between numerical simulation and experimental measurement of the truncated octahedron is observed.

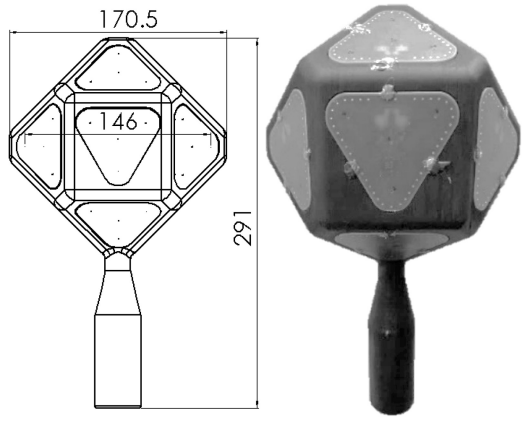

Fig. 14. Side view of the truncated octahedron meta-array drawing (left). The prototype of the truncated octahedron meta-array (right).

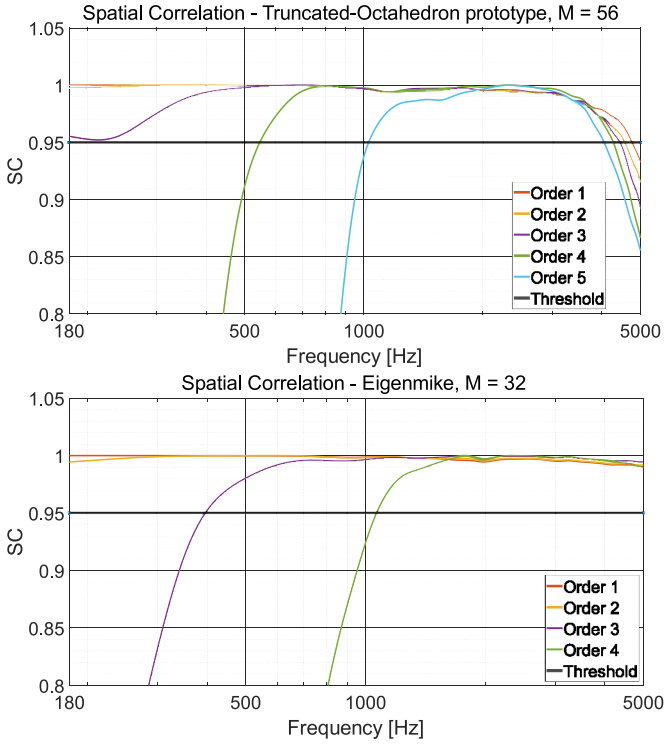

Fig. 15. SC for the truncated octahedron meta-array (above) and for the Eigenmike-32 (below). NB: closer to 1 is better.

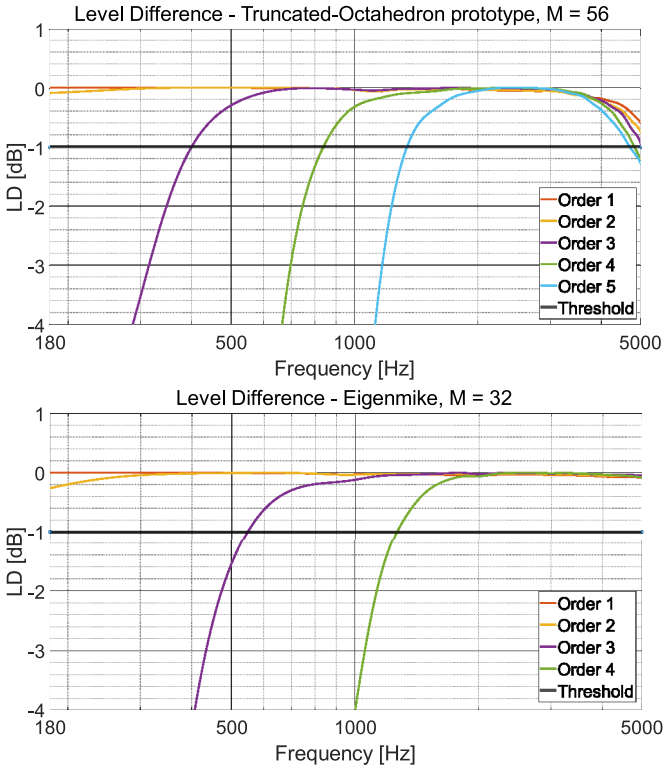

Fig. 16. LD for the truncated octahedron meta-array (above) and for the Eigenmike-32 (below). NB: closer to 0 dB is better.

TABLE IV
MIN. AND MAX. FREQUENCIES FOR AMBISONICS ORDERS O = 1 ÷ 5

O	Truncated-Octahedron M = 56 Measurement	Eigenmike M = 32 Measurement	Truncated-Octahedron M = 56 Measurement	Eigenmike M = 32 Measurement
	Minimum frequency		Maximum frequency	
1	< 180 Hz	< 180 Hz	4.8 kHz	> 5 kHz
2	< 180 Hz	< 180 Hz	4.6 kHz	> 5 kHz
3	400 Hz	550 Hz	4.5 kHz	> 5 kHz
4	850 Hz	1.3 kHz	4.2 kHz	> 5 kHz
5	1.3 kHz	-	4.0 kHz	-

VI. CONCLUSION

Full-digital, multi-network, A^{2}B based meta-arrays of microphones have been proposed. A^{2}B bus revealed optimal for building meta-arrays, thanks to daisy-chain connection and synchronized acquisition. By exploiting the presented multi-network architecture, the 32-channel limitation of A^{2}B bus has been exceeded. In addition, it allows avoiding bulky wiring between capsules, arrays, and acquisition boards, since it communicates by means of UTP cables. At the same time, it offers optimal characteristics for consumer electronics applications such as robust data transmission (up to 40 m), cheap cabling, and expandability.

By connecting in daisy-chain several 4-microphones triangular array units, two meta-arrays have been designed, simulated, prototyped, and measured. A planar one, based on a single A^{2}B network and featuring seven triangular units and 28 capsules, and a 3-dimensional one, which makes use of two A^{2}B networks and includes 14 triangular units and 56 capsules.

The planar meta-array was analyzed with the 3DVMS approach: three virtual cardioid microphones of fourth order were synthesized. Directivity factor Q and Half-Power Beam Width BW parameters were employed for evaluating the spatial performance of the virtual microphones, showing an optimal agreement between numerical and experimental approach.

The 3-dimensional meta-array was analyzed by encoding Ambisonics format up to order five. Spatial Correlation and Level Difference parameters were employed for evaluating the quality of the Spherical Harmonics. At first, a regular dodecahedron (12 triangular units, 48 capsules) and a truncated

octahedron (14 triangular units, 56 capsules) geometries were compared through numerical simulations for assessing the optimal configuration. Since the truncated octahedron provided better performance, it was chosen for building a prototype. It was measured and compared to the numerical model, showing an excellent agreement between the two approaches.

Eventually, the truncated octahedron meta-array prototype was compared with the Eigenmike-32, demonstrating better performance in the vocal band, which is of main interest for telecommunication and ANC applications. At the same time, making use of MEMS capsules instead of analog ones, it allows reducing considerably the cost of electronics.

REFERENCES

[1] M. Fukui, K. Kobayashi, S. Shimauchi, Y. Hioka, and H. Ohmuro, "Hands-free audio conferencing unit with low-complexity dereverberation," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2015, pp. 128–129, doi: 10.1109/ICCE.2015.7066349.

[2] Y. Hioka, M. Okamoto, K. Kobayashi, Y. Haneda, and A. Kataoka, "A display-mounted high-quality stereo microphone array for high-definition videophone system," in *Dig. Tech. Paper IEEE Int. Conf. Consum. Electron.*, Feb. 2008, pp. 1–2, doi: 10.1109/ICCE.2008.4587858.

[3] J. Hong, "Stereophonic acoustic echo suppression for speech interfaces for intelligent TV applications," *IEEE Trans. Consum. Electron.*, vol. 64, no. 2, pp. 153–161, May 2018, doi: 10.1109/TCE.2018.2843293.

[4] J. Hong, S. Jeong, and M. Hahn, "Wiener filter-based echo suppression and beamforming for intelligent TV interface," *IEEE Trans. Consum. Electron.*, vol. 59, no. 4, pp. 825–832, Nov. 2013, doi: 10.1109/TCE.2013.6689695.

[5] T. Kawase, M. Okamoto, T. Fukutomi, and Y. Takahashi, "Speech enhancement parameter adjustment to maximize accuracy of automatic speech recognition," *IEEE Trans. Consum. Electron.*, vol. 66, no. 2, pp. 125–133, May 2020, doi: 10.1109/TCE.2020.2986003.

[6] C.-Y. Chang, A. Siswanto, C.-Y. Ho, T.-K. Yeh, Y.-R. Chen, and S. M. Kuo, "Listening in a noisy environment: Integration of active noise control in audio products," *IEEE Consum. Electron. Mag.*, vol. 5, no. 4, pp. 34–43, Oct. 2016, doi: 10.1109/MCE.2016.2590159.

[7] V. Patel, J. Cheer, and S. Fontana, "Design and implementation of an active noise control headphone with directional hear-through capability," *IEEE Trans. Consum. Electron.*, vol. 66, no. 1, pp. 32–40, Feb. 2020, doi: 10.1109/TCE.2019.2956634.

[8] S. M. Kim, C. J. Chun, and H. K. Kim, "Multi-channel audio recording based on superdirective beamforming for portable multimedia recording devices," *IEEE Trans. Consum. Electron.*, vol. 60, no. 3, pp. 429–435, Aug. 2014, doi: 10.1109/TCE.2014.6937327.

[9] N. D. Gaubitch, J. Martinez, W. B. Kleijn, and R. Heusdens, "On near-field beamforming with smartphone-based ad-hoc microphone arrays," in *Proc. 14th Int. Workshop Acoust. Signal Enhanc. (IWAENC)*, Feb. 2014, pp. 94–98, doi: 10.1109/IWAENC.2014.6953345.

[10] D. Pinardi, L. Ebri, C. Belicchi, A. Farina, and M. Binelli, "Direction specific analysis of psychoacoustics parameters inside car cockpit: A novel tool for NVH and sound quality," in *SAE Tech. Papers*, 2020, pp. 1–8, doi: 10.4271/2020-01-1547.

[11] D. Pinardi, "A human head shaped array of microphones and cameras for automotive applications," in *Proc. Immersive 3D Audio Architect. Automot. (I3DA)*, Sep. 2021, pp. 1–8, doi: 10.1109/I3DA48870.2021.9610879.

[12] D. Pinardi, A. Farina, and J.-S. Park, "Low frequency simulations for ambisonics auralization of a car sound system," in *Proc. Immersive 3D Audio Architect. Automot. (I3DA)*, Sep. 2021, pp. 1–10, doi: 10.1109/I3DA48870.2021.9610959.

[13] A. Toscani et al., "Low-cost structural health monitoring system for smart buildings," in *Proc. 2nd Int. Conf. Sustain Mobility Appl. Renew. Technol. (SMART)*, Nov. 2022, pp. 1–7, doi: 10.1109/SMART55236.2022.9990379.

[14] A. D. Firoozabadi et al., "Three-dimensional sound source localization by distributed microphone arrays," in *Proc. 29th Eur. Signal Process. Conf. (EUSIPCO)*, Feb. 2021, pp. 196–200, doi: 10.23919/EUSIPCO54536.2021.9616326.

[15] T. Damarla, "Detection of gunshots using microphone array mounted on a moving platform," in *Proc. IEEE SENSORS*, Feb. 2015, pp. 1–4, doi: 10.1109/ICSENS.2015.7370532.

[16] R. Kerstens, D. Laurijssen, and J. Steckel, "Low-cost one-bit MEMS microphone arrays for in-air acoustic imaging using FPGA's," in *Proc. IEEE Sensors*, Feb. 2017, pp. 1–3, doi: 10.1109/ICSENS.2017.8234087.

[17] M. Turqueti, J. Saniie, and E. Oruklu, "Scalable acoustic imaging platform using MEMS array," in *Proc. IEEE Int. Conf. Electr. Eng. Inform. Commun. Technol. (EIT)*, 2010, pp. 1–4, doi: 10.1109/EIT.2010.5612131.

[18] H. A. Sanchez-Hevia, R. Gil-Pita, and M. Rosa-Zurera, "FPGA-based real-time acoustic camera using PDM MEMS microphones with a custom demodulation filter," in *Proc. IEEE 8th Sens. Array Multichannel Signal Process. Workshop (SAM)*, Feb. 2014, pp. 181–184, doi: 10.1109/SAM.2014.6882370.

[19] R. González, J. Pearce, and T. Lokki, "Modular design for spherical microphone arrays," in *Proc. AES Int. Conf.*, 2018, pp. 1–7. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=19701

[20] I. Andras, P. Dolinsky, L. Michaeli, and J. Saliga, "Beamforming with small diameter microphone array," in *Proc. 28th Int. Conf. Radioelektronika*, Feb. 2018, pp. 1–5, doi: 10.1109/RADIOELEK.2018.8376368.

[21] Z. I. Skordilis, A. Tsiami, P. Maragos, G. Potamianos, L. Spelgatti, and R. Sannino, "Multichannel speech enhancement using MEMS microphones," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Feb. 2015, pp. 2729–2733, doi: 10.1109/ICASSP.2015.7178467.

[22] R. Schultz-Amling, F. Kuech, M. Kallinger, G. del Galdo, J. Ahonen, and V. Pulkki, "Planar microphone array processing for the analysis and reproduction of spatial audio using directional audio coding," in *Proc. 124th AES Conv.*, 2008, pp. 1–10. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=14505

[23] J. Lopez-Ballester, M. Cobos, J. J. Perez-Solano, G. Moreno, and J. Segura, "General purpose modular microphone array for spatial audio acquisition," in *Proc. 138th AES Conv.*, 2015, pp. 1–6. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=17674

[24] N. v Ilawe, M. B. Oviedo, and B. M. Wong, "Real-time quantum dynamics of long-range electronic excitation transfer in plasmonic nanoantennas," *J. Chem. Theory Comput.*, vol. 13, no. 8, pp. 3442–3454, Feb. 2017, doi: 10.1021/acs.jctc.7b00423.

[25] E. T. Enikov and J. G. Boyd, "A thermodynamic field theory for anodic bonding of micro electro-mechanical systems (MEMS)," *Int. J. Eng. Sci.*, vol. 38, no. 2, pp. 135–158, Feb. 2000, doi: 10.1016/S0020-7225(99)00027-0.

[26] D. Pinardi et al., "An innovative architecture of full-digital microphone arrays over A^2B network for consumer electronics," *IEEE Trans. Consum. Electron.*, vol. 68, no. 3, pp. 200–208, Aug. 2022, doi: 10.1109/TCE.2022.3187453.

[27] Analog Devices. "A2B Audio Bus: An easier, simpler solution for audio design." Accessed: Feb. 2023. [Online]. Available: https://www.analog.com/en/applications/technology/a2b-audio-bus.html

[28] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Feb. 2002, pp. 1781–1784, doi: 10.1109/ICASSP.2002.5744968.

[29] 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," in *Proc. 40th Int. AES Conf. Spatial Audio Sense Sound of Space*, Feb. 2010. pp. 1–10, [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=15577

[30] M. Gerzon, "Ambisonics. Part two: Studio techniques," *Studio Sound*, vol. 17, no. 8, pp. 24–30, Aug. 1975. [Online]. Available: https://www.michaelgerzonphotos.org.uk/articles/Ambisoni%202.pdf

[31] O. Kirkeby, F. Orduna, P. A. Nelson, and H. Hameda, "Inverse filtering in sound reproduction," *Meas. Control*, vol. 26, no. 9, pp. 261–266, Feb. 1993, doi: 10.1177/002029409302600902.

[32] H. Tokuno, O. Kirkeby, P. A. Nelson, and H. Hamada, "Inverse filter of sound reproduction systems using regularization," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E80-A, no. 5, pp. 809–820, 1997. [Online]. Available: https://www.melaudia.net/zdoc/kirkebyInverseFilter.PDF

[33] R. H. Hardin and N. J. A. Sloane, "McLaren's improved snub cube and other new spherical designs in three dimensions," *Discr. Comput. Geom.*, vol. 15, no. 4, pp. 429–441, Apr. 1996, doi: 10.1007/BF02711518.

[34] D. Pinardi, "Spherical t-design for characterizing the spatial response of microphone arrays," in *Proc. Immersive and 3D Audio Arch. Automot. (I3DA)*, Sep. 2021, pp. 1–8, doi: 10.1109/I3DA48870.2021.9610850.
[35] 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," in *Proc. 144th AES Conv.*, Feb. 2018, pp. 294–303. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=19456
[36] L. McCormack et al., "Applications of spatially localized active-intensity vectors for sound-field visualization," *J. Audio Eng. Soc.*, vol. 67, no. 11, pp. 840–854, 2019, doi: 10.17743/JAES.2019.0041.
[37] A. Farina. "Explicit ambix formulas for high order ambisonics." Aug. 2017. [Online]. Available: http://pcfarina.eng.unipr.it/Aurora/HOA_explicit_formulas.htm
[38] S. Bertet, J. Daniel, and S. Moreau, "3D sound field recording with higher order ambisonics - objective measurements and validation of spherical microphone," in *Proc. 120th AES Conv.*, Feb. 2006. pp. 1–24, [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=13661
[39] D. Pinardi and A. Farina, "Metrics for evaluating the spatial accuracy of microphone arrays," in *Proc. Immersive 3D Audio Arch. Autom. I3DA*, Sep. 2021, pp. 1–9, doi: 10.1109/I3DA48870.2021.9610887.
[40] COMSOL. "Acoustics module user's guide." Accessed: Feb. 2023. [Online]. Available: https://doc.comsol.com/5.5/doc/com.comsol.help.aco/AcousticsModuleUsersGuide.pdf
[41] Audio Engineering Society. "AES5–2018: AES recommended practice for professional digital audio—Preferred sampling frequencies for applications employing pulse-code modulation (revision of AES5–2003)." 2018. [Online]. Available: https://www.aes.org/tmpFiles/aessc/20230222/aes05–2018-i.pdf
[42] A. Farina, S. Campanini, L. Chiesi, A. Amendola, and L. Ebri, "Spatial sound recording with dense microphone arrays," in *Proc. 55th AES Int. Conf.*, 2014, pp. 1–8. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=17362
[43] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Proc. 108th AES Conv.*, 2000, pp. 1–24. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=10211
[44] J. Svacina, J. Drinovsky, and R. Videnka, "Virtual anechoic room an useful tool for EMI pre-compliance testing," in *Proc. 17th Int. Conf. Radioelektronika*, Feb. 2007, pp. 1–4, doi: 10.1109/RADIOELEK.2007.371702.
[45] N. Rocchi, A. Toscani, G. Chiorboli, D. Pinardi, M. Binelli, and A. Farina, "Transducer arrays over A^2B networks in industrial and automotive applications: Clock propagation measurements," *IEEE Access*, vol. 9, pp. 118232–118241, 2021, doi: 10.1109/ACCESS.2021.3106710.
[46] N. Rocchi et al., "A modular, low latency, A^2B-based architecture for distributed multichannel full-digital audio systems," in *Proc. Immersive 3D Audio Arch. Automot. (I3DA)*, Sep. 2021, pp. 1–8, doi: 10.1109/I3DA48870.2021.9610947.
[47] J. Dmochowski, J. Benesty, and S. Affes, "On spatial aliasing in microphone arrays," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1383–1395, Apr. 2009, doi: 10.1109/TSP.2008.2010596.
[48] J. Ahonen, V. Pulkki, and T. Lokki, "Teleconference application and B-format microphone array for directional audio coding," in *Proc. AES 30th Int. Conf. Intell. Audio Environ.*, Feb. 2007, pp. 1–10. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=13927
[49] Q. Zou, X. Zou, M. Zhang, and Z. Lin, "A robust speech detection algorithm in a microphone array teleconferencing system," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 5, 2001, pp. 3025–3028, doi: 10.1109/ICASSP.2001.940295.
[50] Infineon Technologies. "Im69D130." 2017. Accessed: Feb. 2023. [Online]. Available: https://www.infineon.com/dgdl/Infineon-IM69D130-DataSheet-v01_00-EN.pdf?fileId=5546d462602a9dc801607a0e46511a2e
[51] E. W. Weisstein. "Regular dodecahedron, MathWorld—A wolfram Web resource." Accessed: Feb. 2023. [Online]. Available: https://mathworld.wolfram.com/RegularDodecahedron.html
[52] E. W. Weisstein. "Truncated octahedron, MathWorld—A wolfram Web resource." Accessed: Feb. 2023. [Online]. Available: https://mathworld.wolfram.com/TruncatedOctahedron.html
[53] A. Lengyel, Z. Gáspár, and T. Tarnai, "The roundest polyhedra with symmetry constraints," *Symmetry*, vol. 9, no. 3, p. 41, Mar. 2017, doi: 10.3390/sym9030041.
[54] F. Martellotta, "Optimizing stepwise rotation of dodecahedron sound source to improve the accuracy of room acoustic measures," *J. Acoust. Soc. Amer.*, vol. 134, no. 3, pp. 2037–2048, Sep. 2013, doi: 10.1121/1.4817879.
[55] B. N. Gover, J. G. Ryan, and M. R. Stinson, "Designing a spherical microphone array for the directional analysis of reflections and reverberation," in *Proc. 115th AES Conv.*, Oct. 2003, pp. 1–11. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=12447

D. Pinardi received the M.S. degree (cum laude) in mechanical engineering from the University of Parma, Italy, in July 2016, with a thesis on loudspeaker modeling and the Ph.D. degree in industrial engineering from the University of Parma in March 2020, with a thesis on the design of microphone, hydrophone, and camera arrays for spatial audio recording.

He has been a Research Assistant of Prof. Angelo Farina with the University of Parma since 2016, mainly specialized in spatial audio, design of transducer arrays, acoustics simulations and 3-D auralization, applied to automotive field, and underwater acoustics.

Andrea Toscani received the M.S. degree (cum laude) in electronic engineering and the Ph.D. degree in information technology from the University of Parma, Italy, in 2004 and 2008, respectively.

Since 2004, he has been working with the Department of Information Engineering (currently, the Department of Engineering and Architecture), University of Parma, where he is currently a Research Fellow. He is the author of two patents. His research activity is mainly focused on power electronics, high-performance electric drives, diagnostic techniques for industrial electric systems, and power converter for audio application.

M. Binelli received the B.S. and M.S. degrees in electronic engineering from the University of Parma, Italy, in 2003 and 2006, respectively, and the Ph.D. degree from the University of Parma, in 2010, with a thesis on analysis and equalization techniques in automotive acoustics.

Since 2010, he has been working as Research Fellow with the University of Parma. His research interests include equalization, microphones and loudspeakers arrays, spatial audio psychoacoustics, and active noise control.

L. Saccenti received the M.S. degree (cum laude) in electronic engineering from the University of Parma, Italy, in 2020, where he is currently pursuing the Ph.D. degree in automotive for intelligent mobility.

His research interests include acoustic simulation, 3-D auralization, spatial audio, and acoustical characterization of materials applied to automotive field.

A. Farina received the M.S. degree in civil engineering and the Ph.D. degree in technical physics from the University of Bologna in 1982 and 1987, respectively.

In May 2005, he became a Full Professor of Environmental Applied Physics with the University of Parma. He is the author of more than 300 scientific papers. His research activity includes several fields of applied acoustics, including noise and vibration, concert hall acoustics, simulation software, advanced measurement systems, microphone, and loudspeaker arrays. He was awarded with the Audio Engineering Society Fellowship for his pioneering work on electroacoustic measurements based on exponential sine sweeps.

L. Cattani received the Ph.D. degree in information technologies from the University of Parma, Italy, in 2016.

He is currently a Project Leader in the Research and Development Department, Ask Industries SpA. His research interests include advanced signal processing (audio and radio frequency) for automotive applications.