

Remote Microphone Virtual Sensing with Nested Microphone Sub-Arrays

Achilles Kappis, Jihui (Aimee) Zhang, and Jordan Cheer^{a)}

*Institute of Sound and Vibration Research, University of Southampton,
Southampton, SO17 1BJ, UK*

1 This study uses the Remote Microphone Technique to investigate the use of nested
2 sub-arrays that incorporate either pressure, or both pressure and pressure gradient
3 information to estimate the pressure at remote locations in a random sound field. The
4 sub-arrays of either pressure sensors or closely spaced microphone pairs are nested
5 to form both uniform linear and circular arrays. The performance of the different
6 configurations is evaluated through both experiments and simulations in terms of
7 the level of estimation error and the spatial extent over which a low estimation error
8 can be achieved. The presented results show that the use of nested arrays of closely
9 spaced microphone pairs outperforms conventional arrays that use pressure alone,
10 both in terms of the estimation error and the size of the estimation zone. Overall,
11 the circular configurations are shown to outperform the equivalent linear configura-
12 tions. The gains in nominal performance, however, are paid for by an increase in the
13 condition number, which influences the robustness of these arrays to uncertainties.
14 The paper highlights the importance of array topology and the advantages provided
15 by the inclusion of pressure gradient information into the estimation of the pressure
16 at remote locations.

^{a)}j.cheer@soton.ac.uk

I. INTRODUCTION

Active Noise Control (ANC) has been the subject of ongoing research for more than three decades and has been proven to be effective in addressing various noise-related challenges where passive control measures are either impractical due to weight and size restrictions, or do not achieve sufficient levels of performance¹. However, global control of noise throughout an acoustic environment using active systems is limited due to both the observability and controllability of the sound field^{2,3}. As a result, research has explored the use of ANC to achieve attenuation of noise at specific locations in space, providing local control⁴. In conventional local ANC applications, such as the active headrest⁵, the disturbance is controlled at the location of monitoring microphones, however, in certain applications, placing microphones at the position of interest is not feasible. For example, with the active headrest it is not possible to position the monitoring microphones in or very close to the ears of the user. To address this problem, various Virtual Sensing (VS) techniques have been proposed to estimate the pressure at points away from the monitoring sensors and shift the zone of local control to some desired remote position⁶.

Regardless of the VS method used, accurately estimating the sound field at the remote positions is of paramount importance. Erroneous estimation can significantly impact the performance of the system, limiting the maximum attainable attenuation⁷. Conventionally, VS methods have primarily been investigated and implemented using omnidirectional sensors to sample the disturbance field and acquire pressure information at discrete positions^{8,9}. However, incorporating pressure gradient information into the estimation process can improve

the accuracy¹⁰ and enhance observability¹¹. Moreover, strategies utilising pressure gradient information to control the total acoustic energy^{12–15} and intensity¹⁶ have been shown to outperform conventional pressure minimisation methods.

While sensors that directly measure pressure and particle velocity exist and have been used in past studies^{10,16}, pressure gradient components can also be estimated via the pressure differential between closely spaced matched pressure microphones^{17,18}. The performance of acoustic energy sensors based on this principle has been investigated both theoretically^{19–21} and experimentally²¹ and it has been shown that they can provide accurate measurements of the pressure, pressure gradient, intensity and total acoustic energy.

Moreau et al.¹⁰ used the pressure at two closely spaced microphones to control the pressure at a virtual microphone location along the axis of the microphones and showed that the control performance is comparable to using the pressure and pressure gradient at a single point. The study described in the current paper expands upon the virtual sensing aspect of this previous work, by employing the Remote Microphone Technique (RMT) VS method⁶ with nested microphone arrays consisting of multiple pairs of closely spaced microphones to estimate the pressure at virtual microphone locations over a two-dimensional grid. The aim of this work is to investigate how nested microphone arrays built up from closely spaced microphone pairs can reduce the estimation error and increase the spatial extent of the region over which effective estimation can be achieved. Circular and linear arrays are considered, highlighting the effects the number of microphones, array and sub-array geometries have on estimation performance in each case.

The paper is structured as follows: [section II](#) outlines the formulation of the RMT estimation method; the considered experimental setup is presented in [section III](#), which includes a description of the different nested microphone array geometries considered; in [section IV](#) the measured responses are used to implement the RMT and the estimation performance is interrogated over a spatial grid of virtual microphone positions; to allow further insight into the limitations of the different nested microphone array configurations, [section V](#) introduces numerical simulations of the experimentally implemented microphone arrays, which allows the estimation error to be explored over a greater spatial extent with finer resolution than was possible in the experimental implementation; finally, [section VI](#) summarises the main conclusions.

II. REMOTE MICROPHONE TECHNIQUE VIRTUAL SENSING

The generalised block diagram for a virtual sensing system is shown in [Figure 1](#). In this general case it is assumed that the pressure field is generated by N_v primary sources, whose complex source strengths are $\mathbf{v} = [v_1, v_2, \dots, v_{N_v}]^T$, where $[\cdot]^T$ denotes transposition, and the frequency dependence has been suppressed for notational convenience. The signals are assumed to be realisations of uncorrelated wide-sense stationary random processes and are characterised by their power spectral densities. These primary sources generate the disturbance signals measured at N_m monitoring sensors, $\mathbf{d}_m = [d_{m_1}, d_{m_2}, \dots, d_{m_{N_m}}]^T$, and N_e virtual microphones, $\mathbf{d}_e = [d_{e_1}, d_{e_2}, \dots, d_{e_{N_e}}]^T$. The disturbance signals can be expressed

78 as

$$\mathbf{d}_m = \mathbf{P}_m \mathbf{v} \quad (1a)$$

$$\mathbf{d}_e = \mathbf{P}_e \mathbf{v}, \quad (1b)$$

79 where $\mathbf{P}_m \in \mathbb{C}^{N_m \times N_v}$ and $\mathbf{P}_e \in \mathbb{C}^{N_e \times N_v}$ are the Frequency Response Functions (FRFs) between
 80 the primary sources and the monitoring and virtual microphones respectively.

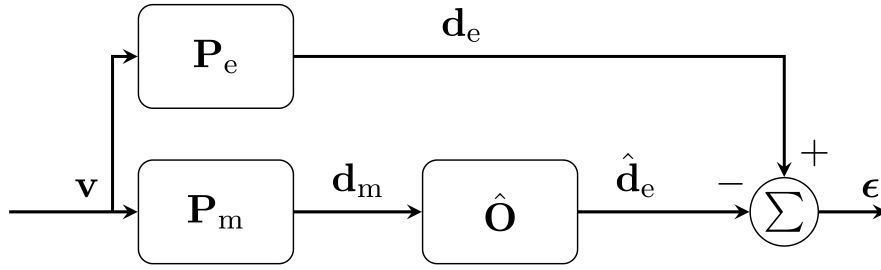


FIG. 1. Block diagram of a virtual sensing system, where the disturbance signals measured at the error microphones, $\mathbf{d}_e$, are estimated from the disturbance signals measured at the monitoring microphones $\mathbf{d}_m$, to give the estimated disturbance signals at the virtual locations, $\hat{\mathbf{d}}_e$.

81 As shown in Figure 1, according to the RMT, the disturbance field $\mathbf{d}_e$ at the virtual
 82 microphone locations is estimated by applying an observation filter $\hat{\mathbf{O}}$ to the monitoring
 83 microphone responses $\mathbf{d}_m$. The estimation error is defined as the difference between the true
 84 and estimated disturbance at the virtual microphones, given as

$$\boldsymbol{\epsilon} = \mathbf{d}_e - \hat{\mathbf{d}}_e = \mathbf{d}_e - \hat{\mathbf{O}}\mathbf{d}_m = \mathbf{P}_e \mathbf{v} - \hat{\mathbf{O}}\mathbf{P}_m \mathbf{v}. \quad (2)$$

85 The symbol $[\hat{\cdot}]$ denotes that the value is an estimate. The optimal observation filter can be
 86 calculated in the least-squares sense by minimising the cost function²²

$$\begin{aligned}
 J &= \mathbb{E} [\boldsymbol{\epsilon}^H \boldsymbol{\epsilon}] = \text{tr} \{ \mathbb{E} [\boldsymbol{\epsilon} \boldsymbol{\epsilon}^H] \} \\
 &= \text{tr} \left\{ \mathbf{S}_{\text{ee}} - \mathbf{S}_{\text{me}} \hat{\mathbf{O}}^H - \hat{\mathbf{O}} \mathbf{S}_{\text{me}}^H + \hat{\mathbf{O}} \mathbf{S}_{\text{mm}} \hat{\mathbf{O}}^H \right\},
 \end{aligned} \tag{3}$$

where $\mathbb{E}[\cdot]$ is the expectation operator, $\text{tr}\{\cdot\}$ denotes the trace of a matrix, $[\cdot]^H$ denotes Hermitian transposition, $\mathbf{S}_{\text{ee}}$ and $\mathbf{S}_{\text{mm}}$ are the power spectral density matrices of the disturbance field at the virtual and monitoring microphones respectively and $\mathbf{S}_{\text{me}}$ is the cross spectral density matrix between the monitoring and virtual microphone signals. Similarly, the power spectral density matrix of the primary source strengths is $\mathbf{S}_{\text{vv}}$ and the power and cross spectral density matrices can be expressed as

$$\mathbf{S}_{\text{ee}} = \mathbb{E}[\mathbf{d}_e \mathbf{d}_e^H] = \mathbf{P}_e \mathbf{S}_{\text{vv}} \mathbf{P}_e^H \tag{4a}$$

$$\mathbf{S}_{\text{mm}} = \mathbb{E}[\mathbf{d}_m \mathbf{d}_m^H] = \mathbf{P}_m \mathbf{S}_{\text{vv}} \mathbf{P}_m^H \tag{4b}$$

$$\mathbf{S}_{\text{me}} = \mathbb{E}[\mathbf{d}_e \mathbf{d}_m^H] = \mathbf{P}_e \mathbf{S}_{\text{vv}} \mathbf{P}_m^H \tag{4c}$$

$$\mathbf{S}_{\text{vv}} = \mathbb{E}[\mathbf{v} \mathbf{v}^H]. \tag{4d}$$

If the number of independent disturbance sources is greater than the number of microphones in the monitoring array configuration, it is guaranteed that the power spectral density matrix of the monitoring microphones, $\mathbf{S}_{\text{mm}}$, will be invertible and the optimal observation filter $\hat{\mathbf{O}}_{\text{opt}}$ can be calculated by minimisation of [Equation 3](#), which gives²²

$$\begin{aligned}
 \hat{\mathbf{O}}_{\text{opt}} &= \mathbf{S}_{\text{me}} (\mathbf{S}_{\text{mm}} + \beta \mathbf{I})^{-1} \\
 &= \mathbf{P}_e \mathbf{S}_{\text{vv}} \mathbf{P}_m^H (\mathbf{P}_m \mathbf{S}_{\text{vv}} \mathbf{P}_m^H + \beta \mathbf{I})^{-1},
 \end{aligned} \tag{5}$$

where equations (4b) and (4c) have been used to expand $\mathbf{S}_{\text{mm}}$ and $\mathbf{S}_{\text{me}}$. The non-negative real number β is a regularisation parameter, which can be frequency dependent and $\mathbf{I}$ is an $N_m \times N_m$ identity matrix. Regularisation is used to constrain the magnitude of the

observation filter⁸, which can increase robustness to uncertainties at the expense of reduced
 estimation accuracy⁷ and is equivalent to introducing white noise to the measured monitoring
 signals²³. The solution becomes increasingly less sensitive to uncertainties for higher values
 of β , but a bias is introduced leading to higher estimation errors.

If the system instead has more monitoring microphones than independent disturbance
 sources, and so is overdetermined, the power spectral density matrix of the monitoring
 microphones will become singular. In this case, the minimum-norm observation filter can
 instead be used, which is given by

$$\hat{\mathbf{O}}_{\text{opt,u}} = \mathbf{P}_e (\mathbf{P}_m^H \mathbf{P}_m + \beta \mathbf{I})^{-1} \mathbf{P}_m^H. \quad (6)$$

The derivation of this form of the observation filter is provided in [Appendix A](#). Substituting
 the observation filter given by [Equation 6](#) into [Equation 2](#) gives the estimation error as

$$\begin{aligned}
 \mathbf{e} &= \mathbf{P}_e \mathbf{v} - \hat{\mathbf{O}}_{\text{opt,u}} \mathbf{P}_m \mathbf{v} \\
 &= \mathbf{P}_e \mathbf{v} - \mathbf{P}_e (\mathbf{P}_m^H \mathbf{P}_m + \beta \mathbf{I})^{-1} \mathbf{P}_m^H \mathbf{P}_m \mathbf{v} \\
 &\stackrel{\beta=0}{=} \mathbf{P}_e \mathbf{v} - \mathbf{P}_e \mathbf{v} = \mathbf{0}.
 \end{aligned} \quad (7)$$

Therefore, in the absence of noise, and when the regularisation parameter β is set to zero,
 the observation filter in the overdetermined case given by [Equation 6](#) will always result in
 perfect estimation, regardless of the sound field or the spatial arrangement of the system
 components.

The normalised mean squared estimation error is used in this work to quantify the esti-
 mation performance and is defined as

$$L_e = 10 \log_{10} \left(\frac{\text{tr} \{ \mathbf{S}_{\epsilon\epsilon} \}}{\text{tr} \{ \mathbf{S}_{ee} \}} \right), \quad (8)$$

where $\mathbf{S}_{\epsilon\epsilon} = \mathbb{E}[\epsilon\epsilon^H]$ is the mean squared estimation error. The metric will acquire negative values for good estimation with $-\infty$ representing perfect estimation accuracy.

III. EXPERIMENTAL SETUP AND PROCEDURE

A schematic diagram of the physical two-dimensional arrangement of sources and sensors used in this work is illustrated in Figure 2. The setup was implemented in the anechoic chamber at the Institute of Sound and Vibration Research, University of Southampton; this allows the performance of the nested microphone arrays to be evaluated in a well-defined acoustic environment before future work considers specific applications. A constellation of sixteen Genelec 8020D loudspeakers were used as primary sources equispaced on the circumference of a circle with a radius of 3 m. GRAS 40PL-10 and Brüel & Kjær Type 4958-A 1/4" omnidirectional microphones were used to implement two monitoring microphone configurations, each consisting of sub-arrays of closely spaced microphones, which have been defined to detect sound field contributions related to the pressure and the pressure gradient in the different coordinate directions. A four-element Uniform Linear Array (ULA), whose elements are denoted with L in the schematic, and a six-element Uniform Circular Array (UCA), whose elements are marked with C, have been considered. The number of elements in the UCA is chosen to be higher than the ULA to take into account the increased distance between the array elements and the estimation positions. The distinct characteristics of the arrays enable evaluation of the performance against the number of elements and inter-element spacing for both array topologies. The ULA is positioned parallel to the global x -axis, translated 0.19 m towards the negative y -direction. The UCA has a radius of 0.45 m

and is centred at the origin with two of its elements being on the x -axis. The sub-arrays consist of two orthogonal pairs of microphones, with the distance between the microphones in each of the two pairs being 0.05 m, giving an aliasing frequency of 3.43 kHz. The ULA has the sub-arrays orientated so that one pair of microphones is parallel to the y -axis and the other to the x -axis. The sub-arrays forming the UCA have one pair of microphones aligned radially and the other arranged circumferentially. The virtual microphone positions, at which the monitoring arrays are utilised to estimate the disturbance pressure, are arranged on a $0.4 \text{ m} \times 0.2 \text{ m}$ grid. The distance between the positions is 0.025 m in each Cartesian direction resulting in a 17×9 uniform grid, giving a total of 153 positions at which the estimation performance is evaluated.

To investigate the estimation accuracy for the different monitoring microphone configurations utilising the RMT described in [section II](#), the responses between each of the sixteen primary sources and all of the monitoring and virtual microphone positions have been measured. This was achieved using synchronised logarithmic sine sweeps²⁴ with a duration of 2 s and energy at frequencies from 50 Hz to 5 kHz and the impulse responses were then calculated via deconvolution. The signal acquisition was performed with a National Instruments PXI-1033 system with a sampling frequency of 24 kHz. The FRFs from the sources to the virtual microphone positions, $\mathbf{P}_e$, and to the monitoring microphones, $\mathbf{P}_m$, were calculated via Fourier transformation of the corresponding impulse responses.

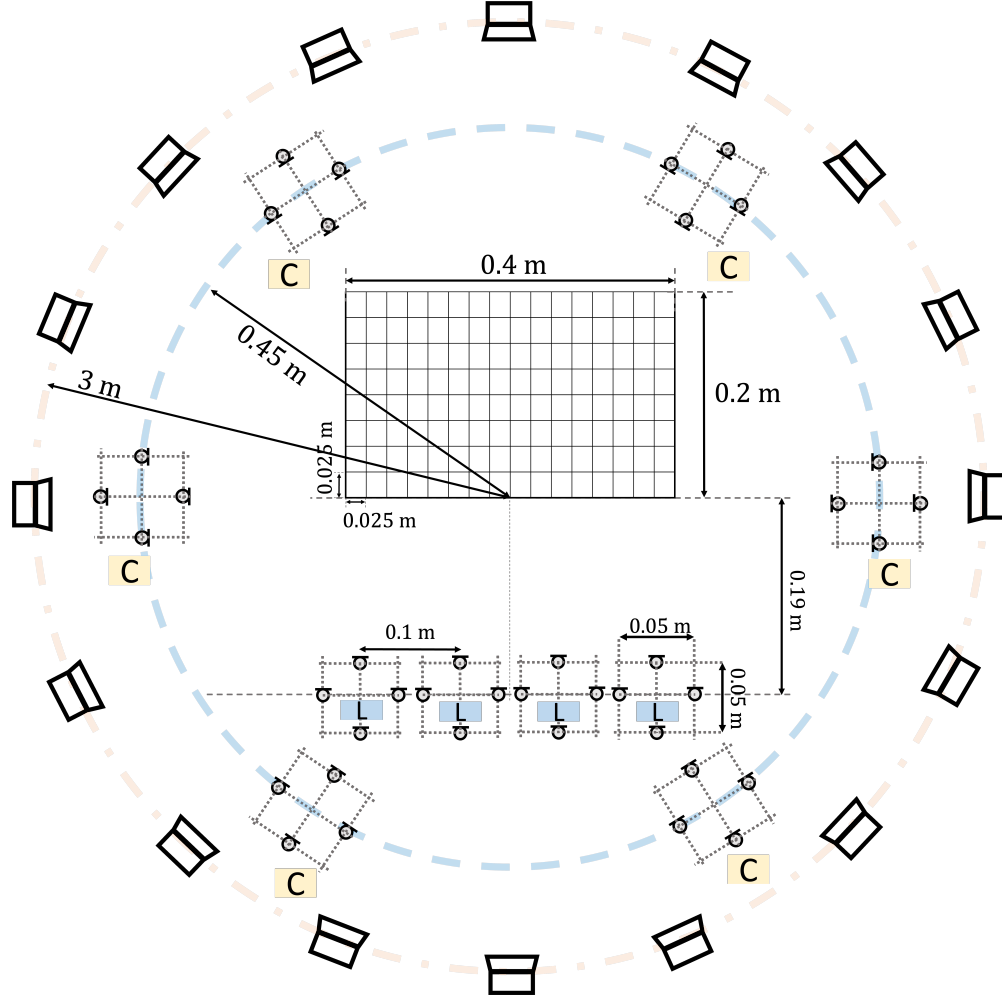


FIG. 2. Schematic diagram of the experimental configuration realised in an anechoic chamber. Sixteen primary sources are located on the outer circle with a radius of 3 m. A four-element uniform linear array denoted with L, and a six-element uniform circular array denoted with C, are deployed as monitoring microphone configurations. The individual elements comprise sub-arrays with four pressure sensors arranged on a cross shape. The virtual microphone positions are uniformly arranged on a 0.4 m $\times$ 0.2 m rectangular grid.

IV. EXPERIMENTAL RESULTS

This section presents an evaluation of the performance of the RMT estimation accuracy for the experimental configuration described in [section III](#). Deriving from the ULA and UCA microphone configurations shown in [Figure 2](#), six distinct monitoring microphone arrangements are investigated by using different subsets of the sub-array microphones, as shown in [Figure 3](#). Firstly, to facilitate comparison with previous literature utilising single pressure sensors^{8,25,26}, the two array configurations shown in [Figure 3](#) (A) were implemented by averaging the responses of the four sub-array microphones to approximate a single pressure sensor²⁰ at the centre of each sub-array; these two configurations are referred to as $ULA_{\hat{p}}$ and $UCA_{\hat{p}}$ with four and six emulated sensors respectively. Secondly, the two array configurations shown in [Figure 3](#) (B) utilise a single microphone pair in each sub-array, thus providing pressure gradient information in one direction^{10,27}; ULA_y utilises the four microphone pairs parallel to the y -axis (giving a total of eight microphones) and UCA_r utilises the six radially aligned microphone pairs (giving a total of twelve microphones). The final two microphone array configurations are shown in [Figure 3](#) (C) and consist of the fully populated sub-arrays, which provide pressure gradient information in two directions, which results in the ULA_{xy} and $UCA_{r\phi}$ configurations, with sixteen and twenty-four microphones, respectively.

The optimal observation filters were calculated for each of the six microphone array configurations using the measured FRFs and [Equation 5](#), except in the case of $UCA_{r\phi}$, which due to there being less primary sources than monitoring microphones utilises the overdetermined solution provided by [Equation 6](#). In these calculations a random field was

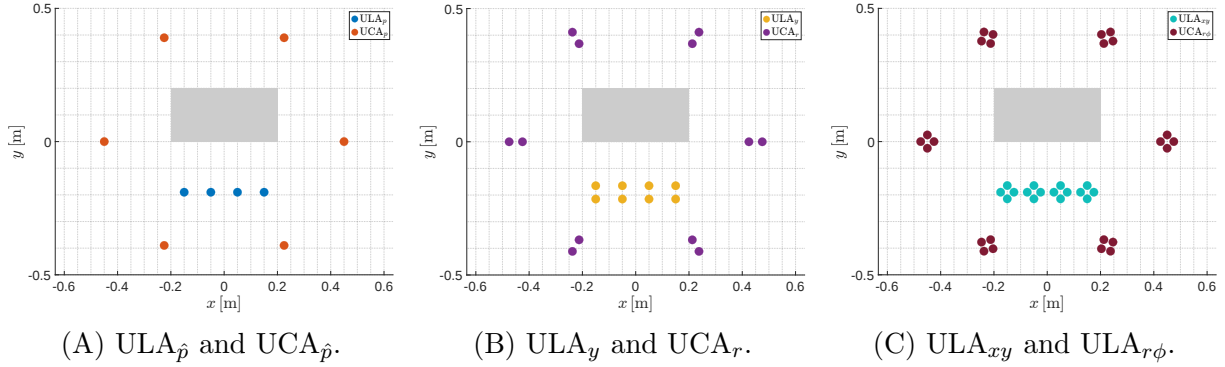


FIG. 3. The six different monitoring microphone configurations investigated. Microphone pairs in each sub-array are coloured differently and the grey rectangle corresponds to the grid of virtual microphone positions. The configurations in (A) are implemented by averaging the response of all four microphones of each sub-array to emulate a pressure sensor. The configurations in (B) comprise a pair of microphones at each array element position and the configurations in (C) have fully populated sub-arrays.

generated by assuming that the sixteen primary sources were uncorrelated with unity source strength, which results in the power spectral density matrix $\mathbf{S}_{vv}$ being an identity matrix and the filters depending only on the FRFs²².

A. Spatial variation in the estimation performance

The estimation performance of the six monitoring microphone configurations described in Figure 3 has initially been assessed over the grid of virtual microphone positions. The normalised estimation error, as defined in Equation 8, averaged over octave frequency bands with centre frequencies of 250 Hz, 500 Hz and 1 kHz, is depicted in Figure 4 for all configurations. Regularisation was applied during the calculation of the observation filters to minimise

numerical errors, with a frequency-independent regularisation parameter of $\beta = 10^{-2}$. The term *10 dB estimation zone* will be used here to define the area where the error is less than -10 dB, with the solid lines in Figure 4 representing these zones. When the estimation zone exceeds the virtual microphone positions, a solid line appears at the border of the grid.

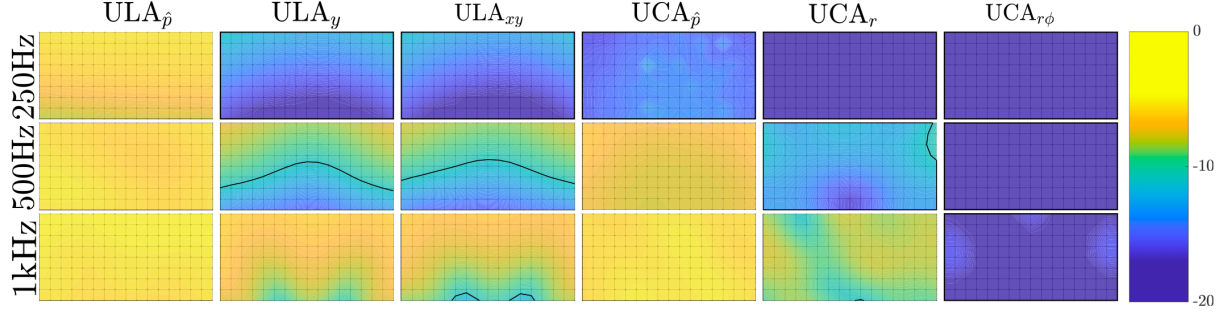


FIG. 4. Spatial distribution of the normalised mean squared estimation error over the virtual microphone grid, as shown in Figure 3. The contours show the error averaged over the octave bands with centre frequencies of 250 Hz, 500 Hz and 1 kHz. The solid lines denote the limits of the area where the error is less than -10 dB. The array configurations are denoted on the top and the subscripts represent the pairs in the sub-arrays, with $\hat{p}$ denoting the configuration with emulated single pressure sensors.

The results presented in Figure 4 demonstrate that the estimation performance deteriorates with increasing frequency for all configurations. Moreover, increasing the number of microphones within the arrays monotonically enhances the estimation accuracy and results in larger estimation zones. There is good agreement between these results and the performance of microphone arrays in simulated diffuse fields²⁸, where the estimation zones extended along the direction of microphone pairs, and the coherence between the monitor-

ing and virtual positions was identified as the primary factor affecting estimation^{10,29}. It is worth noting that, even though the arrays are designed to be symmetrical, the spatial distribution of the error, especially for the UCAs, shows some degree of asymmetry, which can be attributed to positioning errors in the implementation of the experimental setup.

From the results presented in Figure 4 it can be seen that the ULA configuration with single pressure sensors, $ULA_{\hat{p}}$, fails to generate a 10 dB estimation zone in any frequency band. The linear array configurations with sub-arrays containing two, ULA_y , and four, ULA_{xy} , microphones show similar levels of performance to each other. At low frequencies, both configurations achieve a 10 dB estimation zone extending beyond the area of the virtual microphone grid. In the mid and high frequency octave bands, the ULA_{xy} array exhibits a slightly larger 10 dB estimation zone compared to the ULA_y array. However, this extension comes at the expense of doubling the total number of microphones from eight to sixteen. In summary, it can be seen that for the ULA configurations increasing the number of microphones from one to two per sub-array provides significantly higher estimation performance, both in accuracy and spatial extent of the estimation zone, than increasing the number of microphones in each sub-array from two to four.

Considering the results presented in Figure 4 for the UCA configurations, it can be seen that in general they achieve higher spatial uniformity and estimation accuracy than their ULA counterparts. The configuration with pressure sensing in each sub-array element exhibits an estimation zone extending beyond the virtual microphone position grid at low frequencies, which is a significant improvement when compared to the corresponding linear array, although it does utilise six sensors rather than the four used by the $ULA_{\hat{p}}$ configu-

ration. The 10 dB estimation zone for the UCA configuration with two sensor sub-arrays, UCA_r , covers the majority of the virtual microphone positions in the 500 Hz octave band, but in the high frequency band the error is greater than -10 dB over the whole virtual microphone grid. Finally, the fully populated circular array, $UCA_{r\phi}$, achieves a 10 dB estimation zone larger than the virtual microphone grid over all three frequency bands, outperforming all other configurations, but requiring 24 microphones.

The increased performance of the circular microphone configurations can be partly attributed to the greater number of individual sub-arrays in each case. However, comparing the fully populated linear array ULA_{xy} , with sixteen microphones, and the circular configuration with two-sensor sub-arrays, UCA_r , with twelve microphones, can provide more insight into the impact of the spatial distribution of the sensors on the estimation process. The circular setup achieves a lower estimation error over the virtual microphone grid in the low octave band and generates a larger 10 dB estimation zone with higher uniformity in the mid frequency band. In the high frequency band, both configurations show comparable performance in the extent of the 10 dB estimation zones and minimum achievable error. These results indicate, unsurprisingly, that the topology of the monitoring arrangement significantly influences the estimation performance, which is consistent with previous studies^{25,28,30}. In particular, these results demonstrate that both the overall geometry of the array (i.e. linear versus circular), as well as the configuration of the sub-arrays influences the estimation performance. For example, the circular configurations provide a more distributed sampling of the sound field than the linear arrays and, thus, tend to require less element positions to accurately estimate the sound field at remote locations³¹. This tendency may

differ depending on the generated sound field, for example, if the sound field was generated by sources located exclusively below the ULA as shown in Figure 3 and therefore incident from a single direction, the ULA may in fact outperform the UCA.

B. Frequency dependence of the estimation performance

To this point, the evaluation of the monitoring microphone array configurations has focused on the spatial distribution of the estimation error over three octave bands. However, it is insightful to consider the performance in more detail over frequency. Figure 5 presents the range and the mean value of the estimation error across the virtual microphone grid for frequencies between 50 Hz and 1.5 kHz. From Figure 5(A) it can be seen that the performance of the linear array with single pressure sensors, $ULA_{\hat{p}}$, exhibits the lowest variability over the entire frequency range, however, with evidently very limited overall performance. An error of -10 dB at all virtual microphone positions is achieved only at very low frequencies, up to approximately 60 Hz. Increasing the number of microphones in the sub-arrays of the ULA offers a significant increase in the performance, as also shown in Figure 4, but the differences between the two, ULA_y , and four, ULA_{xy} , microphone sub-array configurations is relatively small. There is a slight estimation performance increase for ULA_{xy} at frequencies exceeding about 1.1 kHz, but this is probably not sufficient to justify the significant increase in the number of microphones required. It is also worth noting that for all of the linear arrays, the range in the performance around the mean slightly decreases with increasing frequency.

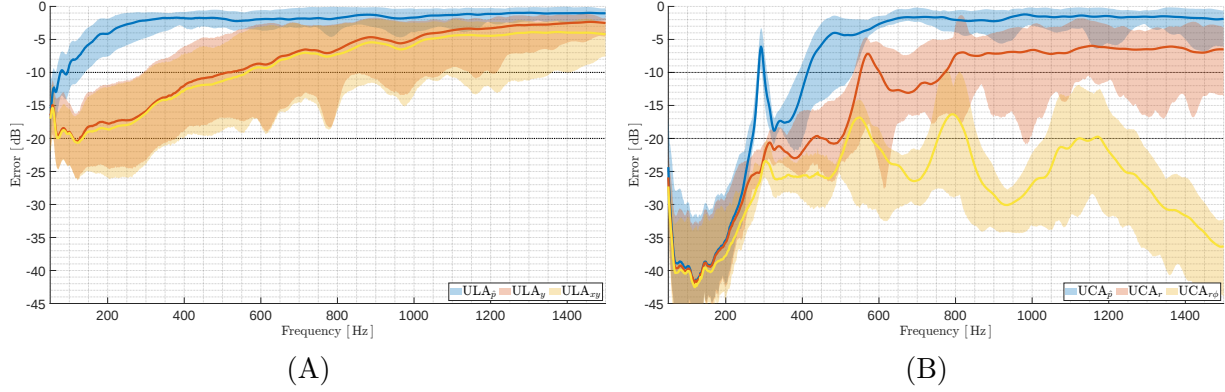


FIG. 5. Estimation error over the virtual microphone grid achieved by the (A) linear configurations and (B) circular configurations for frequencies from 50 Hz to 1.5 kHz. The solid lines indicate the mean estimation error over the virtual microphone grid, while the shaded regions indicate the range in estimation error over the full grid.

Figure 5(B) shows the corresponding results for the UCA configurations, which clearly demonstrates the significant difference in the overall behaviour compared to the linear setups. Compared to their ULA counterparts, the lowest frequency for which the error is less than the -10 dB threshold is shifted higher in frequency for all UCA configurations. However, the variation in the mean error over frequency is considerably larger over the investigated bandwidth, showing increased variability in the estimation. For frequencies up to approximately 250 Hz, the estimation error and the spatial variation across the virtual microphone grid are similar for all UCA sub-array configurations, and all circular setups outperform the ULAs irrespective of the total number of microphones in the configuration.

Moving to higher frequencies, the $UCA_{\hat{p}}$ configuration shows a pronounced reduction in the estimation accuracy at around 280 Hz. This drop in performance occurs at the frequency for which the zeroth order Bessel function associated with the array radius exhibits its first

zero^{29,32}. The two UCA monitoring microphone configurations comprising multi-microphone elements, UCA_r and $\text{UCA}_{r\phi}$, do not suffer from this problem, because the arrangement of the sub-arrays means that the microphones lie on circles of different radii³³. However, both setups exhibit additional peaks in the estimation error, for example at around 550 Hz and 800 Hz, which can be associated with zeros of higher-order Bessel functions. Except for the peak around 280 Hz, the $\text{UCA}_{\hat{p}}$ array achieves a -10 dB error across the entire grid up to approximately 380 Hz. This is a considerable frequency range extension compared to the linear arrays, which at best achieved an estimation error below -10 dB for all virtual microphone positions up to only 250 Hz. It is worth highlighting that the error for the UCA with four-microphone sub-arrays, $\text{UCA}_{r\phi}$, is consistently less than -10 dB across the entire grid for the full frequency range considered. Furthermore, the error remains below -20 dB for a larger part of the presented frequency range. The performance of this configuration is mainly attributed to the difference in the underlying formulation of the observation filter, which for this array is given by Equation 6, because the number of microphones deployed in this setup is larger than the number of independent disturbance sources. This behaviour is discussed in greater detail in section V A below.

Overall, increasing the number of microphones improves performance both over space and frequency. However, higher computational cost will inherently result from the increased number of microphones. As discussed in this section and in section IV A, the topology of the microphone array and the constituent sub-arrays significantly influences performance, suggesting that careful configuration design can minimise the number of microphones required to achieve a certain level of performance.

V. SIMULATION STUDY

The results presented in the previous section have shown, using FRFs from the experimentally implemented array configurations shown in [Figure 3](#), how the performance of the different ULAs and UCAs incorporating different sub-array configurations varies over both frequency and space. In particular, it is clear that the UCAs typically outperform the ULA configurations, but more pertinently to this study is how the inclusion of closely spaced pairs of microphones in the sub-arrays improves the estimation performance by effectively incorporating pressure gradients into the estimation. The experimental study, however, is limited by the practical density and size of the virtual microphone grid. To provide further insight into the limits of the various nested microphone arrays considered here, this section will present a numerical simulation study into the spatial limitations of the arrays. In the first instance, the numerical simulations are described and their behaviour is tuned to be consistent with the experimental implementation in terms of the condition number. Subsequently, the average estimation error over the virtual microphone grid is compared between the numerical and experimental results, before the simulated system is utilised to explore the spatial limits of the estimation error over a wider region of space than was possible in the experimental system.

A. Simulated arrays

In this section, the performance of the considered arrays is numerically simulated with the constituent microphones being modelled as ideal omnidirectional receivers. The disturbance

field is generated by 64 monopoles evenly spaced on the circumference of a circle with a radius of $r_s = 3$ m. The frequency response from a source to a receiver is expressed as³⁴

$$P_{nm}(\omega, r) = \frac{j\omega\rho}{4\pi r_{nm}} e^{-jk r_{nm}}, \quad (9)$$

where ω is the angular frequency, equal to $2\pi f$ with f being the temporal frequency, $k = \omega/c$ is the wavenumber, with c the speed of sound, j the imaginary unit for which $j^2 = -1$ is true, ρ is the density of the medium and r_{nm} denotes the distance between the n th source and m th receiver. The simulated configurations are consistent with the experimental implementations, except a larger number of sources are used to generate the disturbance sound field to overcome issues related to the matrix inversion required to calculate the optimal observation filters.

In the simulated setup, the optimal observation filters for all array configurations can be calculated using Equation 5. As briefly mentioned in section IV B, the calculation of the optimal observation filters for the circular array with four-element sub-arrays, $UCA_{r\phi}$, in the experimental investigation uses Equation 6 and this may lead to differences between the experimental and simulated array performance which will be discussed in this section. However, it is worth noting that, as illustrated in Figure 5(B), the normalised estimation error achieved by the experimental setup does not reach $-\infty$, as predicted by Equation 7, due to experimental uncertainty introduced by electrical noise and positioning errors effectively regularising the matrix inversion²³.

To facilitate a meaningful comparison between the experimental and simulated results in the following sections, the condition number, κ , of the regularised power spectral density matrices has been matched by adjusting the regularisation in the simulated case. This has

335 been achieved by applying a frequency-dependent regularisation in the calculation of the
 336 observation filters using the simulated responses. The regularisation was tuned at the edge
 337 frequencies of the octave bands with centre frequencies ranging from 62.5 Hz to 1 kHz and
 338 shape-preserving cubic interpolation was subsequently applied to calculate the regularisation
 339 factor for all frequencies of interest. The resulting condition numbers for the experimental
 340 and simulated cases for all array configurations are shown in Figure 6. From these results
 341 it can be seen that the simulated results are able to accurately match the predominant
 342 characteristics of the experimentally derived condition number. It is important to highlight
 343 that the ULA and UCA microphone configurations formed using the same sub-arrays exhibit
 344 condition numbers of comparable magnitude, indicating the influence of the constituent sub-
 345 arrays on the overall condition number. Additionally, it is worth noting that the general
 346 modal behaviour of the circular arrays has been successfully reproduced in the simulations.

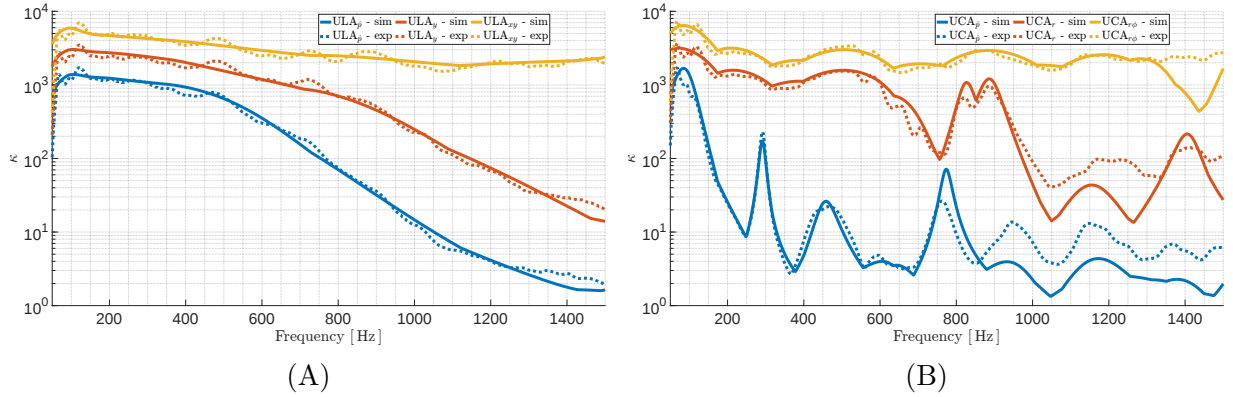


FIG. 6. Condition number κ for the simulated and experimental monitoring microphone configurations for the (A) linear, and (B) circular configurations. The condition number for the experimental setups is shown with dashed lines (- -) and that for the simulated arrays with solid lines (-).

B. Comparison of estimation for the simulated and experimental arrays

In this section, the performance of the experimental and simulated arrays is interrogated over the virtual microphone grid shown in [Figure 2](#). The frequency-dependent regularisation introduced in [section V A](#) for the simulated case has been utilised here such that the experimental and simulated arrays exhibit consistent conditioning. The average estimation error is presented in [Figure 7](#) and although the disturbance field was generated with a different number of sources in the simulations compared to the experimental setup, the performance of the simulated arrays is generally consistent over the majority of the presented spectrum. On this note, according to Kennedy *et al*³⁵, the sound field in a circular region of radius $r_e = 0.2$ m, effectively covering the virtual microphone grid utilised in the experiments, can be reproduced with 16 sources up to a frequency of about 1 kHz with a theoretical upper normalised error bound of approximately -20 dB, regardless of the sound field complexity. Thus, it may be relevant to note that discrepancies between the simulation and experimental results at frequencies above around 1 kHz may be at least partially due to differences in the generated sound fields, but are likely to have another explanation at lower frequencies.

Focusing initially on the ULA results presented in [Figure 7\(A\)](#), it is clear that for all three ULA configurations the estimation error over frequency is consistent between the simulations and experiments. It can be seen that some details that are observed in the experimental results do not appear in the simulations, most notably the ripples in the estimation. However, this is consistent with the corresponding deviations between the experimental and simulated

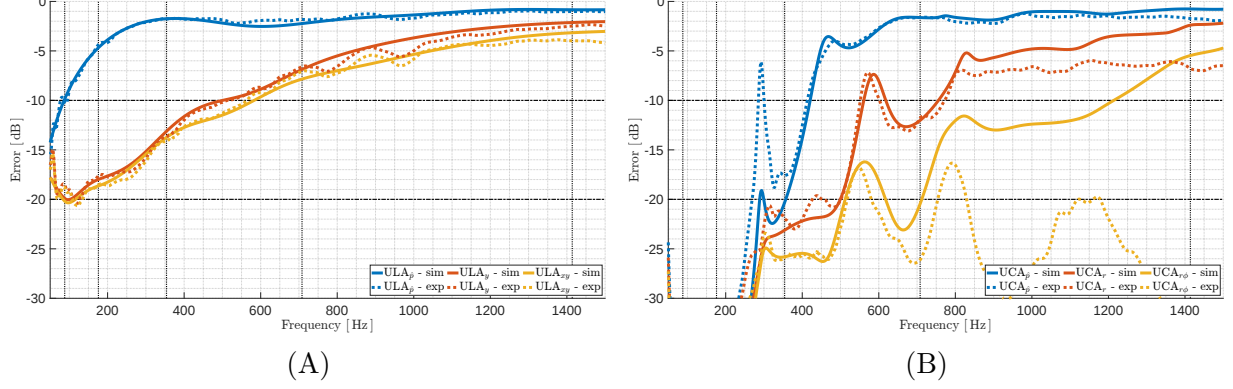


FIG. 7. Average estimation error for the experimental and simulated (A) linear and (B) circular configurations over the virtual microphone grid of Figure 2. The estimation error for the experimental setups is shown with dashed lines (--) and that for the simulated arrays with solid lines (—). The vertical lines denote the edge frequencies of each octave and the horizontal lines the -20 dB and -10 dB error

condition numbers, as shown in Figure 6(A), and can be related to the smoothly interpolated regularisation factors utilised in the simulations.

In the case of the UCA results presented in Figure 7(B), it is clear that more significant deviations between the simulation and experimental results arise compared to the ULA. For the circular array with pressure sensing, $UCA_{\hat{p}}$, the results are consistent at higher frequencies, however, the simulation results significantly outperform the experimental results at lower frequencies. This discrepancy can be related to the high condition number at low frequencies, as shown in Figure 6(B), making the performance susceptible to the uncertainties inherent in the experimental case. In contrast, the estimation performance for the circular arrays with one, UCA_r , and two microphone pairs, $UCA_{r\phi}$, shows exceptional consistency between the simulation and experimental results up to around 700 Hz. However, at higher

frequencies, larger deviations are observed, with significant discrepancies for the fully populated setup, which can primarily be attributed to the difference in the underlying formulation of the observation filters. The experimental results show performance deterioration at frequencies associated with the modal behaviour of the array, but due to the overdetermined nature of this implementation, the performance is otherwise only limited by the inherent uncertainties in the measurements^{19,20,23}. On the contrary, in the simulated case the system is underdetermined and the performance deteriorates as frequency increases, which is consistent with the other microphone arrays. Nevertheless, this comparison between experimental and simulated results helps to justify the use of the simulated systems to provide more insight into the limitations of the various array configurations over a larger spatial extent.

C. Simulated spatial estimation performance limits

To provide further insight into the spatial performance limits of the considered monitoring microphone configurations, the simulated system described in [section V B](#) is used here to assess the estimation performance over a larger region, defined by a $2\text{ m} \times 2\text{ m}$ area containing an 81×81 square grid of equally spaced virtual microphone positions centred at the origin. The disturbance field and the regularisation of the different array configurations are kept as described in [section V B](#). The -10 dB estimation zones for the octave bands with centre frequencies of 250 Hz , 500 Hz and 1 kHz are illustrated in [Figure 8](#). To facilitate comparison with the error maps depicted in [Figure 4](#), the area of the virtual microphone grid deployed in the experimental setup is overlaid as a grey rectangle.

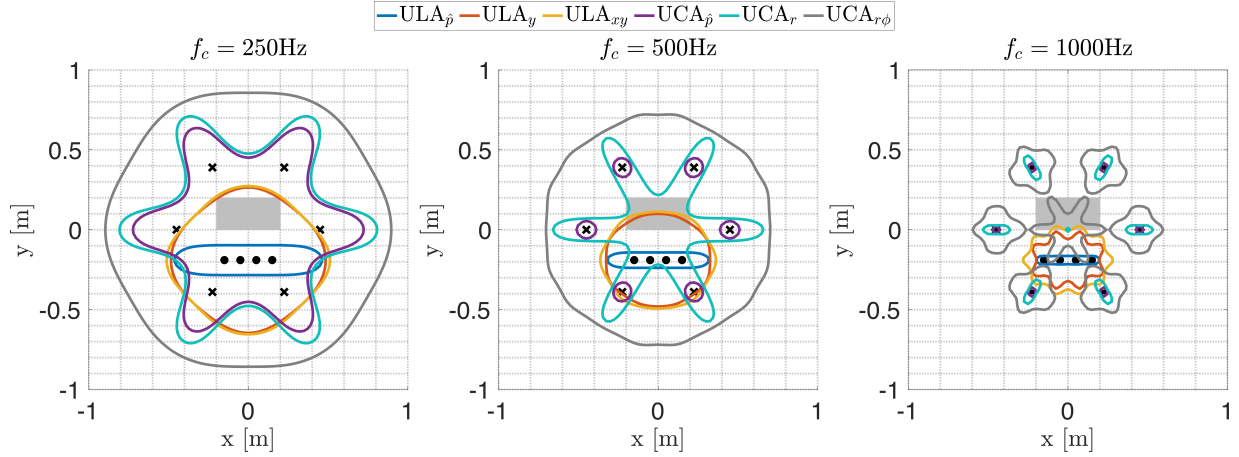


FIG. 8. Estimation zones of the simulated array configurations averaged over octave bands with centre frequencies $f_c = 250$ Hz, 500 Hz and 1 kHz. The circles denote the sub-array element positions for the ULA and the ‘x’ symbols denote the element sub-array positions for the UCA. The virtual microphone grid deployed in the experiment is represented with a grey rectangle for reference.

From the results presented in Figure 8 it can be seen that the -10 dB estimation zone for the conventional ULA array using pressure sensors, ULA_p , is an ellipsoid, with its principal axis aligned along the axis of the array and its secondary axis orientated perpendicularly, measuring about $0.1\lambda_s$ in length, where λ_s denotes the wavelength at the centre frequency of the octave band. As described in previous studies on both tonal and broadband noise, comparably sized estimation zones were generated for bandpass filtered noise and tonal disturbances with frequencies equal to the mid-frequency of the bands considered here^{8,27,36}. Consistent with the experimental results, the simulated -10 dB estimation zone for ULA_p does not encompass the grey rectangle denoting the experimental virtual microphone grid. The simulated -10 dB estimation zones for the two ULAs comprising nested microphone

arrays, ULA_y and ULA_{xy} , are also consistent with the experimental results. In the 250 Hz octave band the zone encompasses the full extent of the experimental virtual microphone grid and it can be seen in the simulated results that it encloses a much more significant region of space covering an area of approximately 0.61 m^2 . In the 500 Hz octave band, the estimation zones for ULA_y and ULA_{xy} cover about half of the experimental virtual microphone grid area in the vertical direction, which is consistent with the experimental results presented in [Figure 4](#). However, the full extent of the simulated zone is significantly larger than the experimental virtual microphone grid area, occupying an area of approximately 0.29 m^2 , centred around the array. Finally, in the 1 kHz octave band the fully populated linear array configuration, ULA_{xy} , includes a small region of the experimental virtual microphone grid in the estimation zone, which is again consistent with the experimental results. However, the simulated estimation zone is in fact comparable in size to the experimental virtual microphone grid, albeit translated in the y-direction to be centred around the microphone array. As previously observed in the experimental results, both linear array configurations incorporating sub-arrays with closely spaced microphone pairs exhibit comparable performance to each other, with notable differences primarily appearing in the 1 kHz octave band. Additionally, it is interesting to highlight that the extent of the estimation zones along the horizontal Cartesian axis are comparable, despite the higher number of microphones in this direction for the ULA_{xy} configuration.

In the case of the UCA configurations, in the 250 Hz frequency band, the estimation zone for all three sub-array configurations encompasses the experimental virtual microphone grid, which is consistent with the experimental results presented in [Figure 4](#). However, it

can be seen from the simulation results that the estimation zones do differ between the three configurations. The conventional circular array, $UCA_{\hat{p}}$, approximates that of the configuration with two-microphone sub-arrays, UCA_r , with the zone in both cases extending outwards from the microphone array in the radial directions aligning with the sub-array elements. In the case of the circular array utilising four-element sub-arrays, $UCA_{r\phi}$, a unified circular estimation zone covering a significant region extending outwards from the array is achieved with an area of approximately 2.4 m^2 . In the 500 Hz frequency band, the three circular arrays show distinct behaviour, with the array utilising only pressure, $UCA_{\hat{p}}$, generating estimation zones that, in two-dimensions, are circles concentrated around the microphones with a radius of about $0.1\lambda_s$. These zones resemble the performance of distinct pressure microphones as has been described in the literature on remote microphone virtual sensing in random sound fields^{8,10,25,28,36}. In the mid-frequency band, the UCA_r configuration is similar in form to that observed in the lower frequency band, but with a smaller overall area and sharper regions extending in the radial directions towards the sub-array elements. The fully populated circular array configuration, however, maintains an extended estimation zone, with an area that is reduced to approximately 1.6 m^2 compared to the low frequency band. Finally, in the 1 kHz octave band the $UCA_{\hat{p}}$ configuration continues to generate small estimation zones around each element, with dimensions of around $0.1\lambda_s$. The UCA_r configuration also generates individual zones around the sub-arrays in this frequency band, but in this case they are elongated radially along the direction of the sensor pair, which gives an ellipsoid shape as described in [27]. Additionally, a small area of effective estimation is also generated at the origin, which is also visible in the experimental

results for the UCA_r configuration. For the fully populated circular array, $\text{UCA}_{r\phi}$, the simulated behaviour is somewhat different from the experimental results, which showed that the estimation zones encompassed the full experimental virtual microphone grid. In the simulations, the -10 dB estimation zone includes a unified zone centred at the origin, covering a circular area with a radius of about $0.3\lambda_s$ with ellipsoid extensions towards the array element positions and distinct zones centred at the element positions, extending about $0.4\lambda_s$ outwards. The difference between the simulated and experimental results in this case can again be related to the different optimal observation filter calculations required in each case, as discussed in relation to [Figure 7](#).

Overall, the presented results agree well with previous studies on microphone arrays in three-dimensional diffuse fields^{[10,27,28](#)}, demonstrating an extension of the -10 dB estimation zones along the array axes. This suggests that sub-array pairs should be, in general, orientated towards the virtual microphone positions to achieve an estimation zone extension along that direction. Moreover, spherical sub-array configurations would provide angularly uniform pressure gradient information, increasing spatial coherence radially around the array elements resulting in uniform estimation zone extension in three dimensions.

VI. CONCLUSIONS

This paper has presented an investigation into the use of multi-microphone configurations to estimate a stationary, random sound field at remote locations using the Remote Microphone virtual sensing technique. The performance of monitoring microphone arrays composed of nested sub-arrays is investigated with closely spaced microphone pairs used to

incorporate pressure-gradient information in the estimation process¹⁰. Using subsets of the sub-arrays, three linear and three circular configurations have been investigated.

The estimation performance for each array was initially assessed over a rectangular grid of virtual microphones over a frequency range spanning almost five octaves using measurements from an experimental implementation. In general, these results have shown that monitoring microphone array configurations with sub-arrays containing closely spaced microphone pairs outperform conventional arrays with single pressure sensors, both in terms of the level of accuracy and the spatial extent of the region over which an accurate estimation is provided. However, inclusion of closely spaced microphone pairs results in a greater variance in the estimation performance over space and will also require a higher computational complexity. The performance of linear array configurations showed significant degradation as the distance between the array and the position at which the pressure is estimated is increased. The circular configurations exhibit higher levels of performance compared to their linear counterparts, but with higher variation over frequency due to their modal behaviour. The presented results have highlighted that the microphone topology has a significant impact on the estimation performance and the spatial distribution of the estimation error, with circular configurations achieving higher estimation accuracy and larger estimation zones with fewer microphones for the considered disturbance field.

Numerical simulations were carried out to assess the performance of the different array configurations over a larger spatial grid. The behaviour of the simulated arrays was matched to the experimental realisations by applying frequency-dependent regularisation to match the condition number over frequency. To circumvent the numerical issue arising

when the number of deployed microphones exceeds the number of sources in the system, the simulations were performed with a larger array of primary disturbances compared to the experimental implementation. Despite this difference in the number of independent primary sources, the performance of the simulated systems aligned well with the experimental results. Finally, the simulations were used to investigate the performance of the considered arrays over a larger area and these results have demonstrated the extent of the performance gains that may be achieved by incorporating pressure gradient into the estimation process. In the case of the linear arrays, there is a significant increase in the -10 dB estimation zone by incorporating pressure gradient information in the y -coordinate direction, but further inclusion of pressure gradient information in the x -coordinate direction only provides modest performance gains. In the case of the circular arrays, significant extensions in the size of the -10 dB estimation zone are achieved by incorporating either a single pair of closely spaced microphones or two closely spaced microphone pairs. However, in all cases the arrays become less well-conditioned when closely spaced microphone pairs are utilised and this may limit the robustness of the array performance to practical uncertainties. Although this study has only investigated the performance of nested arrays in sound fields without scattering or reflecting objects, it has previously been demonstrated that the spatial correlation and coherence increase in the vicinity of rigid scattering bodies. This means that the -10 dB estimation zones are expected to increase in size close to the reflecting or scattering objects, with the extension depending on the shape of the reflective object and its distance from the estimation position³⁷. However, as noted variously throughout this paper, future work is required to further explore practical aspects of the considered nested microphone arrays.

ACKNOWLEDGMENTS

This work was supported by the project “IN-NOVA: Active reduction of noise transmitted into and from enclosures through encapsulated structures”, which has received funding from the European Union’s Horizon Europe programme under the Marie Skłodowska-Curie grant agreement no. 101073037 and by UK Research and Innovation under the UK government’s Horizon Europe funding agreement with grant number EP/X027767/1. Jordan Cheer was supported by the Department of Science, Innovation and Technology (DSIT) Royal Academy of Engineering under the Research Chairs and Senior Research Fellowships programme.

CONFLICT OF INTEREST

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX: OVERDETERMINED SYSTEM

When the number of microphones exceeds the number of sources in the system, the power spectral density matrix, $\mathbf{S}_{mm}$ as given in [Equation 4b](#), is singular. Using [Equation 1a](#) and [Equation 1b](#) the error in [Equation 2](#) can be expressed as

$$\boldsymbol{\epsilon} = \mathbf{d}_e - \mathbf{O}\mathbf{d}_m = (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m) \mathbf{v}. \quad (\text{A1})$$

Since we want the error to be zero for all $\mathbf{v}$, the trivial solution of $\mathbf{v} = \mathbf{0}$ is dropped. A minimum norm observation filter that sets Equation A1 to zero can be formulated by solving the following optimisation problem

$$\min \|\mathbf{O}\|_F^2, \quad \text{s.t. } \mathbf{P}_e - \mathbf{O}\mathbf{P}_m = \mathbf{0}, \quad (\text{A2})$$

where $\mathbf{0} \in \mathbb{R}^{N_e \times N_v}$ here is the zero matrix and $\|\cdot\|_F$ denotes the Frobenius norm. The problem can be solved using the method of Lagrange multipliers with the Lagrangian being³⁸

$$\mathcal{L}(\mathbf{O}, \mathbf{\Lambda}) = \text{tr} \left\{ \mathbf{O}\mathbf{O}^H + \mathcal{R} \left\{ \mathbf{\Lambda}^H (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m) \right\} \right\}, \quad (\text{A3})$$

where $\mathbf{\Lambda} \in \mathbb{C}^{N_e \times N_v}$ is a matrix whose entries are the complex Lagrange multipliers λ and $\mathcal{R}\{\cdot\}$ denotes the real part of an expression. The real part of $\mathbf{\Lambda}^H (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m)$ is taken because the Lagrangian has to be a real function.

Using the fact that for a complex number z , the sum $z + z^* = 2\mathcal{R}\{z\}$, with $[\cdot]^*$ denoting complex conjugation, to express the real part of $\mathbf{\Lambda}^H (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m)$ we get

$$\begin{aligned} \mathcal{R} \left\{ \mathbf{\Lambda}^H (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m) \right\} &= \frac{1}{2} \left[\mathbf{M}^H (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m) \right. \\ &\quad \left. + \mathbf{M}^T (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m)^* \right] \\ &= \mathbf{\Lambda}^H (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m) \\ &\quad + \mathbf{\Lambda}^T (\mathbf{P}_e^* - \mathbf{O}^* \mathbf{P}_m^*), \end{aligned} \quad (\text{A4})$$

where $\mathbf{\Lambda} = \frac{1}{2}\mathbf{M}$. Plugging Equation A4 into Equation A3 the Lagrangian is expressed as

$$\begin{aligned} \mathcal{L}(\mathbf{O}, \mathbf{\Lambda}) &= \text{tr} \left\{ \mathbf{O}\mathbf{O}^H \right\} + \text{tr} \left\{ \mathbf{\Lambda}^H (\mathbf{P}_e - \mathbf{O}\mathbf{P}_m) \right\} \\ &\quad + \text{tr} \left\{ \mathbf{\Lambda}^T (\mathbf{P}_e^* - \mathbf{O}^* \mathbf{P}_m^*) \right\}. \end{aligned} \quad (\text{A5})$$

546 Treating $\mathbf{O}$, $\mathbf{O}^*$, $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^*$ as independent variables and taking the partial derivatives of the
 547 Lagrangian we get³⁹

$$\frac{\partial \mathcal{L}}{\partial \mathbf{O}} = \mathbf{O}^* - (\mathbf{\Lambda} \mathbf{P}_m^H)^* = (\mathbf{O} - \mathbf{\Lambda} \mathbf{P}_m^H)^* \quad (\text{A6})$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{O}^*} = \mathbf{O} - \mathbf{\Lambda} \mathbf{P}_m^H \quad (\text{A7})$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{\Lambda}} = \mathbf{P}_e^* - \mathbf{O}^* \mathbf{P}_m^* = (\mathbf{P}_e - \mathbf{O} \mathbf{P}_m)^* \quad (\text{A8})$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{\Lambda}^*} = \mathbf{P}_e - \mathbf{O} \mathbf{P}_m, \quad (\text{A9})$$

548 where the dependency of the Lagrangian on the variables has been dropped for notational
 549 convenience. Next, summing Equation A6 with Equation A7 and using the relation of the
 550 sum of a complex variable with its conjugate used before, we can form the partial derivative
 551 with respect to the real part of $\mathbf{O}$ like³⁹

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathcal{R}\{\mathbf{O}\}} &= \frac{\partial \mathcal{L}}{\partial \mathbf{O}} + \frac{\partial \mathcal{L}}{\partial \mathbf{O}^*} \\ &= (\mathbf{O} - \mathbf{\Lambda} \mathbf{P}_m^H)^* + (\mathbf{O} - \mathbf{\Lambda} \mathbf{P}_m^H) \\ &= 2\mathcal{R}\{(\mathbf{O} - \mathbf{\Lambda} \mathbf{P}_m^H)\}. \end{aligned} \quad (\text{A10})$$

552 Similarly, by subtracting Equation A7 from Equation A6 and using the fact that $z - z^* =$
 553 $2\mathcal{I}\{z\}$, where $\mathcal{I}$ denotes the imaginary part of an expression, we can form the partial deriva-
 554 tive with respect to the imaginary part of $\mathbf{O}$. Working in the same way for the derivatives

555 with respect to $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^*$ we get

$$\frac{\partial \mathcal{L}}{\partial \mathcal{R}\{\mathbf{O}\}} = 2\mathcal{R}\{(\mathbf{O} - \mathbf{\Lambda}\mathbf{P}_m^H)\} \quad (\text{A11})$$

$$\frac{\partial \mathcal{L}}{\partial \mathcal{I}\{\mathbf{O}\}} = 2\mathcal{I}\{(\mathbf{O} - \mathbf{\Lambda}\mathbf{P}_m^H)\} \quad (\text{A12})$$

$$\frac{\partial \mathcal{L}}{\partial \mathcal{R}\{\mathbf{\Lambda}\}} = 2\mathcal{R}\{(\mathbf{P}_e - \mathbf{O}\mathbf{P}_m)\} \quad (\text{A13})$$

$$\frac{\partial \mathcal{L}}{\partial \mathcal{I}\{\mathbf{\Lambda}\}} = 2\mathcal{I}\{(\mathbf{P}_e - \mathbf{O}\mathbf{P}_m)\}. \quad (\text{A14})$$

556 Combining Equation A11 with Equation A12 and Equation A13 with Equation A14 we get

557 the complex derivatives of the Lagrangian as

$$\frac{\partial \mathcal{L}}{\partial \mathbf{O}} = \frac{\partial \mathcal{L}}{\partial \mathcal{R}\{\mathbf{O}\}} + j \frac{\partial \mathcal{L}}{\partial \mathcal{I}\{\mathbf{O}\}} = 2\mathbf{O} - 2\mathbf{\Lambda}\mathbf{P}_m^H \quad (\text{A15})$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{\Lambda}} = \frac{\partial \mathcal{L}}{\partial \mathcal{R}\{\mathbf{\Lambda}\}} + j \frac{\partial \mathcal{L}}{\partial \mathcal{I}\{\mathbf{\Lambda}\}} = 2\mathbf{P}_e - 2\mathbf{O}\mathbf{P}_m. \quad (\text{A16})$$

558 Next, setting Equation A15 equal to zero and solving for $\mathbf{O}$ gives

$$\mathbf{O} = \mathbf{\Lambda}\mathbf{P}_m^H. \quad (\text{A17})$$

559 Plugging Equation A17 into Equation A16, setting equal to zero and solving for $\mathbf{\Lambda}$ gives

$$\mathbf{\Lambda} = \mathbf{P}_e (\mathbf{P}_m^H \mathbf{P}_m)^{-1}. \quad (\text{A18})$$

560 Using Equation A18 to express $\mathbf{\Lambda}$ in Equation A17 gives the minimum norm solution for

561 the observation filter

$$\mathbf{O}_{\text{opt,u}} = \mathbf{P}_e (\mathbf{P}_m^H \mathbf{P}_m)^{-1} \mathbf{P}_m^H = \mathbf{P}_e \mathbf{P}_m^\dagger, \quad (\text{A19})$$

562 where the term $\mathbf{P}_m^\dagger = (\mathbf{P}_m^H \mathbf{P}_m)^{-1} \mathbf{P}_m^H$ is recognised as the Moore-Penrose, left inverse, of $\mathbf{P}_m$

563 and this concludes the derivation.

References

- ¹C. Hansen, S. Snyder, X. Qiu, B. Laura, and D. Moreau, *Active control of noise and vibration*, 2 ed. (CRC Press, Taylor & Francis Group, 2012), pp. 1–1537.
- ²A. J. Bullmore, P. A. Nelson, A. R. D. Curtis, and S. J. Elliott, “The active minimization of harmonic enclosed sound fields, part ii: A computer simulation,” *Journal of Sound and Vibration* **117**, 15–33 (1987).
- ³S. J. Elliott, A. R. D. Curtis, A. J. Bullmore, and P. A. Nelson, “The active minimization of harmonic enclosed sound fields, part iii: Experimental verification,” *Journal of Sound and Vibration* **117**, 35–58 (1987).
- ⁴S. J. Elliott and P. A. Nelson, “Active noise control,” *IEEE Signal Processing Magazine* **10**, 12–35 (1993) doi: [mu10.1109/79.248551](https://doi.org/10.1109/79.248551).
- ⁵B. Rafaely, S. J. Elliott, and J. Garcia-Bonito, “Broadband performance of an active headrest,” *The Journal of the Acoustical Society of America* **106**, 787–793 (1999) doi: [mu10.1121/1.427134](https://doi.org/10.1121/1.427134).
- ⁶D. Moreau, B. Cazzolato, A. Zander, and C. Petersen, “A review of virtual sensing algorithms for active noise control,” *Algorithms* **1**, 69–99 (2008) doi: [mu10.3390/a1020069](https://doi.org/10.3390/a1020069).
- ⁷J. Zhang, S. J. Elliott, and J. Cheer, “Robust performance of virtual sensing methods for active noise control,” *Mechanical Systems and Signal Processing* **152** (2021) doi: [mu10.1016/j.ymssp.2020.107453](https://doi.org/10.1016/j.ymssp.2020.107453).
- ⁸W. Jung, S. J. Elliott, and J. Cheer, “Estimation of the pressure at a listener’s ears in an active headrest system using the remote microphone technique,” *The Journal of the*

Acoustical Society of America **143**, 2858–2869 (2018) doi: [mu10.1121/1.5037363](https://doi.org/10.1121/1.5037363).

⁹W. Jung, S. J. Elliott, and J. Cheer, “Local active control of road noise inside a vehicle,”
Mechanical Systems and Signal Processing **121**, 144–157 (2019) doi: [mu10.1016/j.ymssp.
2018.11.003](https://doi.org/10.1016/j.ymssp.2018.11.003).

¹⁰D. J. Moreau, J. Ghan, B. S. Cazzolato, and A. C. Zander, “Active noise control in a pure
tone diffuse sound field using virtual sensing,” The Journal of the Acoustical Society of
America **125**, 3742–3755 (2009) doi: [mu10.1121/1.3123404](https://doi.org/10.1121/1.3123404).

¹¹B. Xu, S. D. Sommerfeldt, and T. W. Leishman, “Generalized acoustic energy density,”
The Journal of the Acoustical Society of America **130**, 1370–1380 (2011) doi: [mu10.1121/
1.3624482](https://doi.org/10.1121/1.3624482).

¹²J. W. Parkins, S. D. Sommerfeldt, and J. Tichy, “Narrowband and broadband active
control in an enclosure using the acoustic energy density,” The Journal of the Acoustical
Society of America **108**, 192–203 (2000) doi: [mu10.1121/1.429456](https://doi.org/10.1121/1.429456).

¹³B. Xu and S. D. Sommerfeldt, “Generalized acoustic energy density based active noise
control in single frequency diffuse sound fields,” The Journal of the Acoustical Society of
America **136**, 1112–1119 (2014) doi: [mu10.1121/1.4892754](https://doi.org/10.1121/1.4892754).

¹⁴D. J. Moreau, B. S. Cazzolato, and A. C. Zander, “Active noise control at a virtual
acoustic energy density sensor in a three-dimensional sound field,” in *Proceedings of the
20th International Congress on Acoustics*, Sydney, Australia (August, 2010).

¹⁵Y. C. Park and S. D. Sommerfeldt, “Global attenuation of broadband noise fields using
energy density control,” The Journal of the Acoustical Society of America **101** (1997) doi:

[mu10.1121/1.417980](https://doi.org/10.1121/1.417980).

¹⁶N. Han and X. Qiu, “A study of sound intensity control for active noise barriers,” *Applied Acoustics* **68**, 1297–1306 (2007) doi: [mu10.1016/j.apacoust.2006.07.002](https://doi.org/10.1016/j.apacoust.2006.07.002).

¹⁷J. Ghan, B. S. Cazzolato, and S. D. Snyder, “Expression for the estimation of time-averaged acoustic energy density using the two-microphone method (I),” *The Journal of the Acoustical Society of America* **113**, 2404–2407 (2003) doi: [mu10.1121/1.1567273](https://doi.org/10.1121/1.1567273).

¹⁸B. S. Cazzolato and J. Ghan, “Frequency domain expressions for the estimation of time-averaged acoustic energy density,” *The Journal of the Acoustical Society of America* **117**, 3750–3756 (2005) doi: [mu10.1121/1.1904505](https://doi.org/10.1121/1.1904505).

¹⁹B. S. Cazzolato and C. H. Hansen, “Errors in the measurement of acoustic energy density in one-dimensional sound fields,” *Journal of Sound and Vibration* **236**, 801–831 (2000) doi: [mu10.1006/jsvi.2000.3002](https://doi.org/10.1006/jsvi.2000.3002).

²⁰B. S. Cazzolato and C. H. Hansen, “Errors arising from three-dimensional energy density sensing in one-dimensional sound fields,” *Journal of Sound and Vibration* **236**, 375–400 (2000) doi: [mu10.1006/jsvi.1999.2992](https://doi.org/10.1006/jsvi.1999.2992).

²¹J. W. Parkins, S. D. Sommerfeldt, and J. Tichy, “Error analysis of a practical energy density sensor,” *The Journal of the Acoustical Society of America* **108** (2000) doi: [mu10.1121/1.429458](https://doi.org/10.1121/1.429458).

²²S. J. Elliott and J. Cheer, “Modeling local active sound control with remote sensors in spatially random pressure fields,” *The Journal of the Acoustical Society of America* **137**, 1936–1946 (2015) doi: [mu10.1121/1.4916274](https://doi.org/10.1121/1.4916274).

²³S. J. Elliott, J. Cheer, J. W. Choi, and Y. Kim, “Robustness and regularization of personal audio systems,” *IEEE Transactions on Audio, Speech and Language Processing* **20**, 2123–2133 (2012) doi: [mu10.1109/TASL.2012.2197613](https://doi.org/10.1109/TASL.2012.2197613).

²⁴A. Novak, P. Lotton, and L. Simon, “Synchronized swept-sine: Theory, application and implementation,” *AES: Journal of the Audio Engineering Society* **63**, 786–798 (2015) doi: [mu10.17743/jaes.2015.0071](https://doi.org/10.17743/jaes.2015.0071).

²⁵X. Zheng, Z. Jia, B. Wan, M. Zeng, and Y. Qiu, “A study on hybrid active noise control system combined with remote microphone technique,” *Applied Acoustics* **205** (2023) doi: [mu10.1016/j.apacoust.2023.109296](https://doi.org/10.1016/j.apacoust.2023.109296).

²⁶S. J. Elliott, J. Cheer, and W. Jung, “The effect of remote microphone technique and head-tracking on local active sound control,” in *Proceedings of the 23rd International Congress on Sound & Vibration*, Athens, Greece (July, 2016).

²⁷S. Elliott and J. Garcia-Bonito, “Active cancellation of pressure and pressure gradient in a diffuse sound field,” *Journal of Sound and Vibration* **186**, 696–704 (1995) doi: [mu10.1006/jsvi.1995.0482](https://doi.org/10.1006/jsvi.1995.0482).

²⁸A. Kappis, J. Cheer, and J. Zhang, “Remote sensing using multi-microphone configurations for local active noise control applications,” in *Proceedings of the 30th International Congress on Sound and Vibration*, Amsterdam, Netherlands (July 08–11, 2024).

²⁹P. Zhang, S. Wang, H. Duan, J. Tao, H. Zou, and X. Qiu, “A study on coherence between virtual signal and physical signals in remote acoustic sensing,” *The Journal of the Acoustical Society of America* **152**, 2840–2848 (2022) doi: [mu10.1121/10.0015140](https://doi.org/10.1121/10.0015140).

³⁰W. Jung, “Mid-frequency local active control of road noise,” Ph.D. thesis, University of Southampton, 2018.

³¹S. A. Verburg, F. Elvander, T. Van Waterschoot, and E. Fernandez-Grande, “Optimal sensor placement for the spatial reconstruction of sound fields,” *EURASIP Journal on Audio, Speech, and Music Processing* **2024**(1), 41 (2024) doi: [mu10.1186/s13636-024-00364-4](https://doi.org/10.1186/s13636-024-00364-4).

³²H. Teutsch, “Wavefield decomposition using microphone arrays and its application to acoustic scene analysis,” Ph.D. thesis, Friedrich-Alexander University, 2006.

³³I. Balmages and B. Rafaely, “Open-sphere designs for spherical microphone arrays,” *IEEE Transactions on Audio, Speech and Language Processing* **15**, 727–732 (2007) doi: [mu10.1109/TASL.2006.881671](https://doi.org/10.1109/TASL.2006.881671).

³⁴A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications, Third Edition*, 3rd ed. (Springer, 2019).

³⁵R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, “Intrinsic Limits of Dimensionality and Richness in Random Multipath Fields,” *IEEE Transactions on Signal Processing* **55**(6), 2542–2556 (2007) doi: [mu10.1109/TSP.2007.893738](https://doi.org/10.1109/TSP.2007.893738).

³⁶B. Rafaely, “Zones of quiet in a broadband diffuse sound field,” *The Journal of the Acoustical Society of America* **110**, 296–302 (2001) doi: [mu10.1121/1.1377632](https://doi.org/10.1121/1.1377632).

³⁷J. Garcia-Bonito, S. Elliott, and M. Bonilha, “Active Cancellation of Pressure at a Point in a Pure Tone Diffracted Diffuse Field,” *Journal of Sound and Vibration* **201**(1), 43–65 (1997) doi: [mu10.1006/jsvi.1996.0742](https://doi.org/10.1006/jsvi.1996.0742).

³⁸D. G. Luenberger, *Optimization by vector space methods*, Series in decision and control (J.
Wiley, New York, 1969).

³⁹D. Brandwood, “A complex gradient operator and its application in adaptive array theory,”
IEE Proceedings F Communications, Radar and Signal Processing **130**(1), 11 (1983) doi:
[mu10.1049/ip-f-1.1983.0003](https://doi.org/10.1049/ip-f-1.1983.0003).