30th International Congress on Sound and Vibration

ICSV30 AMSTERDAM 8-11 JULY 2024

REMOTE SENSING USING MULTI-MICROPHONE CONFIGURATIONS FOR LOCAL ACTIVE NOISE CONTROL APPLICATIONS

Achilles Kappis*, Jordan Cheer, Jihui (Aimee) Zhang

Institute of Sound and Vibration Research, University of Southampton (ISVR), University of Southampton, Southampton, SO17 1BJ, United Kingdom.

E-mail: A.Kappis@soton.ac.uk

In local active noise control applications it is often impractical or even impossible to place sensors at the locations where the sound field is to be controlled. To overcome this limitation, virtual sensing methods are used to project the zones of quiet away from the physical sensors to the position of interest. This study investigates the use of multi-microphone configurations for the estimation of a wide-sense stationary, single-frequency, diffuse sound field with the Remote Microphone Technique. Linear and circular arrays of three-dimensional microphone configurations were evaluated through numerical simulations, and the findings indicate that multi-microphone setups lead to increased estimation accuracy over larger areas compared to conventional single-microphone arrays. Additionally, the effect of regularisation on the estimation error and spatial extent of the estimation zones was also studied and is shown to significantly affect the performance of dense microphone configurations.

Keywords: Active noise control, Virtual sensing, Remote microphone technique, Microphone arrays

1. Introduction

Traditional active noise control (ANC) algorithms assume that error microphones are located at the physical positions where control is needed. However, practical systems such as the headrest application [1] often do not meet this requirement. To address this limitation, virtual sensing (VS) algorithms have been developed to move the controlled position away from the monitoring sensors. One such method is the Remote Microphone Technique (RMT) [2], in which an observation filter is applied to measured signals to estimate the sound field at the location of virtual microphones.

Conventionally, omnidirectional microphones have been used in the implementation of ANC and VS systems. However, past research [3, 4] shows that total acoustic energy sensing can yield superior results compared to pressure estimation. The total acoustic energy is the sum of potential and kinetic energy, which are in turn proportional to the square of the pressure and pressure gradient. Controlling these two quantities is equivalent to controlling the acoustic energy [5]. Elliott and Garcia-Bonito [6] have demonstrated that controlling the pressure at two closely spaced locations in a diffuse field provides similar results to controlling the pressure and pressure gradient at a single point.

This study utilises the approach described in [6] to incorporate the pressure gradient to the sound field estimation within the RMT framework. Expanding upon prior research on estimation with linear and circular microphone arrays [1, 7], this work focuses on the estimation of a single-frequency diffuse

sound field using multi-microphone configurations. The estimation performance is assessed and compared to estimation with single-microphone configurations. Furthermore, the study explores the effect of regularisation on estimation performance and the spatial extent of the areas where accurate estimation is achieved.

2. Problem formulation

The block diagram of the system used for the evaluation of the estimation performance of the microphone configurations is shown in *Fig.*1. The primary disturbances are considered to be random, and all signals are characterised by their spectral densities. The primary disturbance generated by N_v sources with strengths $\boldsymbol{v} = [v_1, v_2, \ldots, v_{N_\mathrm{v}}]^T$ pass through the transfer functions $\boldsymbol{P}_\mathrm{m}$ and $\boldsymbol{P}_\mathrm{e}$ to generate the disturbance vectors $\boldsymbol{d}_\mathrm{m} = \left[d_{\mathrm{m}_1}, d_{\mathrm{m}_2}, \ldots, d_{\mathrm{m}_{N_\mathrm{m}}}\right]^T$ and $\boldsymbol{d}_\mathrm{e} = \left[d_{\mathrm{e}_1}, d_{\mathrm{e}_2}, \ldots, d_{\mathrm{e}_{N_\mathrm{e}}}\right]^T$ at N_m monitoring and N_e virtual microphones, respectively. The observation filters $\hat{\boldsymbol{O}}$ are applied to the signals captured at the monitoring microphones to produce the estimated disturbance at the virtual positions, where $[\hat{\cdot}]$ denotes an estimated value. The difference between the true and estimated value at the virtual microphone is the estimation error $\boldsymbol{e}$ given by

$$\boldsymbol{e} = \boldsymbol{d}_\mathrm{e} - \hat{\boldsymbol{d}}_\mathrm{e} = \boldsymbol{d}_\mathrm{e} - \hat{\boldsymbol{O}}\boldsymbol{d}_\mathrm{m} = \boldsymbol{P}_\mathrm{e}\boldsymbol{v} - \hat{\boldsymbol{O}}\boldsymbol{P}_\mathrm{m}\boldsymbol{v}. \tag{1}$$

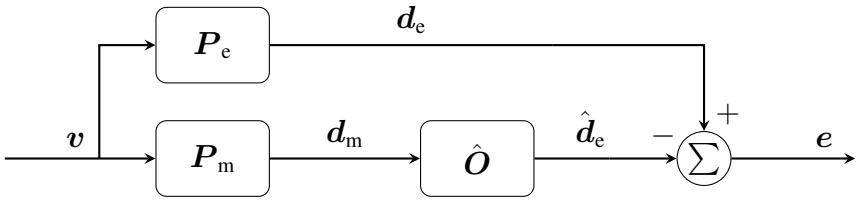

Figure 1: Block diagram of the sound field estimation problem.

The optimal filter is calculated in the least squares sense by minimising the function [8]

$$J_2 = \mathrm{E}\left[\boldsymbol{e}^H\boldsymbol{e}\right] = \mathrm{trace}\left\{\mathrm{E}\left[\boldsymbol{e}\boldsymbol{e}^H\right]\right\} = \mathrm{trace}\left\{\boldsymbol{O}\boldsymbol{S}_\mathrm{mm}\boldsymbol{O}^H - \boldsymbol{O}\boldsymbol{S}_\mathrm{me}^H - \boldsymbol{S}_\mathrm{me}\boldsymbol{O}^H + \boldsymbol{S}_\mathrm{ee}\right\}, \tag{2}$$

where $[\cdot]^H$ denotes Hermitian conjugation and $\boldsymbol{S}_\mathrm{me} = \mathrm{E}\left[\boldsymbol{d}_\mathrm{e}\boldsymbol{d}_\mathrm{m}^H\right]$ and $\boldsymbol{S}_\mathrm{ee} = \mathrm{E}\left[\boldsymbol{d}_\mathrm{e}\boldsymbol{d}_\mathrm{e}^H\right]$ are the cross spectral densities of the monitoring and virtual microphone signals respectively. The optimal solution of *eq.*(2) with respect to the observation filter is given by [8]

$$\hat{\boldsymbol{O}}_\mathrm{opt} = \boldsymbol{S}_\mathrm{me}\left(\boldsymbol{S}_{mm} + \beta\boldsymbol{I}\right)^{-1} = \boldsymbol{P}_\mathrm{e}\boldsymbol{v}\boldsymbol{v}^H\boldsymbol{P}_\mathrm{m}^H\left(\boldsymbol{P}_\mathrm{m}\boldsymbol{v}\boldsymbol{v}^H\boldsymbol{P}_\mathrm{m}^H\right)^{-1} = \boldsymbol{P}_\mathrm{e}\boldsymbol{S}_\mathrm{vv}\boldsymbol{P}_\mathrm{m}^H\left(\boldsymbol{P}_\mathrm{m}\boldsymbol{S}_\mathrm{vv}\boldsymbol{P}_\mathrm{m}^H\right)^{-1}, \tag{3}$$

with $\boldsymbol{I}$ being the identity matrix of appropriate dimensions, β a real scalar regularisation factor, and $\boldsymbol{S}_\mathrm{vv} = \mathrm{E}\left[\boldsymbol{v}\boldsymbol{v}^H\right]$ the power spectral density of the primary disturbances. The term $\beta\boldsymbol{I}$ imposes a penalty to the maximum attainable values of $\hat{\boldsymbol{O}}_\mathrm{opt}$, reducing its sensitivity to perturbations, and is tantamount to introducing uncorrelated noise to the monitoring signals [9].

3. Microphone Array Geometries

To facilitate comparison with previously reported results in the literature, a uniform linear array (ULA) and a uniform circular array (UCA), have been implemented to estimate a simulated diffuse sound

field. Two different multi-microphone configurations have been used at each measurement position in the arrays resulting in the setups shown in *Fig.*2. Specifically, the first configuration is comprised of three pairs of omnidirectional microphones, with each orientated along one of the Cartesian axes and all having an inter-element spacing $d = 0.05\,\text{m}$. The second configuration features microphones located on the vertices of a normal tetrahedron, with edge length equal to d. The term *sub-array* will be used herein to refer to the closely clustered arrays of microphones, which then form the full linear and circular arrays.

The distance between successive sub-array centres in the ULA has been set to $L_{\text{ULA}} = 0.1\,\text{m}$. To ensure a symmetric configuration with tetrahedral sub-arrays, half of them have been mirrored with respect to the $y - z$ plane. In addition, the three-axis sub-arrays have been aligned to the global coordinate system axes. The UCA radius has been chosen to be $L_{\text{UCA}} = 0.4\,\text{m}$. The three-axis sub-arrays have been rotated so that one of their axes is oriented radially. Similarly, the tetrahedral sub-arrays have been rotated to ensure that the microphone aligned with the vector parallel to the radial direction passing through the microphone and the sub-array centre lies outside of the array's circumference. All arrays have been placed with their centres at the origin of the global coordinate system.

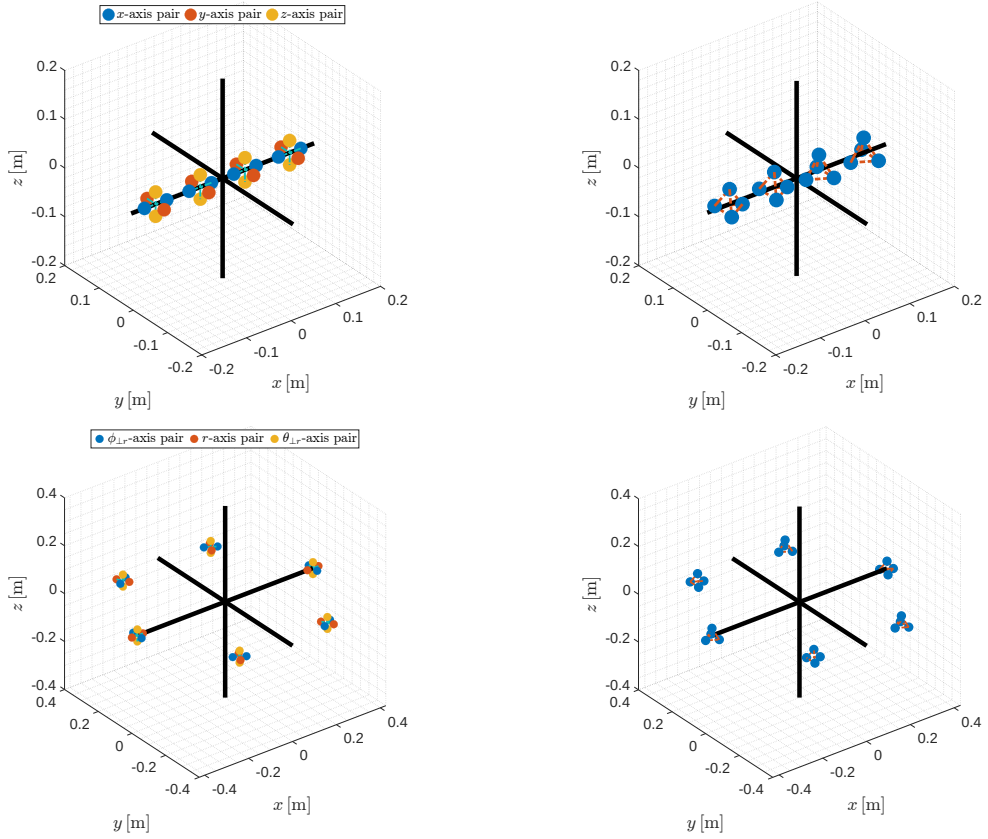

Figure 2: Multi-microphone arrays used in this work formed from two different sub-array clusters: the two upper plots show the ULA configurations and the two lower plots show the UCA configurations. The distance between sub-arrays in the ULA is $L_{\text{ULA}} = 0.1\,\text{m}$ and the UCA radius is $L_{\text{UCA}} = 0.4\,\text{m}$. The inter-element distance between each pair in the three-pair sub-array is $d = 0.05\,\text{m}$. The microphones in the tetrahedral sub-arrays are placed on the vertices of a normal tetrahedron with edge length equal to d.

A diffuse field is approximated using 408 monopole sources uniformly distributed on a sphere with radius $r_{\text{s}} = 3\,\text{m}$, centred at the origin. The sources are assumed to be driven with uncorrelated random signals and have unity strength so that $\boldsymbol{S}_{\text{vv}}$ is an identity matrix. The sound field is estimated on an 85×85 square grid of virtual microphones with side length $L_{\text{e}} = 2\,\text{m}$, centred at the origin. The normalised mean

square error level is used to assess the performance at each virtual microphone position, calculated as

$$L_\epsilon = 10 \log_{10} \left(\frac{S_{\epsilon\epsilon}}{S_{\text{ee}}} \right), \tag{4}$$

where $S_{\epsilon\epsilon} = \mathrm{E}\left[\boldsymbol{e}^H \boldsymbol{e}\right]$ is the power spectral density of the estimation error given in *eq.*(1). The area where $L_\epsilon \leq -10\,\text{dB}$ has been chosen in this work to denote the *estimation zone*. This has been chosen to be consistent with the *zone of quiet* widely adopted in the literature to be the area within which $10\,\text{dB}$ of attenuation is achieved [8].

4. Simulation results

4.1 Estimation zones

The estimation zones achieved using the ULA setups are shown in *Fig.*3 for virtual microphones on the $x-y$ and $y-z$ planes for frequencies $f = 250\,\text{Hz}$, $500\,\text{Hz}$ and $1\,\text{kHz}$. In addition to the arrays shown in *Fig.*2, two further configurations have also been used: one comprising of only the y-axis pairs and one using pressure microphones at each sub-array position.

As depicted in *Fig.*3, the alignment of microphones along a certain direction considerably increases the size of the estimation zone in that direction. For example, the single-pair sub-array configuration zone extends along the axis of the microphone pairs, but in other directions, it is similar to that of the pressure microphone array. Extension is achieved along the axes connecting all microphones in the configuration, which is especially evident with the three-axis and tetrahedral setups. The larger number of microphones in the former provides higher order pressure gradient estimates, yielding better performance further from the array, as is visible along the axis of the ULA. However, the single-axis setup does not provide any pressure gradient information in the x- and z-directions, so it has no advantage over the pressure array along those axes. The configuration of the microphones in the sub-array topologies of the three-axis and tetrahedral setups are similar on the $y-z$ plane and the estimation zone matches very well.

The estimation zones for the UCA setups are depicted in *Fig.*4 and show similar trends to the ULA examples. The estimation using single microphones at the sub-array positions and with an array comprising of sub-arrays using microphone pairs only along the radial direction are also shown. The estimation zone extension is predominantly along the axes of the microphone pairs. At low frequencies, where the distance between the sub-array positions is considerably shorter than the wavelength, the estimation zone with the pressure microphone array covers almost the entirety of the space interior to the array. However, when the condition $d \leq 0.4\lambda$ is not satisfied, the zone closely resembles that corresponding to single pressure microphones at the measurement positions [6]. Similar behaviour is evident for the other arrays, where, at high frequencies, the shape and size of the estimation zone reduces to that of the individual sub-arrays concentrated around the monitoring positions.

The reduction of the estimation zone extent with increasing frequency is a common feature for all setups. These findings are consistent with those discussed in [6], where it is proposed that the estimation accuracy depends on the spatial correlation between two positions which, in a diffuse field, deteriorates with increasing frequency and distance.

4.2 Estimation robustness

In addition to the maximum performance, it is important to consider the robustness of remote sensing strategies. The condition number κ of the cross spectral density matrix of the monitoring microphones has been used in this work to assess the sensitivity of the setups to perturbations in the measured signals

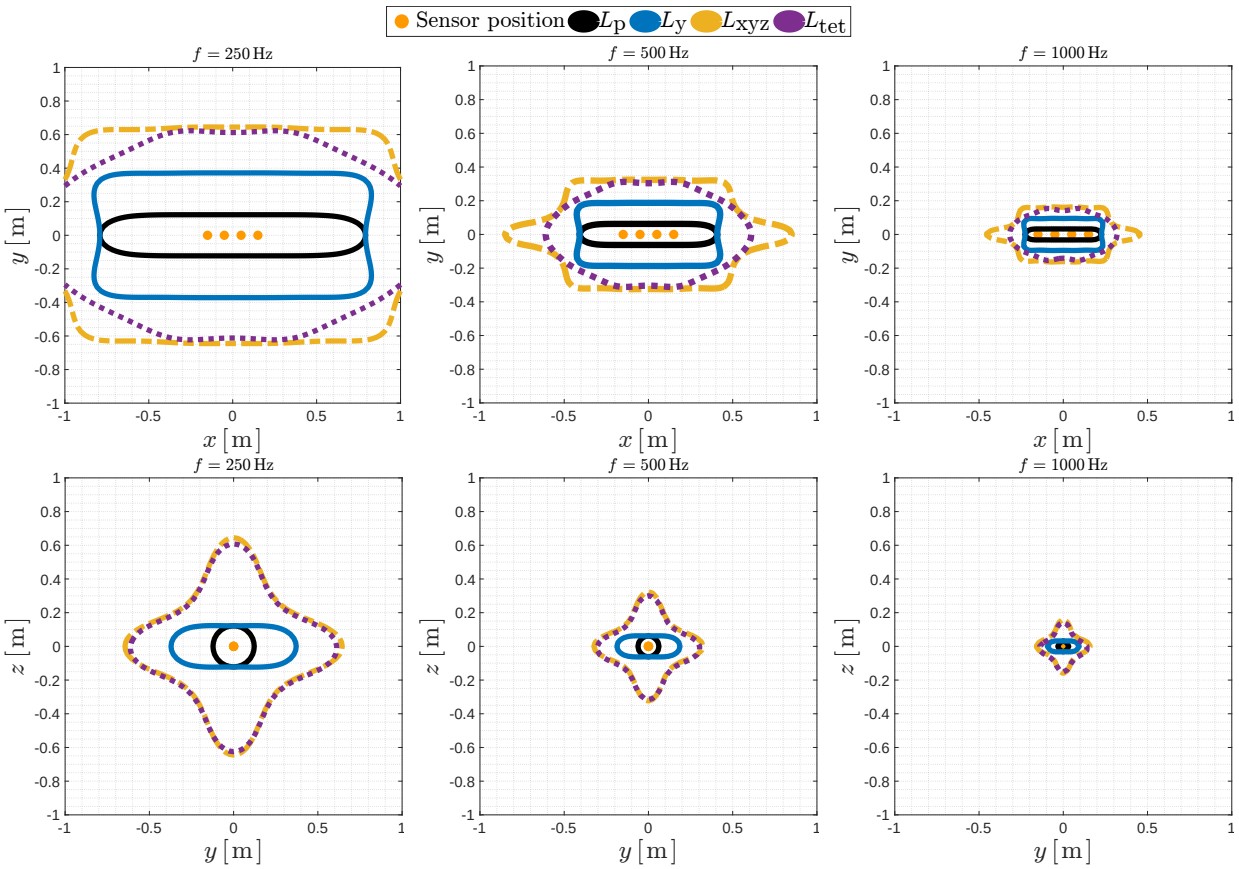

Figure 3: Estimation zones achieved using the ULA setups for frequencies $f = 250$ Hz, 500 Hz and 1 kHz on the $x-y$ and $y-z$ planes. L_{p} denotes the estimation zone achieved with single microphones at the sub-arrays positions, L_{y} those with sub-arrays comprised of only the y-axis pairs, L_{xyz} those of the sub-arrays with three-axis pairs and L_{tet} those with tetrahedral sub-arrays.

[10]. The condition number and spatially averaged error level over the estimation area on the $x-y$ plane are plotted against the regularisation factor β for a frequency of 500 Hz in *Fig.*5. The results indicate that configurations with a higher density of microphones have significantly larger condition numbers when the regularisation factor is less than 10^4, as expected due to the similarities between the responses of closely located microphones. Furthermore, low regularisation values affect the arrays minimally, with a performance plateau evident. However, the performance and condition of the array start to decline for higher levels of regularisation. The proximity of the microphones in the array has a pronounced effect on the value of the regularisation number for which the performance starts to decline. In particular, the ULA is found to be more sensitive to perturbations due to its higher microphone density compared to the UCA.

The effect of regularisation on the estimation can be seen in *Fig.*6, where the zones corresponding to the regularised arrays are shown at a frequency of $f = 500$ Hz. Appropriate regularisation values have been applied to ensure that all arrays have a condition number of $\kappa = 10^3$, except for the pressure UCA, which consistently exhibits low conditioning irrespective of regularisation. Compared to *Fig.*3 and *Fig.*4, the estimation zones have been significantly reduced, and all arrays comprising of multi-microphone sub-arrays exhibit similar performance. Although the extent of the estimation zones in *Fig.*6 is similar for most configurations, as shown in *Fig.*5, the spatially averaged error for the multi-microphone sub-array configurations is shown to be lower for sub-arrays with a higher number of microphones. This

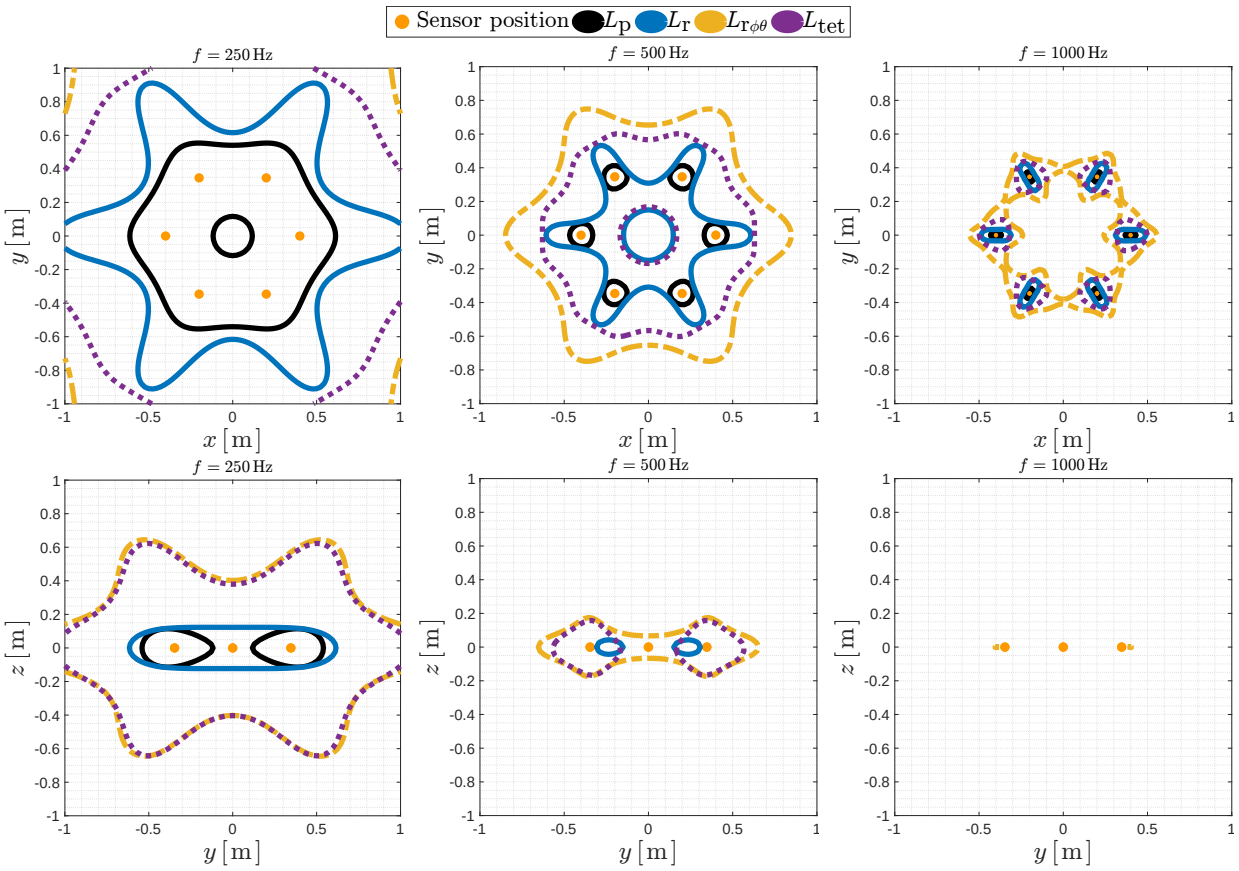

Figure 4: Estimation zones achieved using the UCA setups for frequencies $f = 250$ Hz, 500 Hz and 1 kHz on the $x-y$ and $y-z$ planes. L_{p} denotes the estimation zone achieved with single microphones at the sub-array positions, L_{r} those with sub-arrays comprised of only the radially aligned pairs, $L_{\mathrm{r}\phi\theta}$ those of the sub-arrays with three-axis pairs and L_{tet} those with tetrahedral sub-arrays.

reduction in average error can be attributed to the increase in performance inside the estimation zone when increasing the number of microphones per sub-array.

The impact of regularisation on the estimation error at frequencies ranging from 50 Hz to 1.25 kHz is shown in *Fig.*7. The estimation is performed at a virtual microphone located at $(0.15\,\mathrm{m}, 0.2\,\mathrm{m}, 0\,\mathrm{m})$. The performance deteriorates throughout the spectrum after regularisation is applied, with denser configurations being affected more severely. This severity is reflected both in the sub-array configurations as well as the density of the array as a whole. The pressure arrays seem to be affected only at low frequencies with minimal impact at mid and high frequencies. While the impact of regularisation is more severe for more compact configurations, these same configurations consistently provide superior estimation even after regularisation, at the expense of a larger number of physical microphones.

5. Conclusions

The performance of multi-microphone sub-arrays that incorporate pressure gradient information has been evaluated in the context of the RMT. A ULA and a UCA have been employed to estimate a diffuse sound field. The results suggest that the area for which the mean square estimation error is less than -10 dB extends along the direction of the pressure gradient estimate. Furthermore, utilising sub-arrays that can provide pressure differential estimates along all three Cartesian coordinate axes can extend the

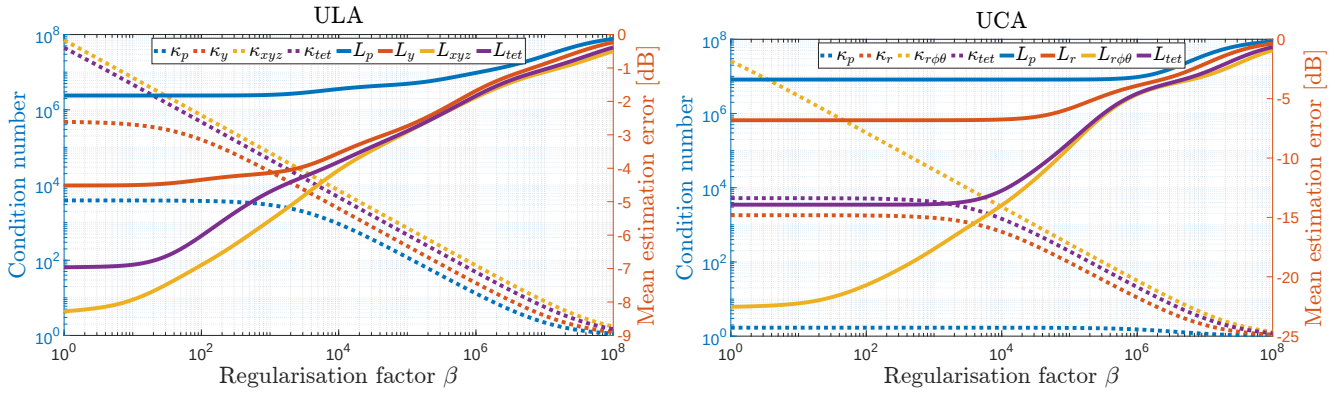

Figure 5: Condition number and spatially averaged mean square estimation error level achieved with the ULA and UCA configurations, at a frequency of 500 Hz.

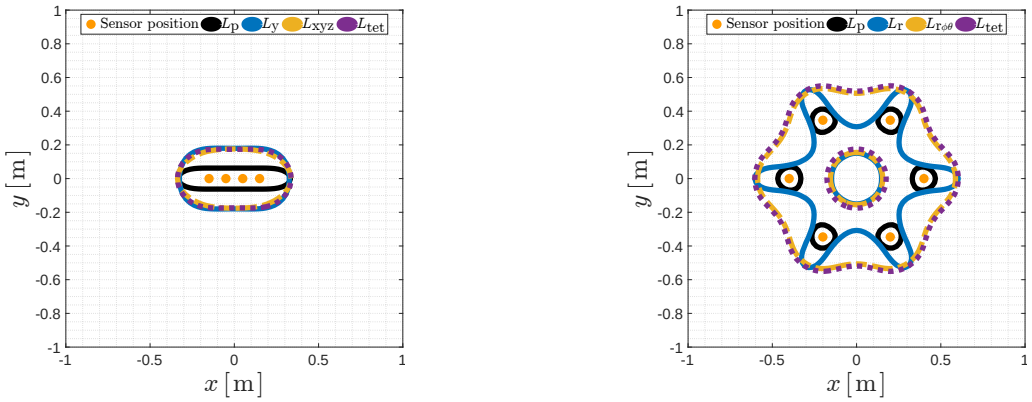

Figure 6: Estimation zones of the regularised ULA and UCA setups. The frequency of excitation is $f = 500$ Hz and all arrays have condition number $\kappa = 10^3$ except for the pressure UCA, which constantly exhibits a lower κ value.

estimation zones in the three-dimensional space.

Increasing the number of microphones in a setup enhances estimation performance, albeit at the expense of sensitivity to perturbations in the measured signals. Regularisation can alleviate the problem while affecting the estimation performance, with denser sub-arrays and setups being impacted more severely. Arrays with sub-arrays comprised of a larger number of microphones maintain superior performance even after regularisation. Increasing the distance between measurement positions can significantly improve the robustness of the observation filters. Using appropriately spaced multi-microphone sub-arrays the estimation performance can be enhanced without significant impact on sensitivity to uncertainties, which can be further mitigated through regularisation.

6. Acknowledgments

The 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 guarantee [grant number EP/X027767/1].

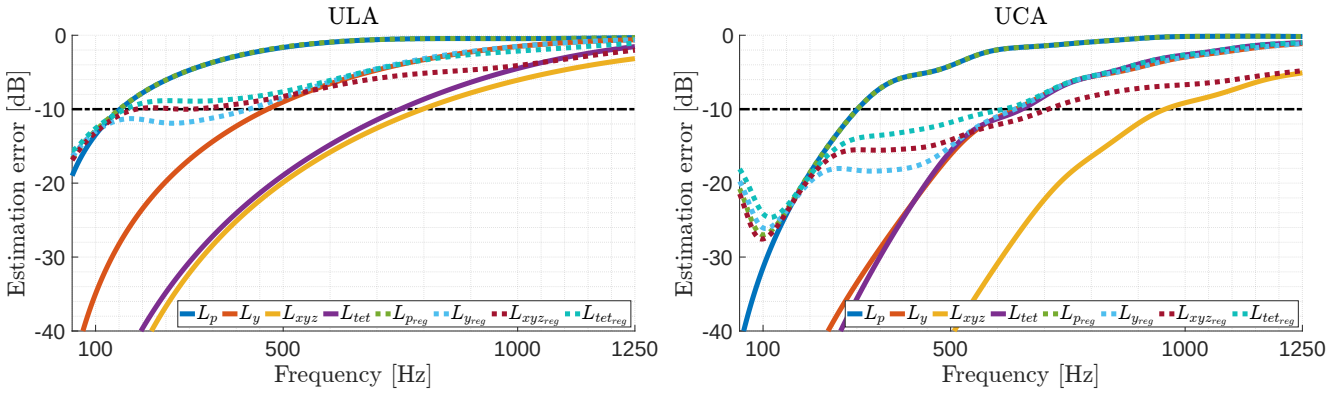

Figure 7: Mean squared estimation error level of the multi-microphone arrays, with and without regularisation, for a virtual microphone located at $(0.15\,\mathrm{m}, 0.2\,\mathrm{m}, 0\,\mathrm{m})$ for frequencies $f \in [50\,\mathrm{Hz}, 1.25\,\mathrm{kHz}]$. All arrays are regularised as in *Fig*.6.

REFERENCES

1. Jung, W., Elliott, S. J. and Cheer, J. 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).

2. Moreau, D., Cazzolato, B., Zander, A. and Petersen, C. A review of virtual sensing algorithms for active noise control, *Algorithms*, **1**, 69–99, (2008).

3. Xu, B. and Sommerfeldt, S. D. 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).

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

5. Moreau, D. J., Ghan, J., Cazzolato, B. S. and Zander, A. C. 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).

6. Elliott, S. and Garcia-Bonito, J. Active cancellation of pressure and pressure gradient in a diffuse sound field, *Journal of Sound and Vibration*, **186**, (1995).

7. Zheng, X., Jia, Z., Wan, B., Zeng, M. and Qiu, Y. A study on hybrid active noise control system combined with remote microphone technique, *Applied Acoustics*, **205**, (2023).

8. Elliott, S. J. and Cheer, J. 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).

9. Elliott, S. J., Cheer, J., Choi, J. W. and Kim, Y. Robustness and regularization of personal audio systems, *IEEE Transactions on Audio, Speech and Language Processing*, **20**, 2123–2133, (2012).

10. Zhang, J., Elliott, S. J. and Cheer, J. Robust performance of virtual sensing methods for active noise control, *Mechanical Systems and Signal Processing*, **152**, (2021).