



FORUM ACUSTICUM EURONOISE 2025

TOPOLOGY OPTIMISATION OF MICROPHONE ARRAYS FOR REMOTE MICROPHONE VIRTUAL SENSING IN DIFFUSE SOUND FIELDS

Achilles Kappis*

Jordan Cheer

Institute of Sound and Vibration Research, University of Southampton, Southampton, UK

ABSTRACT

Past research has shown that virtual sensing techniques can enhance the performance of active noise control systems by projecting the control points towards remote locations of interest. However, accurate sound field estimation using virtual sensing is critical to the performance of such active control systems and depends on both signal processing and the physical microphone array. Typically, microphone placement is determined by practical considerations such as convenience, spatial constraints and cost, resulting in limited exploration of optimal microphone positioning. The current study employs a genetic algorithm to identify optimal microphone array configurations for accurate estimation of the pressure in a diffuse sound field, utilising the Remote Microphone Technique. The optimality criterion is defined as the estimation performance or the robustness of the derived topologies to practical uncertainties. The resulting optimal configurations are evaluated against a conventional uniform linear microphone array, which consist of sub-array elements capable of utilising both pressure and pressure gradient information for enhanced estimation accuracy.

Keywords: *Virtual sensing, Remote Microphone Technique, Microphone arrays*

1. INTRODUCTION

Virtual Sensing (VS) has been effectively integrated in to local Active Noise Control (ANC) systems such as the ac-

tive headrest and has been shown to improve control performance [1, 2]. However, this is reliant on accurate estimation of the sound field at the remote location of interest [1, 3]. In general, VS studies have focused on the use of physical microphone arrays with rather straightforward geometries, such as linear or circular arrays [4, 5]. Additionally, in many practical applications, microphone placement is dictated by constraints in the system such as limited space availability and convenience. Thus, limited research has been undertaken investigating more complex or flexible microphone positioning that is solely based on providing optimal estimation performance. The problem of finding the optimal combination of sensor positions, in the ANC context, has been investigated in the past [6] with the use of Genetic Algorithms (GAs) providing promising results [5]. However, in most of the studies presented in the literature the optimal configuration was generated by selecting the combination of microphone positions out of a predefined set. The current work constitutes a preliminary study to investigate the suitability of using a GA to calculate the optimal microphone configurations from a continuous set of available positions. The Remote Microphone Technique (RMT) is used as the VS method employed for the estimation. The estimation error over a grid of virtual microphones and the robustness to uncertainties are used as performance metrics for the calculated microphone configurations and the solutions are compared to a uniform linear array consisting of sub-arrays capable of incorporating pressure and pressure gradient for the estimation of the pressure at the virtual microphones [7–9].

*Corresponding author: A.Kappis@soton.ac.uk.

Copyright: ©2025 A. Kappis et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.





FORUM ACUSTICUM EURONOISE 2025

2. REMOTE MICROPHONE TECHNIQUE

2.1 Virtual sensing formulation

The block diagram of a generalised virtual sensing system using the RMT is shown in Fig. 1. The sound field is generated by N_v primary sources with complex strengths $\mathbf{v} = [v_1, v_2, \dots, v_{N_v}]^T$, with $[\cdot]^T$ denoting transposition. The source strengths correspond to realisations of uncorrelated wide sense stationary random processes. The signals at the monitoring microphones, $\mathbf{d}_m = [m_1, m_2, \dots, m_{N_m}]^T$ and at the virtual microphones, $\mathbf{d}_\epsilon = [\epsilon_1, \epsilon_2, \dots, \epsilon_{N_\epsilon}]^T$, are given by

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

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

with $\mathbf{P}_m \in \mathbb{C}^{N_m \times N_v}$ and $\mathbf{P}_\epsilon \in \mathbb{C}^{N_\epsilon \times N_v}$ denoting the fre-

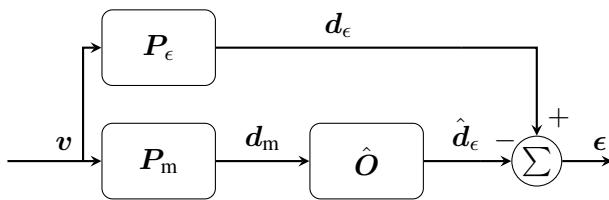


Figure 1. Block diagram of a generalised virtual sensing system. The sound field at the error microphones, \mathbf{d}_ϵ , is estimated by applying the observation filter $\hat{\mathbf{O}}$ to the signals measured at the monitoring microphones \mathbf{d}_m .

quency response functions from the sources to the monitoring and virtual microphones respectively. The sound field at the virtual microphone positions is estimated by applying the observation filter $\hat{\mathbf{O}}$ to the monitoring microphone signals. The estimation error is then given by

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

The optimal filter, in the least squares sense, is calculated by minimising the cost function [3]

$$\begin{aligned} J_2 &= \mathbb{E}[\mathbf{d}_\epsilon^H \mathbf{d}_\epsilon] \\ &= \text{tr} \left\{ \mathbf{S}_{\epsilon\epsilon} - \mathbf{S}_{\text{me}} \hat{\mathbf{O}}^H - \hat{\mathbf{O}} \mathbf{S}_{\text{me}}^H + \hat{\mathbf{O}} \mathbf{S}_{\text{mm}} \hat{\mathbf{O}}^H \right\}, \end{aligned} \quad (3)$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator, $[\cdot]^H$ denotes Hermitian conjugation, the quantities $\mathbf{S}_{\epsilon\epsilon}$ and \mathbf{S}_{mm} correspond to the monitoring and virtual microphone power spectral density matrices respectively and \mathbf{S}_{me} is the cross spectral density matrix between the monitoring and virtual microphones. These matrices can be calculated with the help of the power spectral density matrix of the source signals, \mathbf{S}_{vv} , as

$$\mathbf{S}_{\epsilon\epsilon} = \mathbb{E}[\mathbf{d}_\epsilon \mathbf{d}_\epsilon^H] = \mathbf{P}_\epsilon \mathbf{S}_{vv} \mathbf{P}_\epsilon^H \quad (4a)$$

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

$$\mathbf{S}_{\text{me}} = \mathbb{E}[\mathbf{d}_\epsilon \mathbf{d}_m^H] = \mathbf{P}_\epsilon \mathbf{S}_{vv} \mathbf{P}_m^H \quad (4c)$$

$$\mathbf{S}_{vv} = \mathbb{E}[\mathbf{v} \mathbf{v}^H]. \quad (4d)$$

The unconstrained solution to (3) is then given by [3]

$$\hat{\mathbf{O}}_{\text{opt}} = \mathbf{S}_{\text{me}} \mathbf{S}_{\text{mm}}^{-1} = \mathbf{P}_\epsilon \mathbf{S}_{vv} \mathbf{P}_m^H (\mathbf{P}_m \mathbf{S}_{vv} \mathbf{P}_m^H)^{-1}, \quad (5)$$

where $[\cdot]^{-1}$ denotes matrix inversion. The assumption of uncorrelated sources with equal strengths simplifies the source power spectral density matrix to a scaled identity matrix which can be omitted without loss of generality. This formulation enables the calculation of the filter's expected performance without generating multiple sound field realisations, making it suitable for repetitive calculations, such as those required for the optimisation with a GA used in the this study.

2.2 Performance metrics

The monitoring microphone array configurations calculated with the GA are evaluated based on two metrics, the estimation performance and the robustness to uncertainties. The metric used to evaluate the performance of the estimation is the normalised mean squared estimation error. For a single virtual microphone this is given by

$$L_\epsilon = 10 \log_{10} \left(\frac{\mathbb{E}[\epsilon \epsilon^*]}{\mathbb{E}[d_\epsilon d_\epsilon^*]} \right), \quad (6)$$

where $[\cdot]^*$ denotes complex conjugation. The estimation error gives negative values for good estimation performance with $-\infty$ denoting perfect estimation.

The robustness of the configurations is reported by the condition number of the monitoring microphone power spectral density matrix, $\kappa(\mathbf{S}_{\text{mm}})$. Higher condition numbers signify higher sensitivity to uncertainties with the theoretical maximum value being ∞ for singular matrices and the lowest is unity, for a scaled identity matrix.





FORUM ACUSTICUM EURONOISE 2025

3. GENETIC ALGORITHM

A GA is used in this work to perform a search of the microphone positions that will provide optimal estimation performance under certain constraints. The search is performed over a constrained two-dimensional continuous spatial domain with MATLAB's built-in GA algorithm. When non-linear constraints, such as those used in this study are included, the underlying algorithm used is the Augmented Lagrangian Genetic Algorithm (ALGA) [10]. A Lagrangian consisting of the fitness values of the individuals complemented by the sum of a scaled transformation of the constraints is minimised in each generation, and the original problem is approximately solved. The high level flowchart diagram describing the algorithm is shown in Fig. 2, and the steps and their associated parameters are described below.

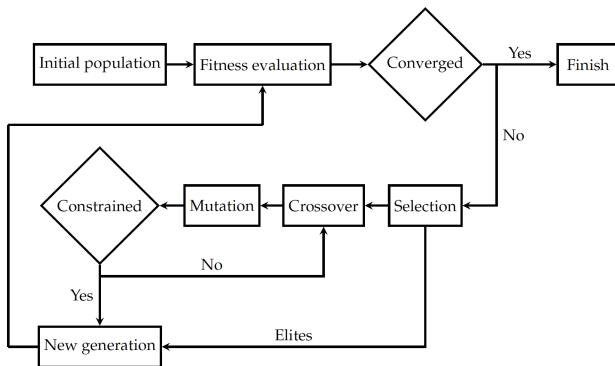


Figure 2. Flowchart diagram describing the genetic algorithm used in the current study.

Population and coding scheme The parameters to be optimised are the positions of the monitoring microphones. The number of microphones is fixed to eight and their x and y coordinates are interleaved in a vector of the form $\mathbf{x} = [x_1, y_1, x_2, y_2, \dots, x_{N_m}, y_{N_m}]$ which constitutes one individual in the population, with one generation consisting of 100 individuals. The population of the first generation is drawn at random from a uniform distribution scaled and shifted to satisfy any imposed constraints.

Fitness function The fitness function defines the performance of each individual. In the current study, two distinct fitness functions are used. The function evaluating the performance of the configurations is defined as the

sum of the estimation error values over a spatial virtual microphone grid, given by

$$J_{GA} = \sum_{i=1}^{N_e} L_{\epsilon_i}. \quad (7)$$

For the optimisation of the robustness of the configurations, the condition number is used as the fitness function and is directly calculated using the corresponding built-in function in MATLAB. Smaller values denote better performance for both functions so the optimisation is formulated as a minimisation problem.

Selection After the fitness of each individual in a generation has been evaluated, the individuals are ordered based on a scale associated with their performance rank instead of their fitness score. The values are proportional to $\frac{1}{\sqrt{r}}$, where r is the rank of each individual. Ordering the individuals based on their rank removes the effect of the spread of the score values on the convergence of the search [11].

Elitism To ensure useful information is not lost from one generation to the next, the notion of *elitism* is used. The five individuals that scored highest are passed directly to the next generation. This ensures that the best fitness value of each generation cannot be worse than the previous and facilitates convergence to an optimal solution.

Crossover To traverse the search domain, individuals in new generations are created by combining elements of the parameter vectors of solutions in the current generation. After initial trials the percentage of the individuals in the new generation resulting from crossover was set to 80% which provides a good tradeoff between domain search and variability of solutions. For each individual created by crossover five candidates are picked at random and the two with the highest fitness score are chosen as the “parents”. To decide which part of the vector will be picked from each parent two integers, n and m in the range from zero up to the number of parameters are picked at random. Elements with indices less than n and greater than m are picked from the first parent and the rest from the second parent.

Mutation The new generation individuals that are not created through crossover result from mutation, which is





FORUM ACUSTICUM EURONOISE 2025

the process of applying random changes to some of the values in the parameter vectors of current generation individuals. To perform the mutation, a direction in the space of the parameter vector is picked and scaled at random. If any imposed constraints are not satisfied the direction and scale are adapted.

Constraints and bounds The minimum value of the estimation error given by Eq. (7) has been constrained to be greater than -30 dB for each virtual microphone, which aims to discourage solutions that achieve effective estimation on average across all positions by achieving extremely low error at a small number of microphone positions.

The values describing the coordinates of the microphones have been constrained to satisfy two conditions. Firstly, the microphones must not be located outside of a specified area to avoid overly favourable positioning overlapping with the virtual microphone grid. Secondly, the minimum allowed distance between the microphones is constrained to be approximately 0.0354m. This constraint serves a twofold purpose: first, to avoid overlapping microphones, and second, to ensure comparability with the “conventional” array against which the solutions are compared.

Termination conditions The optimisation is performed until 300 generations are evaluated, or the best fitness value change is smaller than 10^{-1} for 100 successive generations. At this point, the parameter vector associated with the best fitness values is returned as the optimal solution.

4. SIMULATIONS

4.1 Simulation setup

The GA was used in a scenario where monitoring microphones could be positioned over a rectangular spatial domain with side equal to 0.5m. A rectangular $0.5\text{m} \times 0.25\text{m}$ grid of 10×10 virtual microphones was used to evaluate the fitness function of Eq. (7). The complete geometry of the optimisation problem is presented in Fig. 3.

To facilitate comparison with a past study [9], a two element Uniform Linear Array (ULA) of inter-element distance of 0.1m was implemented and positioned parallel to the x Cartesian axis, as shown in Fig. 3. Each array element consists of a sub-array with two pairs of microphones, one orientated parallel to the x -axis and one to

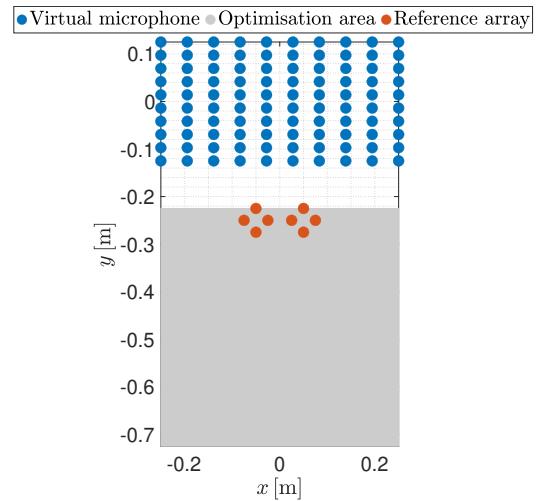


Figure 3. Geometry of the optimisation problem. Virtua microphones are denoted with blue dots and the microphones of the reference linear array with red dots. The grey square denotes the area the algorithm uses for the candidate monitoring configurations.

the y -axis. The distance between the two microphones of each pair is 0.05m.

A diffuse sound field was generated by a constellation of 501 ideal point sources on the surface of a sphere with radius of 10m arranged on a Fibonacci lattice. At low frequencies, the impact of the microphone topology is obscured by the increased performance due to high coherence between the monitoring and virtual microphones [2, 4]. For this reason, the optimisation was performed at a frequency of $f = 1\text{kHz}$, although the performance has been assessed over a broader frequency range.

4.2 Optimising estimation performance

First, the GA was used to calculate microphone configurations that provide optimal performance by minimising the fitness function of Eq. (7). The optimisation was performed five times and the spatial distribution of the estimation error over the virtual microphone grid at the optimisation frequency is presented in Fig. 4, along with the performance of the reference ULA. The black line indicates the area over which the estimation error is less than -10 dB.





FORUM ACUSTICUM EURONOISE 2025

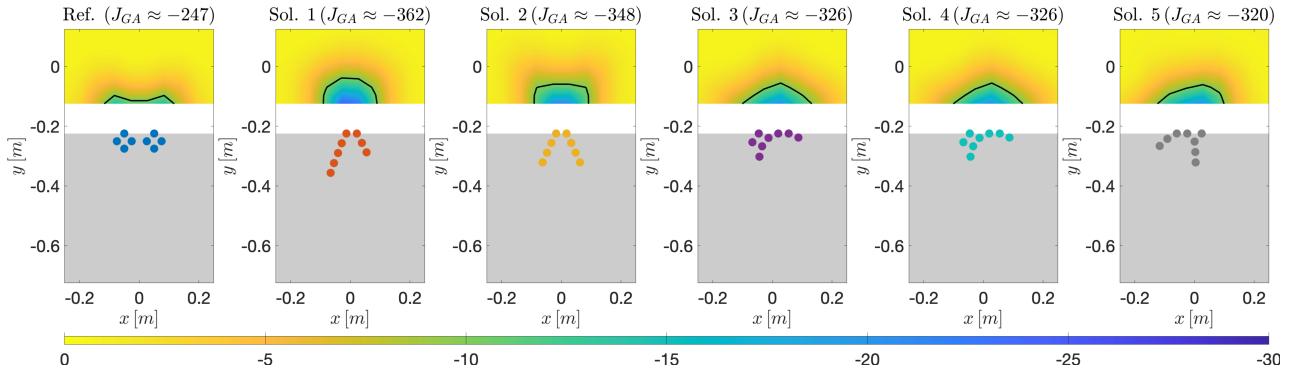


Figure 4. Spatial estimation error distribution for the microphone configurations calculated with the GA at the frequency of 1 kHz. The solid line (—) indicates the area over which the estimation error is less than -10 dB.

The configurations obtained with the GA yield significantly lower fitness values than the reference design. For the GA setups, areas with estimation errors below -10 dB consistently extend further from the microphone array, and the lowest error achieved is consistently lower than that of the reference arrangement.

Distinct patterns are observed in the topologies of the solutions. The configurations with the best fitness scores, solutions 1 and 2, consist of linear sub-arrays oriented toward the estimation area, extending the -10 dB error zone along their axes [5, 9]. The other configurations utilise the same mechanism, to a lesser extent, combined with microphone placement near the virtual microphone grid to achieve a low fitness value, resulting in slightly worse estimation spread over a larger area.

The performance gain achieved with the calculated configurations comes at the cost of increased sensitivity to uncertainties, which is reflected in the condition numbers associated with the arrays. Fig. 5 presents the condition numbers for the calculated arrays at frequencies from 50 Hz to 2 kHz. The condition numbers range from about 10^6 for the reference array to approximately 2.5×10^7 for the solution with the lowest fitness score at the optimisation frequency, with the span increasing significantly at lower frequencies. Additionally, it is important to highlight the large range of the fitness scores being 42 dB at the frequency of 1 kHz.

4.3 Optimisation with constrained condition number

An additional constraint on the upper bound of the condition number was introduced to control the sensitivity of

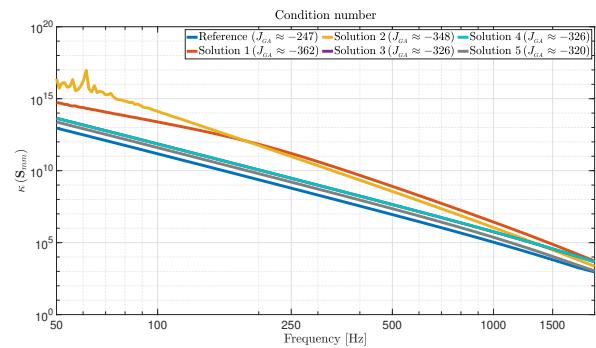


Figure 5. The condition number for the arrays calculated with the genetic algorithm for frequencies from 50Hz to 2kHz.

the calculated arrays. Additionally, solutions 2 and 3 of Fig. 4 were included in the initial population. The optimisation was performed five times for each of the condition numbers from 10^3 to 10^8 with a geometric step of 10. The spatial distribution of the estimation error for the configurations associated with the lowest fitness score from each trial are shown in Fig. 6.

Even with the condition number constrained to low values, the optimised arrays achieve lower function scores than the reference array. The spacing between microphones is larger under strict condition constraints and decreases with more relaxed bounds. The topologies exhibit significant similarities; for condition bounds up to 10^5 , the primary difference between configurations is the micro-





FORUM ACUSTICUM EURONOISE 2025

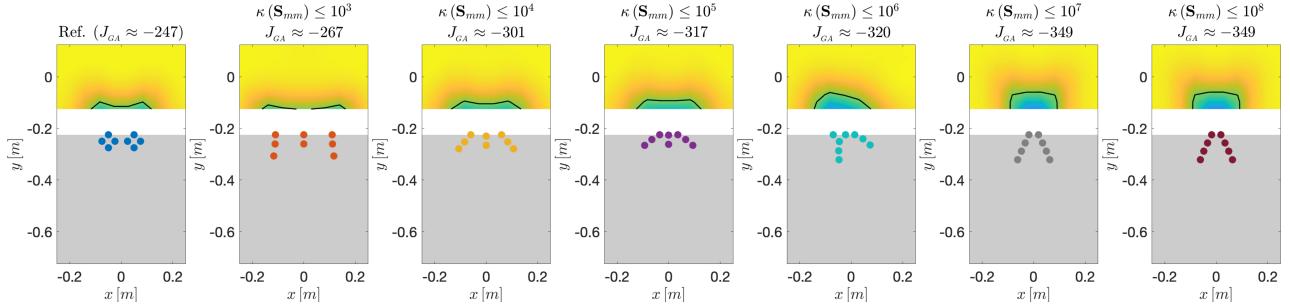


Figure 6. Spatial estimation error distribution for the microphone configurations associated with the lowest fitness score when the condition number is constrained. The solid line (—) indicates the area over which the estimation error is less than -10 dB.

phone spacing. At higher condition constraints, the optimal configurations resemble those calculated without any constraint on the condition, indicating that further relaxation of the bound is unlikely to yield significant performance benefits.

Fig. 7 presents the condition numbers of the microphone configurations of Fig. 6. The conditioning of the arrays is consistent with the imposed constraints up to approximately the optimisation frequency. For constraint values of 10^6 and higher, the configurations with the best fitness scores are those introduced in the initial population, and their condition does not reach the bound. Notably, among the two configurations, only the one returned for $\kappa(S_{mm}) \leq 10^6$ satisfies this constraint. When the constraint is relaxed, the configuration that achieves the best score is returned instead.

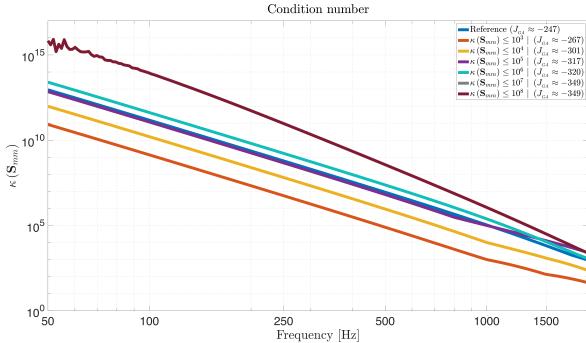


Figure 7. Conditioning of the microphone configurations associated with the lowest fitness score when the condition is constrained.

It is important to investigate the consistency of the GA in calculating the optimal arrays presented in Fig. 6. Fig. 8 presents the spread of the fitness scores for the calculated arrays at each constraint value, revealing a score range of the order of 40 dB for all configurations, consistent with the range for the unconstrained optimisation. The vertical overlap of the boxes indicates that, for a specified constraint, a calculated array may perform worse than those derived with more relaxed condition bounds, which contradicts the purpose of the optimisation.

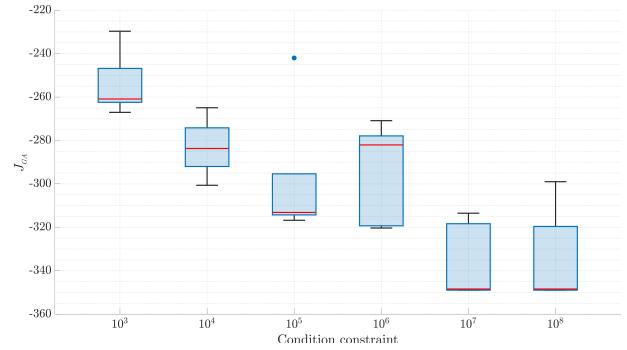


Figure 8. Fitness score spread for the calculated arrays for each condition constraint value.

4.4 Optimisation with constrained estimation error

An alternative approach to the optimisation is to constrain the estimation error to be lower than a specified bound and minimise the condition number. Five trials





FORUM ACUSTICUM EURONOISE 2025

were performed with the total estimation error over the virtual microphone grid being upper bounded by values from -100 dB to -400 dB in steps of 50 dB. The same two arrays were included in the initial population for these trials too.

The spatial distribution of the error for the arrays associated with the lowest condition numbers is shown in Fig. 9. The titles indicate the constraint and the actual value achieved by the array. It can be seen that when the bound was set very low, no feasible solution was found and the array that was closest to satisfying the constraint was returned. The minimisation of the condition number leads to distinctively different arrays compared to those acquired by minimising the error, with inconsistencies in the performance. The array associated with the $J_{GA} \leq -300$ dB has a condition number of about 3.4×10^4 , while the array in Fig. 6 associated with a condition number bound of 10^4 achieves an error of $J_{GA} \approx -301$.

The range and median of the condition number associated with the calculated arrays for each constraint value are presented in Fig. 10. The spread of the values is significant when the consistency in the estimation error is taken into account. The results suggest that the same error values can be achieved with arrays whose conditioning differs significantly from each other. In contrast to the results obtained for the error minimisation, no significant overlap is observed between the condition numbers of the arrays calculated for each constraint bound.

5. SUMMARY

In this study, the applicability of a genetic algorithm for the calculation of optimal microphone configurations used to estimate the pressure in a tonal diffuse sound field at some location remote from the array was investigated.

The sum of estimation errors over a spatial grid of virtual microphones was used as the optimisation criterion. Microphone arrangements with similar topological patterns demonstrated performance gains compared to a reference linear array composed of sub-arrays capable of incorporating pressure and pressure gradient information into the estimation process [9]. However, the performance gains were accompanied by increased sensitivity to uncertainties, as assessed via the numerical conditioning.

Further optimisation trials were conducted with constraints on the condition number or the sum of estimation errors over the grid. Significant inconsistencies were observed in the results, with similar error metrics resulting from distinctly different array topologies. Overall, the

range of the fitness values associated with the calculated configurations was found to be excessive with significant overlap for different constraint values in certain situations. Further refinement of the algorithm and its parameters is required if the method is to be used for the calculation of optimal microphone configurations.

6. 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.

The authors would also like to acknowledge the use of *IRIDIS* High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work. 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.

7. REFERENCES

- [1] W. 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,” *J. Acoust. Soc. Am.*, vol. 143, pp. 2858–2869, May 2018.
- [2] W. Jung, S. J. Elliott, and J. Cheer, “Local active control of road noise inside a vehicle,” *Mech. Syst. Signal Process.*, vol. 121, pp. 144–157, Apr. 2019.
- [3] S. J. Elliott and J. Cheer, “Modeling local active sound control with remote sensors in spatially random pressure fields,” *J. Acoust. Soc. Am.*, vol. 137, pp. 1936–1946, Apr. 2015.
- [4] P. 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,” *J. Acoust. Soc. Am.*, vol. 152, pp. 2840–2848, Nov. 2022.
- [5] X. Zheng, Z. Jia, B. Wan, M. Zeng, and Y. Qiu, “A study on hybrid active noise control system combined with remote microphone technique,” *Appl. Acoust.*, vol. 205, p. 109296, Mar. 2023.





FORUM ACUSTICUM EURONOISE 2025

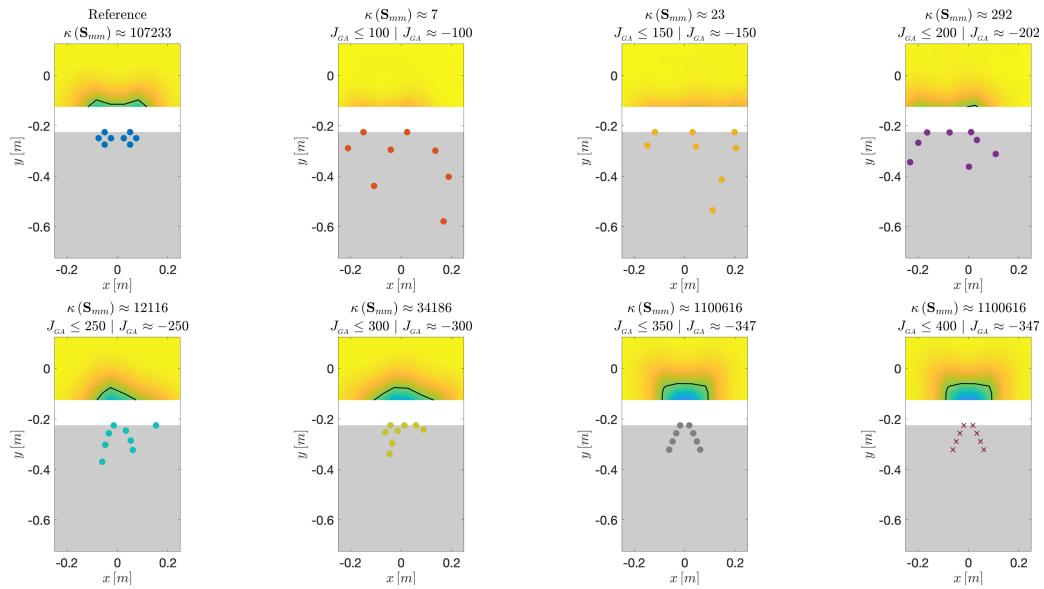


Figure 9. Spatial estimation error distribution for the microphone configurations with the lowest condition number when the total estimation error is upper bounded. The solid line (—) indicates the area over which the estimation error is less than -10 dB.

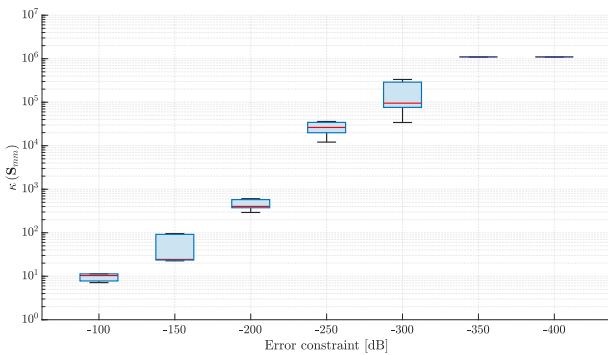


Figure 10. Fitness score spread for the calculated arrays per condition constraint value.

[6] S. L. Padula and R. K. Kincaid, “Optimization Strategies for Sensor and Actuator Placement,” Technical Memorandum 209126, NASA, NASA Langley Research Center Hampton, VA 23681-2199, Apr. 1999.

[7] D. J. Moreau, J. Ghan, B. S. Cazzolato, and A. C. Zander, “Active noise control in a pure tone diffuse

sound field using virtual sensing,” *J. Acoust. Soc. Am.*, vol. 125, pp. 3742–3755, June 2009.

- [8] S. Elliott and J. Garcia-Bonito, “Active cancellation of pressure and pressure gradient in a diffuse sound field,” *Journal of Sound and Vibration*, vol. 186, pp. 696–704, Oct. 1995.
- [9] A. Kappis, J. Cheer, and J. A. Zhang, “Remote Sensing Using Multi-Microphone Configurations for Local Active Noise Control Applications,” in *Proc. 30th Int. Congr. Sound Vibr.*, (Amsterdam, Netherlands), June 2024.
- [10] A. R. Conn, N. Gould, and P. L. Toint, “A globally convergent Lagrangian barrier algorithm for optimization with general inequality constraints and simple bounds,” *Math. Comput.*, vol. 66, pp. 261–289, Jan. 1997.
- [11] S. Katoch, S. S. Chauhan, and V. Kumar, “A review on genetic algorithm: past, present, and future,” *Multimed. Tool. Appl.*, vol. 80, pp. 8091–8126, Feb. 2021.

