PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

On in situ beamforming in an automotive cabin using a planar loudspeaker array

Martin Bo MØLLER[1,2]; Martin OLSEN[3]

[1]Bang & Olufsen a/s, Denmark, mim@bang-olufsen.dk
[2]Aalborg University, Denmark, marbm@es.aau.dk
[3]Harman Lifestyle Audio, Denmark, martin.olsen@harman.com

Abstract
Compact loudspeaker arrays are often used in applications where control of the sound dispersion is required, e.g. in audio reproduction systems and aspects of sound field control. An example of the latter is sound zoning where the goal is to deliver individual audio content to one listener without disturbing another. Typical approaches are based on filter-and-sum beamforming applied to control the resulting directivity of the array. However, directivity is defined in free field while sound zones are generally of interest in listening spaces with reflective boundaries such as domestic rooms and automotive cabins. In order to address this discrepancy, in situ zonal control can be utilized. This relies on measured impulse responses from each array element to the spatial control points defining the zones. Such methods increase the solution specificity and can lead to non-causal control filters. In the present paper, methods based on controlling the flow of acoustic energy are applied to a planar loudspeaker array located on the dashboard inside a car cabin. A comparison study including directivity and in situ zone control is presented, and the solutions are evaluated in terms of the resulting sound zoning performance.
Keywords: Sound field control, Sound zones, Beamforming, Automotive audio

1 INTRODUCTION

Reproduction of high quality audio content in automotive listening spaces is associated with numerous challenges. Prominent examples are the spatially confined and complex acoustical environment as well as restricted loudspeaker placement options. Furthermore, a significant shift might occur in applications and user scenarios compared to today, in response to the expected prevalence of autonomous driving in the near future. Therefore, new functionalities have recently received much attention such as reconfigurable seating areas with preserved sound quality, or the ability to provide personal sound to individual occupants. Many aspects of these applications can be aided by introducing spatial sound field control solutions. The topic of the present paper concerns fundamental investigations of a 19 channel compact loudspeaker array, which was positioned on the dashboard inside an automotive cabin. Different sound field control strategies have been formulated and applied in an in situ comparison study. The objective was to investigate specific setup scenarios and to assess the resulting performance, related to controlling the sound content reproduced at each individual front seat. Performance assessment was based on the acoustic separation achieved between the front seats. This scenario can be attained by means of in situ zonal control, which relies on measured impulse responses from each array loudspeaker to the spatial control points, defining the spatially confined zones, as introduced in e.g. [1]. Other control methods have been proposed for in situ control, where vectorial metrics are included in the optimization cost functions, such as particle velocity [2], and sound intensity [3]. The resulting control filters inherently reflect the acoustical complexity of the listening space, which is a general characteristic of most in situ methods. The listening space information is captured in the measured impulse responses, and consequently the solution specificity is often high.

Another approach is to control the sound field by focusing the sound energy flow towards the target listening position, which effectively increases the acoustical separation. The beamforming control methods based on the assumption of free field sound radiation, implicitly assumes a high degree of resemblance between the setup scenario and the acoustical environment in which the compact array should be operating. The mismatch between a free field setup scenario and operation in reverberant environments has previously been considered in e.g. [4]. However, in automotive audio systems the positioning of control loudspeakers is restricted to confined regions along the interior boundaries, which often results in areas of highly complex geometrical and acoustic impedance features in close proximity to the array. This is likely to affect the control problem since some of the adjacent physical components should be considered a part of the array structure. Some aspects of these effects have been addressed in [5] in relation to cross-talk-cancellation, but not disclosed in details.

In the present paper, a comparative study is presented where a time-domain pressure matching control algorithm has been applied, and control filters were determined in different setup configurations. In addition, a simple delay-and-sum beamforming solution, based on free field analytical plane wave model has been included as reference [7]. The following specific setup methods for the control algorithm are considered:

- **Free field beamforming** where the control filters are based on impulse response measurements conducted under free field conditions in a grid of microphone positions, that captures the radiation characteristics. A subset of microphone positions is selected as the bright and dark points, respectively. In the case of beamforming, the bright points define the main sound energy flow direction whereas the dark points form the directions where ideally no sound radiation should occur. This definition is followed consistently throughout this paper when beamforming methods are considered.
- **In situ beamforming** was applied to include the effects of the reflecting interior boundaries of the reproduction scenario inside the car cabin. Here, the impulse responses are measured in a plane in front of the loudspeaker array. Constraining the sound field in points in this plane, according to the definition of bright and dark directions, resembles the control strategy proposed for the free field case. The hypothesis investigated here was that the in situ setup method potentially provides a beamforming solution due to a relatively short distance between the loudspeaker array and the control points.
- **In situ P-U beamforming** has been considered since controlling the sound field in a single plane introduces the risk of accurately controlling only that plane without focusing sound propagating towards the listener. To cope with this, a method including a vectorial control metric was defined, constraining both the pressure and particle velocity in order to steer the sound energy flow.
- **In situ zonal control** based on measurements of impulse responses between the array loudspeakers and microphones inside the two control regions at the front seats in the unoccupied cabin. A bright zone was defined where the sound pressure level should be higher compared to the dark zone at the adjacent seat.

In general, the setup measurements for the in situ methods were conducted inside a regular unoccupied car cabin. In addition, measurements series were performed with an anechoic termination for setting up special versions of the in situ beamforming and P-U. A head-and-torso-simulator (HATS) was utilized in order to create a more realistic performance assessment in the comparison of the different setup methods investigated in this paper.

2 THEORY

In this section, the applied sound field control algorithms are introduced along with the primary performance evaluation metric. The latter is defined in the frequency domain as the mean square pressure in the bright zone relative to the mean square pressure in the dark zone, commonly referred to as the acoustic contrast, written as

$$\text{contrast}(f) = M_B^{-1} \sum_{m_B=1}^{M_B} |p_{m_B}(f)|^2 \Bigg/ M_D^{-1} \sum_{m_D=1}^{M_D} |p_{m_D}(f)|^2 , \tag{1}$$

where the subscript $(\cdot)_B$ and $(\cdot)_D$ denotes the bright and dark zone respectively, $p_m(f)$ is the complex pressure at microphone position m and frequency f, and M_B and M_D are the number of microphone positions sampling the bright and dark zone respectively.

The fundamental aspects of the control methods are presented in the sections below. The delay-and-sum method (based on the microphone equivalent in [7]) is regarded the simpler beamforming solution, whereas the time-domain pressure (and particle velocity) matching is the basis for all the remaining sound field control methods investigated in the current work.

2.1 Time-domain pressure matching

In the following, it is assumed that the loudspeaker and listening space interaction behaves as a linear time-invariant system. Furthermore, it is assumed that the input signal is a unit sample sequence. The pressure impulse response at microphone m due to loudspeaker ℓ is expressed as

$$\mathbf{p}_{m\ell} = \mathbf{H}_{m\ell}\mathbf{w}_{\ell}, \quad \mathbf{H}_{m\ell} = \begin{bmatrix} h_{m\ell}[0] & \cdots & 0 \\ \vdots & \ddots & \vdots \\ h_{m\ell}[I-1] & \cdots & 0 \\ 0 & \cdots & h_{m\ell}[0] \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{m\ell}[I-1] \end{bmatrix}, \quad \mathbf{w}_{\ell} = \begin{bmatrix} w_{\ell}[0] \\ \vdots \\ w_{\ell}[N_w-1] \end{bmatrix}, \tag{2}$$

where $h_{m\ell}[i]$ is the i^{th} sample of the impulse response. The combined pressure at the microphone due to the L loudspeakers is

$$\mathbf{p}_m = \mathbf{H}_m\mathbf{w}, \quad \mathbf{H}_m = \begin{bmatrix}\mathbf{H}_{m1} & \cdots & \mathbf{H}_{mL}\end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix}\mathbf{w}_1^T & \cdots & \mathbf{w}_L^T\end{bmatrix}^T. \tag{3}$$

Finally, the pressures at all the microphones can be concatenated as

$$\mathbf{p}_B = \mathbf{H}_B\mathbf{w}, \quad \mathbf{H}_B = \begin{bmatrix}\mathbf{H}_1^T & \cdots & \mathbf{H}_{M_B}^T\end{bmatrix}^T, \tag{4}$$

where subscript B denotes the bright zone. The dark zone pressure can equivalently be expressed with the corresponding microphones and denoted by the subscript D. The pressure matching cost-function to be minimized is then

$$J_{\text{p}}(\mathbf{w}) = (1-\xi)\|\mathbf{p}_t - \mathbf{H}_B\mathbf{w}\|_2^2 + \xi\|\mathbf{H}_D\mathbf{w}\|_2^2 + \alpha_{\text{e}}\|\mathbf{R}_{\text{e}}\mathbf{w}\|_2^2 + \alpha_{\text{b}}\|\mathbf{B}\mathbf{w}\|_2^2, \tag{5}$$

where $\xi \in [0,1]$ controls the tradeoff between attaining the target pressure $\mathbf{p}_t$ in the bright zone and silence in the dark zone and the penalties $\|\mathbf{R}_{\text{e}}\mathbf{w}\|_2^2$ and $\|\mathbf{B}\mathbf{w}\|_2^2$ are introduced to control the shape of the resulting finite impulse response (FIR) filters and the frequency range where the loudspeakers should be active as suggested in [8] and [9], respectively. The FIR filters which minimize this quadratic cost-function can be expressed in closed-form as

$$\mathbf{w}_{\text{p}} = \left((1-\xi)\mathbf{H}_B^T\mathbf{H}_B + \xi\mathbf{H}_D^T\mathbf{H}_D + \alpha_{\text{e}}\mathbf{R}_{\text{e}}^T\mathbf{R}_{\text{e}} + \alpha_{\text{b}}\mathbf{B}^T\mathbf{B}\right)^{-1}\mathbf{H}_B^T\mathbf{p}_t. \tag{6}$$

2.2 In situ P-U beamforming

Constraining both the sound pressure and particle velocity at the control points of concern, provides a formulation which includes a vectorial control component. The particle velocity component of the sound field in a given direction can be estimated from the finite difference approximation to Euler's equation of motion between two pressure measurements displaced along the axis of interest as [7]

$$u_x(t) = -\int_{-\infty}^{t} \frac{1}{\rho}\frac{\partial p(t)}{\partial x}dt \approx -\int_{-\infty}^{t} \frac{p_{x_2}(t) - p_{x_1}(t)}{\rho\Delta x}dt. \tag{7}$$

In the above, x_1 and x_2 denotes points on the axis of interest, $\Delta x = x_2 - x_1$, and ρ is the density of air. As the pressures are available as impulse responses from a causal system, the particle velocity is zero before the beginning of the impulse responses, hence, the integral only needs to start at $t = 0$. Furthermore, as the pressure is only known at discrete time intervals, the integral must be evaluated as a numerical integration using e.g. the trapezoidal rule. The particle velocity impulse responses can then be expressed as

$$\mathbf{u}_{m\ell,x} = \mathbf{U}_{m\ell}\mathbf{w}_\ell = -(1-\eta)\frac{\Delta T}{\rho \Delta x}\mathbf{T}(\mathbf{H}_{m\ell,x_2} - \mathbf{H}_{m\ell,x_1})\mathbf{w}_\ell, \quad \mathbf{T} = \begin{bmatrix} 1/2 & 0 & 0 & \cdots & 0 \\ 1/2 & 1/2 & 0 & \cdots & 0 \\ 1/2 & 1 & 1/2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/2 & 1 & 1 & \cdots & 1/2 \end{bmatrix}, \tag{8}$$

where $\mathbf{T}$ is the chosen integration rule (here trapezoidal) expressed in a matrix, $\eta \in [0,1]$ is a weight for controlling the relative importance of controlling the pressure and the particle velocity, and ΔT is the sampling interval (here 1/48 kHz). Extending $\mathbf{U}_{m\ell,x}$ to all microphones and loudspeakers of concern, the P-U matching cost function is written as

$$J_{\mathrm{pu}}(\mathbf{w}) = (1-\xi)\left\| \begin{bmatrix} \eta \mathbf{p}_t \\ (1-\eta)\mathbf{u}_t \end{bmatrix} - \begin{bmatrix} \bar{\mathbf{H}}_B \\ \mathbf{U}_B \end{bmatrix} \mathbf{w} \right\|_2^2 + \xi \left\| \begin{bmatrix} \bar{\mathbf{H}}_D \\ \mathbf{U}_D \end{bmatrix} \mathbf{w} \right\|_2^2 + \alpha_{\mathrm{e}} \|\mathbf{R}_{\mathrm{e}}\mathbf{w}\|_2^2 + \alpha_{\mathrm{b}} \|\mathbf{B}\mathbf{w}\|_2^2, \tag{9}$$

where $\mathbf{u}_t$ is the target particle velocity in the bright zone and $\bar{\mathbf{H}}_B = \eta \mathbf{H}_B$. The resulting solution is

$$\mathbf{w}_{\mathrm{pu}} = \left((1-\xi)(\bar{\mathbf{H}}_B^T \bar{\mathbf{H}}_B + \mathbf{U}_B^T \mathbf{U}_B) + \xi(\bar{\mathbf{H}}_D^T \bar{\mathbf{H}}_D + \mathbf{U}_D^T \mathbf{U}_D) + \alpha_{\mathrm{e}} \mathbf{R}_{\mathrm{e}}^T \mathbf{R}_{\mathrm{e}} + \alpha_{\mathrm{b}} \mathbf{B}^T \mathbf{B}\right)^{-1} (\eta \bar{\mathbf{H}}_B^T \mathbf{p}_t + (1-\eta)\mathbf{U}_B^T \mathbf{u}_t). \tag{10}$$

The performance of the in situ beamforming and in situ P-U beamforming relies on the sound field characteristics captured in the measured impulse responses. Thereby, the main difference between the investigated methods in this paper is the geometry defining the bright and dark zone, as well as the environment in which the impulse responses were measured.

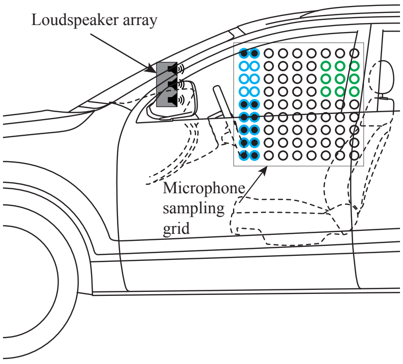
3 EXPERIMENTAL METHODOLOGY

3.1 Setup

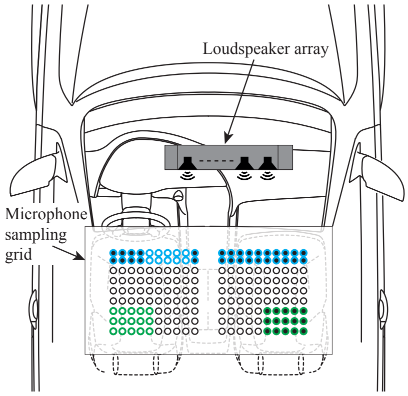
The compact loudspeaker array employed in this work includes a total of 19 elements in a planar configuration, as illustrated in 1d, each comprised by a 35 mm 'full-range' loudspeaker in a closed box cabinet.

To determine control filters for the free field beamforming method, the compact loudspeaker array was positioned in the center of a free field measurement system. Anechoic impulse responses from each loudspeaker were measured to control points on a hemicylindrical surface (half-circles stacked), as illustrated in Fig. 1d. The points representing the bright zone were selected relative to the positions representing the target direction towards the listener, whereas the remainder of the hemicylindrical surface was selected as the dark zone.

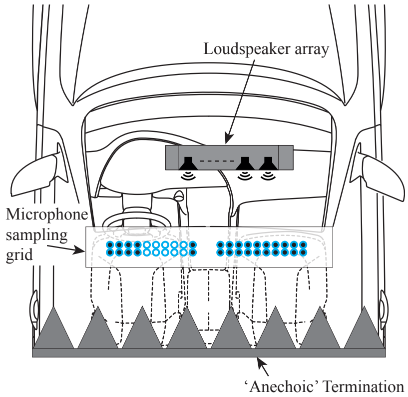
The impulse responses estimated in the car cabin were based on measurements sampling the cabin in a 3-D control point grid, using a planar array of quarter-inch microphones. The loudspeaker array was positioned and centered on the dashboard, and the 3-D measurement grid was covering the front-row seat area from positions in close proximity to the loudspeaker array and along the length-dimension, according to the layout shown in Fig. 1b. The filters for the in situ beamforming and P-U methods were calculated using the impulse responses from the first and the first two planes in front of the loudspeaker array, respectively, as illustrated in Fig. 1b. Furthermore, an additional version of these two methods were based on impulse response measurements where a wall of open-cell foam wedges was installed immediately in front of the front seats. Hence, an anechoic termination was introduced, valid above approximately 700 Hz at normal incidence, and is shown in Fig. 1c. The purpose of introducing the termination was to effectively separate the front seat area from the remainder of the car cabin, and to eliminate reflections from the front seats. The in situ zonal control was based on impulse response measurements inside the desired control regions, as illustrated in Fig. 1b.



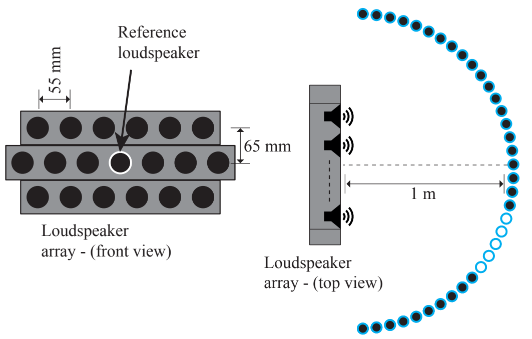
(a) Full cabin - in situ zonal control.



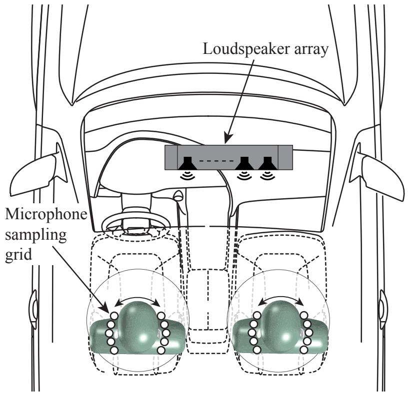
(b) Full cabin - in situ zonal and beamforming.



(c) Anechoic terminated cabin - in situ beamforming and P-U.



(d) Loudspeaker array and free field setup.



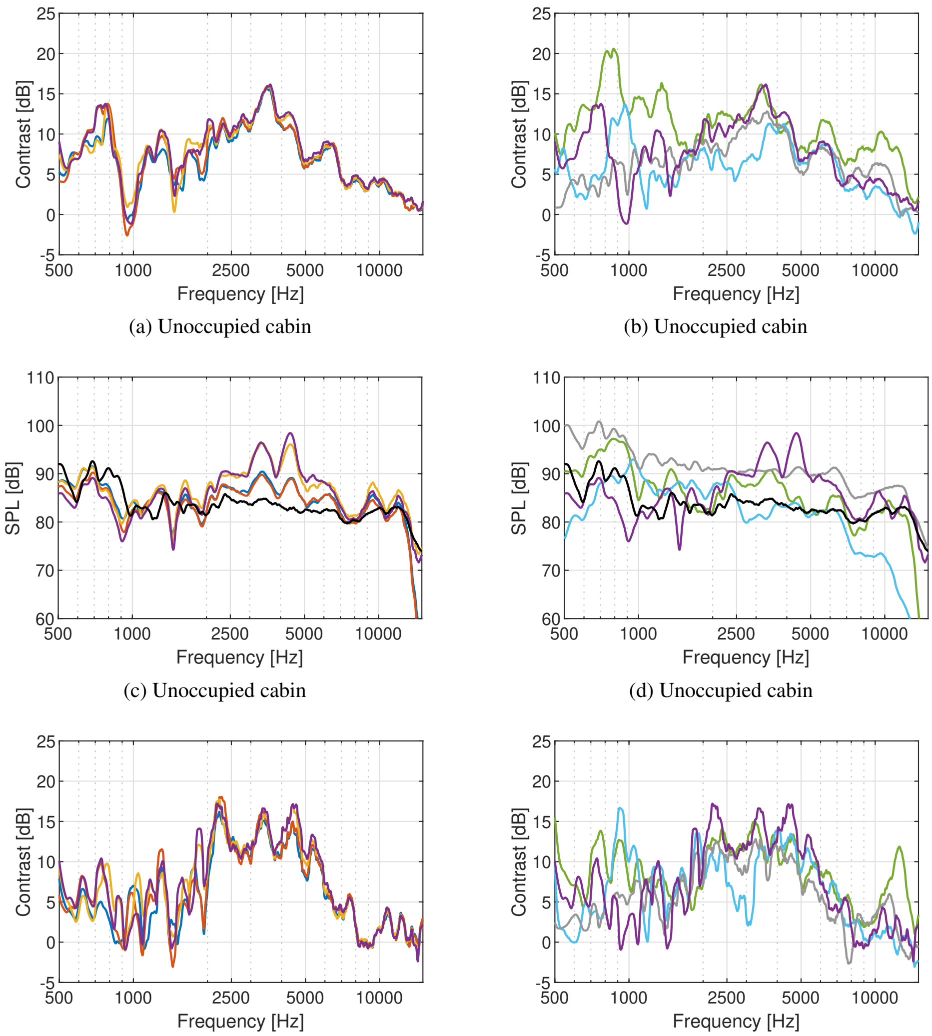
(e) Full cabin - head and torso simulator

Figure 1. Topview and sideview illustrations of the different setups considered for control filter calculation (details provided in sub-captions) and evaluation points. The open black circles represent the unconstrained control points. Open blue circles represent the bright zone control points utilized for the beamforming methods, whereas the closed blue circles indicate the dark points. The control points used for setting up the zonal control method follow the same convention, but with green circles.

In all the methods investigated here, the target pressure and particle velocity were defined as the response from the reference loudspeaker to the points sampling the bright zone, delayed by 100 samples at 48 kHz sampling frequency. The length of the control filters were 350 taps for all methods and the parameter adjusting the weight of bright vs. dark zone was $\xi = 0.9$. The parameter η was $\sqrt{0.04}$, α_e and α_b were set to 10^{-9} and 10^{-3} for the P-U method and to 10^{-7} and 10^{-1} for the remaining methods. Furthermore, the impulse responses used for estimating the particle velocity were lowpass filtered using a 4th order Butterworth filter with cutoff frequency 2 kHz [7], due to the errors inherently introduced in the finite difference approximation of the pressure gradient. A separate set of impulse responses were measured at the positions used for the zonal control. Additionally, this set of measurements was also used to evaluate the performance of the control strategies in driver and passenger seat positions of the unoccupied car. To provide a more realistic evaluation scenario, a measurement series was conducted with a head and torso simulator, Brüel & Kjær Type 4100, with in-ear microphones. The sound field was sampled in 3 heights and 5 head rotations, with superposition between the HATS and a traditional manikin on the front seats, as indicated in Fig. 1e.

4 RESULTS

To evaluate the idealized separation as well as a more realistic separation, the contrast results are evaluated for both the occupied and unoccupied car cabin, as introduced in the previous section. Additionally, the mean square pressure in the bright zone is reported to show the spectral emphasis introduced with the control methods.



(a) Unoccupied cabin (b) Unoccupied cabin

(c) Unoccupied cabin (d) Unoccupied cabin

(e) HATS (f) HATS

Figure 2. Experimental results depicting the contrast between the front seats in the unoccupied cabin or occupied by HATS. The middle plots show the mean square pressure in the unoccupied driver's seat as sound pressure levels (SPL). All results are smoothed by a 1/12 octave moving average fliter. (—): Delay-and-sum beamformer. (—): free field beamformer. (—): In situ beamformer. (—): In situ beamformer, anechoic. (—): In situ p-u beamformer. (—): In situ p-u beamformer, anechoic. (—): Zonal control. (—): Reference loudspeaker response.

The results are shown in Fig. 2 where the in situ beamforming and P-U methods are compared in the left column. The reference loudspeaker response is included as reference in the middle row. The response from this loudspeaker was used as target pressure and particle velocity in the bright zone. Results are reported using impulse responses measured both with and without the anechoic termination. It is seen that the methods all provide similar contrast results, both when evaluated in the unoccupied cabin and when evaluated with the HATS measurements. The main difference is observed in the mean square pressure in the bright zone between 2.5 and 6 kHz, where the in situ P-U method provides a boost.

In the right column of Fig. 2, the zonal control is compared with the delay-and-sum beamforming method, the free field beamforming, and the in situ P-U method based on the impulse responses measured with anechoic termination installed. It is seen that the zonal control is superior in terms of contrast evaluated in the unoccupied cabin. However, when the contrast is evaluated based on the HATS measurements, the methods provide overall more similar results. It is seen that the delay-and-sum solution is slightly suboptimal, when the acoustic contrast is evaluated throughout the frequency range of interest. However, it is also found that the resulting mean square pressure generated in the bright zone in general is more uniform when compared to the other methods.

5 DISCUSSION

From the experimental results, a number of observations can be made regarding the specific application of sound field control in this work. The first point relates to the sound field control by means of the in situ beamforming approaches. Only minor differences in the resulting performance are found, when comparing filters determined based upon impulse responses measured in the full car cabin and with an anechoic termination. Likewise, adding a constraint on the particle velocity on top of the pressure also fails to significantly alter the resulting solution. From these results it is assumed that the measurements were taken in such close proximity to the loudspeaker array, that they are dominated by the direct sound. Thereby, the anechoic termination does not provide significant changes in the responses, as the reverberation is already significantly attenuated relative to the direct sound in the particular source/receiver configuration. Furthermore, the particle velocity component in the length direction of the car is already heavily dominated by the direct sound impinging towards the measurement microphone planes, and the introduction of a vectorial physical control parameter proves to be less significant in this specific case.

It is seen that the conditions, under which the acoustic separation between the two front seats is evaluated, heavily influence the observed differences between the proposed methods. When evaluated in the unoccupied volumes at each seat, the zonal control is seen to provide superior contrast. This is expected as the measurements used in this method very closely resemble the measurements used to evaluate this result. However, when the seats are occupied by a HATS, it is seen that the observed separation is much more similar between the investigated methods. It is interesting to observe that introducing listeners in the intended bright and dark zone, removes most of the expected benefits from impulse responses measured within the car cabin. This implies that in order to attain the additional benefit of in situ measurements, the measurements should closely resemble the reproduction scenario i.e. including information about how the current occupancy of the cabin affects the sound field from the loudspeakers to the ears of the occupants.

The frequency range of the control performance seems to be divided into three regions. Below 1.5 kHz the results indicate that significant interaction occurs between the car cabin and the control loudspeakers. This is observed from the zonal control being the only method providing high separation in that particular range and that the response of the reference driver exhibits increased frequency variation in the mean square pressure in the bright zone. Between 1.5 kHz and 4-5 kHz, the beamforming approaches become capable of effectively focusing sound towards one seat relative to the other as seen from the improved performance of the delay-and-sum method in this region. Finally, above 4-5 kHz the contrast is seen to decrease with frequency. This can be explained by the spatial aliasing due to the horizontal and vertical displacement of the drivers. Given the vertical displacement of 65 mm and horizontal displacement of 27.5 mm, the expected frequencies for onset of sidelobes are around 2.6 kHz and 6.2 kHz respectively.

6 CONCLUSION

Throughout this paper a number of methods were investigated for reproducing audio to one of the front seats in a car using a planar loudspeaker array. The in situ beamforming and P-U methods were shown to yield almost identical results, which can be attributed to the direct sound having the most significant contribution in the measured impulse responses. The acoustic separation performance was seen to lie between the delay-and-sum beamforming and the zonal control method, indicating that constraining the sound field in close proximity to loudspeaker array generates a beamforming solution.

The best choice of control method depends on the availability of the impulse responses accurately representing the reproduction scenario. The delay-and-sum beamforming provides a uniform mean square pressure relative to the in situ methods and is a suitable approach when representative system characterization is not available. If accurate knowledge of the reproduction scenario inside the control regions is available, the zonal control provides the highest possible separation between the zones, among the investigated methods. The associated variation of the mean square pressure in the target zone, could then be equalized in a global sense.

REFERENCES

[1] Coleman, P.; et al. Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array. J. Acoust. Soc. Am., Vol 135(4), 2014.

[2] Buerger, M.; et al. Broadband multizone sound rendering by jointly optimizing the sound pressure and particle velocity. J. Acoust. Soc. Am., Vol 143(3), 2018, pp 1477-1490.

[3] Choi, J.W.; et al. Manipulation of sound intensity within a selected region using multiple sources. J. Acoust. Soc. Am., Vol 116(2), 2004, pp 843-852.

[4] Gálvez, M.F.S.; et al. The effect of reverberation on personal audio devices. J. Acoust. Soc. Am., Vol 135(5), 2014, pp 2654-2663.

[5] House, C.; et al. Personal spatial audio in cars, Proceedings of the Institute of Acoustics, University of South Hampton ISVR, Vol 39 (2), 2017.

[6] Olivieri, F; et al. Loudspeaker array processing for multi-zone audio reproduction based on analytical and measured electroacoustical transfer functions. AES 52nd International Conference, Guildford, UK, 2013 Sep. 2-4.

[7] Jacobsen, F.; Juhl, P. M. Fundamentals of general linear acoustics, Wiley, 2013.

[8] Møller, M. B.; et al. Sound Zones: On envelope shaping of FIR filters. 24th International Congress on Sound & Vibration, London, UK, 2017 Jul. 23-27.

[9] Gálvez, M. F. S.; et al. Time domain optimizatiion of filters used in a loudspeaker array for personal audio, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol 23 (11), 2015, pp. 1869-1878.