DOI: 10.61782/fa.2023.0584

forum acusticum 2023

IDENTIFICATION OF MULTIPOLE SOURCES WITH NEURAL DECONVOLUTION

Thiago Lobato[1*] **Roland Sottek**[1] **Michael Vorländer**[2]

[1] HEAD acoustics GmbH, 52134 Herzogenrath, Germany
[2]Institute for Hearing Technology and Acoustics, RWTH Aachen University, Germany

ABSTRACT

Beamforming techniques for localizing sound sources often assume the radiation of multiple monopoles and use a simplified transfer function between the source and the array. While this approximation provides satisfactory results in many practical cases, it reaches its limits for more complex sources. By considering multipole sources such as dipoles, we can improve the results; however, the orientation of the dipoles must be known a priori. In this paper, the neural deconvolution method, which employs neural networks to help deconvolve the beamforming map, is extended to include multipole sources. This is done by developing an extended version of the DAMAS algorithm, which we call DAMAS-MS, where MS stands for "Multiple-Sources", and by applying a neural grid compression technique to make the whole process run fast enough for real-time applications. We show that it is possible to obtain reasonable estimates of source strength, location and radiation type without prior knowledge of the dipole orientation, and fast enough for real-time applications.

Keywords: *Beamforming, Deconvolution, Neural Networks, Multipole, Neural grid compression.*

1. INTRODUCTION

Beamforming techniques are widely employed to localize sound sources and quantify their strengths [1–4]. Typically, a grid of potential source directions is defined and unwanted directions are filtered out to determine the strength of each source. The technique must assume the radiation of the sources, such as considering them as monopoles [4]. This approximation generally works well but can fail for complex sources with non-monopole directivity [5, 6]. If the complex directivity pattern of the source is known, the method can still be effective [5, 6]. However, as we often lack knowledge about the source itself, a "chicken and egg" problem arises. This is because the directivity cannot be known without quantifying the source, and the source cannot be quantified without knowing its directivity. An alternative is to calculate various predefined sources and use the strongest one [6]. However, this approach is computationally intensive since it involves calculating various full beamforming maps and disregards cases where multiple sources with different propagation types may be present on the map.

To address this challenge, we extended the deconvolution technique called DAMAS [7], originally developed for monopoles, to accommodate sources with multipole directivities. This expansion requires solving a linear system that is highly ill-posed and computationally expensive, making it difficult to implement in practice. We overcome this difficulty by applying an extended neural grid compression as proposed in our recent work [8], which sometimes reduces the size of the linear system to be solved by more than 99%. This enables the development of an efficient approach to identify multipole sources.

2. MULTI-DIRECTIVITY BEAMFORMING

2.1 Classical beamforming

The main approach for beamforming is the delay-and-sum algorithm [9], which exploits the differences in delay and

*Corresponding author: thiago.lobato@head-acoustics.com.

Copyright: ©2023 Lobato 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.

amplitude between microphones in an array to filter out unwanted directions and accurately characterize sources. The starting point of the algorithm is the effect of a single source from direction i on the pressure at an array of microphones, as shown in Eqn. (1):

$$\boldsymbol{p}_i = a_i \cdot \boldsymbol{g}_i \tag{1}$$

where a_i is the source strength, $\boldsymbol{p}_i$ is the $Mx1$ complex pressure vector in the frequency domain, radiated from the single source at direction i to all M microphones in the array, and $\boldsymbol{g}_i$ is a $Mx1$ vector with the transfer functions between the source and the microphones. This transfer function determines the radiation model we use, which we will discuss in the next section. First, we continue with the derivation of the beamforming, which aims to identify auto strengths of the source $A_i = |a_i|^2$ at each direction i from a grid of directions I. To achieve this, we can apply the following formula:

$$|a_i|^2 \approx \tilde{A}_i = 0.5 \cdot \frac{\boldsymbol{g}_i^H \boldsymbol{p}\boldsymbol{p}^H \boldsymbol{g}_i}{|\boldsymbol{g}_i|^4} = \boldsymbol{w}_i^H \boldsymbol{C} \boldsymbol{w}_i \tag{2}$$

For a grid of directions I, this can be written as:

$$\tilde{A}_i = A_i + \sum_{j=1, j\neq i}^{I} A_j |\boldsymbol{g}_j^H \boldsymbol{w}_i|^2 \tag{3}$$

Here, $\boldsymbol{C}$ is the MxM signal cross-spectra matrix, $\boldsymbol{p}$ is the $Mx1$ measured pressure on the array, which is the sum of all $\boldsymbol{p}_i$, $\boldsymbol{w}_i$ is the $Mx1$ normalized steering function, $\tilde{A}_i$ is the least-squares solution for direction i, and $(\cdot)^H$ stands for the conjugate transpose. Eqn. (3) presents the relationship between the real strength A_i and the estimated strength $\tilde{A}_i$. Ideally, the sum term on the right side which represents the interference of the unwanted directions should be minimized.

2.2 Deconvolution

A method to improve the beamforming results and reduce the strong interference in Eqn. (3) is deconvolution, which removes the influence of the array on the source. One approach is the DAMAS method [7], which models the cross-spectra matrix $\boldsymbol{C}$ as a sum of incoherent sources from each direction I. Each source has its own contribution, $\boldsymbol{C}_{mod,i}$, as shown in Eqn. (4).

$$\boldsymbol{C}_{mod,i} = A_i \cdot \boldsymbol{g}_i \boldsymbol{g}_i^H \tag{4}$$

The full matrix is the sum of the individual contributions:

$$\boldsymbol{C}_{mod} = \sum_{i}^{I} \boldsymbol{C}_{mod,i} \tag{5}$$

Substituting this in Eqn. (2) yields Eqn. (6).

$$\tilde{A}_i = \sum_{j}^{I} A_j \cdot \frac{\boldsymbol{g}_i^H \boldsymbol{g}_j \boldsymbol{g}_j^H \boldsymbol{g}_i}{|\boldsymbol{g}_i|^4} = \sum_{j}^{I} A_j \cdot P_{i,j} \tag{6}$$

From this, a linear system relating the beamforming pattern and the real source strength can be derived as:

$$\boldsymbol{P}\boldsymbol{A} = \tilde{\boldsymbol{A}} \tag{7}$$

$\boldsymbol{P}$ contains the IxI point-spread-functions of all grid directions, $\boldsymbol{A}$ is the $Ix1$ vector with all real source strengths, and $\tilde{\boldsymbol{A}}$ is the $Ix1$ vector containing the measured strengths. This linear system can be solved with a non-negative Gauss-Seidel method, but it is not fast enough for real-time applications when using high-resolution grids and only considers one type of source.

2.3 Multipole Transfer Functions

The transfer functions $\boldsymbol{g}$ are defined based on the type of source represented, which is often assumed to be a monopole. However, we can define them to contain different terms of a multipole expansion, such as dipoles and quadrupoles. The equivalent transfer function $\boldsymbol{g}_{mu,i}$ of a multipole expansion is defined in Eqn. (8):

$$\boldsymbol{g}_{mu,i}(k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} c_{n,m}(k) h_n^{(2)}(kr_i) Y_{n,m}(\theta_i, \psi_i) \tag{8}$$

in which $Y_{n,m}(\theta_i, \psi_i)$ is the spherical harmonic of order m and degree n at the given direction, $h_n^{(2)}$ is the spherical hankel function of second kind, k is the wave number and $c_{n,m}$ are the multipole coefficients that define the radiation. We can merge these components into a total of N individual transfer functions and define a MxN transfer function matrix $\boldsymbol{G}$ in which each column $\boldsymbol{G}_k$ is the transfer function of a particular radiation term. In this way we can defined monopoles, dipoles, quadrupoles, etc. Instead of using single terms, we obtain better results by defining combinations of coefficients representing a particular source type. This allows us to consider complex

higher order radiation terms in the calculation. In this work, we consider combinations representing a "vertical" and a "horizontal" dipole.

2.4 DAMAS with multiple transfer functions per direction (DAMAS-MS)

We can easily expand the DAMAS method to consider multipole sources as well. To illustrate this, we rewrite Eqn. (5) as Eqn. (9).

$$\boldsymbol{C}_{mod} = \sum_{k=1}^{N} \sum_{i}^{I} \boldsymbol{C}_{mod,i,k} \tag{9}$$

In which:

$$\boldsymbol{C}_{mod,i,k} = A_{i,k} \cdot \boldsymbol{G}_{i,k} \boldsymbol{G}_{i,k}^{H} \tag{10}$$

$$\boldsymbol{C}_{mod,i} = \sum_{k=1}^{N} A_{i,k} \cdot \boldsymbol{G}_{i,k} \boldsymbol{G}_{i,k}^{H} \tag{11}$$

Here we basically say that each direction can have more than one component, but that all cross-terms are 0. We make this assumption because this work is a first proof of concept of the method, although we could calculate each individual cross-term with a significant increase in computational complexity. Similar to the regular DAMAS, this result can be replaced in Eqn. (2) to obtain:

$$\tilde{A}_i = \sum_{k=1}^{N} \sum_{j}^{I} A_{j,k} \cdot \frac{\boldsymbol{g}_i^H \boldsymbol{G}_{j,k} \boldsymbol{G}_{j,k}^H \boldsymbol{g}_i}{|\boldsymbol{g}_i|^4} = \sum_{k=1}^{N} \sum_{j}^{I} A_{j,k} \cdot P_{i,j,k} \tag{12}$$

The relation in Eqn. (12) yields the following linear system:

$$\boldsymbol{P}_m \boldsymbol{A}_m = \tilde{\boldsymbol{A}} \tag{13}$$

In which $\boldsymbol{P}_m$ is the expanded $Ix(I \cdot N)$ point-spread-function with multiple radiations per point and $\boldsymbol{A}_m$ is the $(I \cdot N)x1$ vector containing the amplitude of each source component. We call this expansion **DAMAS-MS**, where "MS" stands for "Multiple Sources". One of the most important decisions is the choice of the function to use for $\boldsymbol{g}$, since all other terms will be projected on it. To illustrate the implications of this choice, we show in Fig. (1) to (2) the monopole projection of different multipole terms placed at the origin of the scanning grid. In Fig. (2) we see that, by using the monopole projection on a dipole we can have a completely different strength level at the real source position, and even the positions with the strongest radiation show values more than 25 dB weaker than the real source. Fig. (3) shows the result using the correct projection with the 60 dB strength on the center of the grid. This illustrates how important it is to know the correct source type.

One is not restricted to using only one projection, but the projections of all transfer functions in $\boldsymbol{G}$ can be used. This can be done trivially by generating the corresponding linear system in Eqn. (13) for each projection and then concatenating the matrices and vectors, yielding a square $(I \cdot N)x(I \cdot N)$ matrix $\boldsymbol{P}_m$ and a concatenated $(I \cdot N)x1$ vector $\tilde{\boldsymbol{A}}$ containing all beamforming maps, while $\boldsymbol{A}_m$ remains the same. By considering all projections we remove some of the ambiguity of the deconvolution approach. Still, even the original "smaller" DAMAS system in Eqn. (7) is already extremely ill-posed [7], so solving Eqn. (13) directly is cumbersome. Furthermore, the system to be solved becomes very large and we also need to calculate N beamforming maps, which can make the method prohibitively expensive, especially for real-time applications. To address this issue, we combine the DAMAS-MS with a neural grid compression presented in the next section.

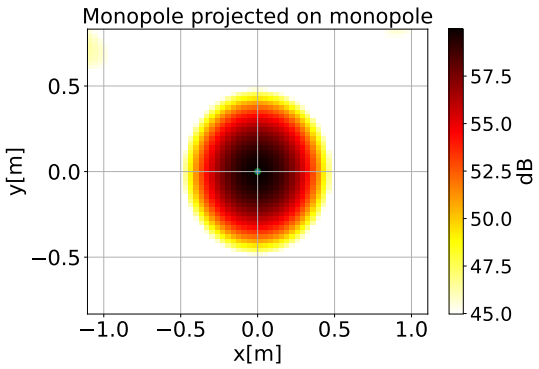

Figure 1: Projected radiation of a monopole source with 60 dB strength on the center of the grid using a monopole transfer function.

2.5 Neural grid compression

We base our work on the neural deconvolution method presented by Lobato et.al [8] that uses a neural network to localize the sources and then solves the inverse deconvolution problem only for the positions where a source is present. The method can be described by the following

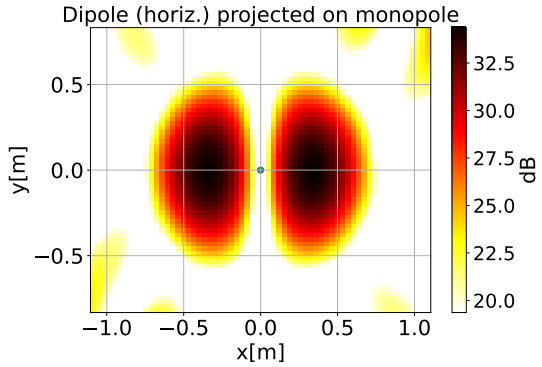

Figure 2: Projected radiation of a horizontal dipole source with 60 dB strength on the center of the grid using a monopole transfer function.

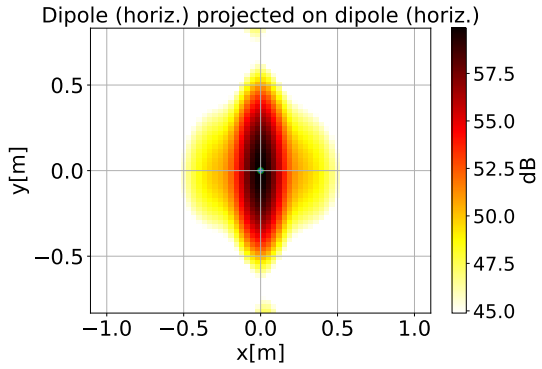

Figure 3: Projected radiation of a horizontal dipole source with 60 dB strength on the center of the grid using a dipole transfer function.

equations:
$$\widehat{\boldsymbol{A}}_{m,c} = NNLS(\boldsymbol{P}_{m,c}; \widetilde{\boldsymbol{A}}) \tag{14}$$

$$\widehat{\boldsymbol{A}}_{\notin c} = 0 \tag{15}$$

$$c = \{i, k \mid i \in I \text{ and } NN(\widetilde{\boldsymbol{A}}_i, k) \geq \gamma \cdot \max(NN(\widetilde{\boldsymbol{A}}, k))\} \tag{16}$$

Here NN is the neural network indicating the probability that a given source of the expansion term k is present at direction i, $NNLS$ is a non-negative least-squares solver, γ controls the sensitivity of the method, and $\widehat{\boldsymbol{A}}_{m,c}$ is the estimated strength of every radiation term and c represents the set of all directions and expansion terms considered in the calculation, so the subscript indicates that we use all elements of the array/matrix related to the direction/expansion term pair. We then sum the mask of all components to obtain a single mask with all positions with sources. In this case, we use only the beamforming map projected onto the monopole transfer function as input for the network. This approach allows us to very efficiently obtain a grid with sources that is often more than 100 times smaller than the full grid [8]. With this compressed grid, we can calculate the additional beamforming projections needed for the full Eqn. (13) only at the compressed directions, which has negligible calculation time compared to the beamforming on the full grid, thus obtaining all the necessary projections in roughly the same time as a single beamforming calculation.

3. METHODOLOGY

We used a similar raw U-Net as presented in [8], now with N outputs in the last layer to account for all source terms. We train the network with an encoder-decoder type of loss. The encoder is the U-Net neural network and the decoder is the linear system in Eqn. (13). The full loss or the method is defined as:

$$L(\boldsymbol{A}, \widetilde{\boldsymbol{A}}) = \frac{1}{N} \cdot \sum_{k=1}^{N} E_{\text{enc}}(NN(\widetilde{\boldsymbol{A}}, k), \boldsymbol{A}_{m,k,bi}) \cdot (1 + \lambda \cdot E_{\text{dec}}(\boldsymbol{P}NN(\widetilde{\boldsymbol{A}}, k), \boldsymbol{P}\boldsymbol{A}_{m,bi})) \tag{17}$$

where the suffix bi indicates a binarized vector in which all values above 0 are set to 1. We set $\lambda = 0.01$ and use the point-spread-function $\boldsymbol{P}$ of a monopole projection independent of the source type for simplification. The encoder loss is:

$$E_{\text{enc}}(NN(\widetilde{\boldsymbol{A}}), \boldsymbol{A}_{bi}) = \frac{1}{N \cdot I} \sum_{k=1}^{N} \sum_{i=1}^{I} (-w \cdot \boldsymbol{A}_{m,i,k,bi} \cdot \log(NN(\widetilde{\boldsymbol{A}}_{i,k})) + (1 - \boldsymbol{A}_{m,i,k,bi}) \cdot \log(1 - NN(\widetilde{\boldsymbol{A}}_{i,k}, k))) \tag{18}$$

where w is a weighting for positive classes, that needs to be high, otherwise the network will predict mostly zeros, since we have mostly sparse sources. Here we set $w =$

$N \cdot 100 = 300$. The decoder loss is defined as:

$$E_{\text{dec}}(\boldsymbol{PNN}(\widetilde{\boldsymbol{A}}), \boldsymbol{PA}_{m,bi}) = \sqrt{\frac{100}{N \cdot I} \cdot \sum_{k=1}^{N} \sum_{i=1}^{I} \left(\log_{10} \left(\frac{\boldsymbol{PNN}(\widetilde{\boldsymbol{A}}_k, k)}{\boldsymbol{PA}_{m,k,bi}} \right) \right)^2} \quad (19)$$

To train the network, we generated the following distribution of sources:

- 1 to 10 uniformly placed sources, where each source has only one uniformly chosen expansion term (monopole, horizontal dipole, or vertical dipole) producing a sine wave of a given frequency.
- Signal-to-noise ratio of the sources between 20 dB and ∞ dB, where we sampled the noise amplitude uniformly between 0 and 1 times the strength of the lowest source on the grid.
- Source dynamic range of maximum 40 dB, where we uniformly sampled for each source amplitude dB values between -40 dB and 0 dB and converted them to linear amplitudes.
- The normalized (maximum equals 1) linear beampattern for the selected array in the defined grid is used as input to the neural network.
- The target of the neural network is the deconvolved map with a value of 1 at grid positions to which sources were assigned and zero otherwise. We assign the sources to the closest grid point while the generated beampattern is from the exact source locations.
- The output of the neural network is the probability of a source at each grid location for each term.
- Uniformly sampled frequency on a Bark scale between 3.5 Bark_{HMS} and 26.5 Bark_{HMS} [10, 11].

We generated 184000 samples for training and 46000 samples for validation and trained the network with Pytorch [12]. After the training, we evaluated the approach against the following metrics using a new test set with 1000 samples that we had not seen during training:

- The average dB error of the sources found.
- The distance error between sources, where we compared the real position of the source term to a centroid (weighted with the estimated source energies) of the estimated source positions.
- The percentage of cases in which the prominent expansion term was correctly estimated by the network.
- The precision and recall of the method.
- The calculation time of the method at inference.
- The confusion matrix of the source type, being the calculated type identified as the strongest component at a direction after deconvolution.

For these evaluation samples, we restricted the number of sources at each frequency to obtain more realistic results, since, for example, 10 dipoles at 300 Hz is an unlikely case in real life. This was done by making the limit of maximum sources on the map a function of frequency by increasing it by 1 for each 1000 Hz. The array used for the analysis was a spiral microphone array with 56 microphones shown in Fig. (4) for a $64x64$ scanning grid at 3 m from the array, where $\gamma = 0.7$ was chosen.

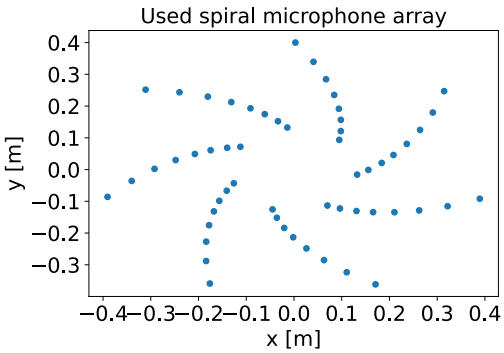

Figure 4: Array used for the experiment.

4. RESULTS

The strength error is shown in Fig. (5). There we have an average error of 0.5 dB, with most errors (80 %) being below ca. 4.5 dB. Considering that errors due to an incorrect choice of source type can go up to the full strength of the source and are around 25 dB even if the maximum apparent source strength is taken into account, we consider this result to be an excellent improvement over using only monopole transfer functions as typically done in beamforming.

The distance error can be seen in Fig. (6), and we see an average error of 0.009 m, which shows that as long as

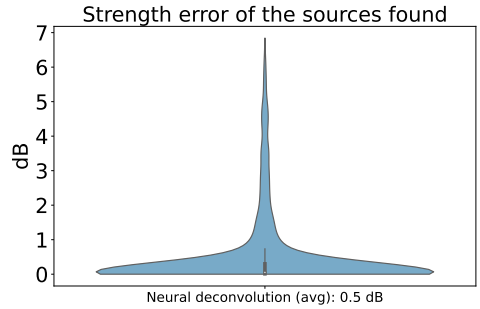

Figure 5: dB error of the sources found.

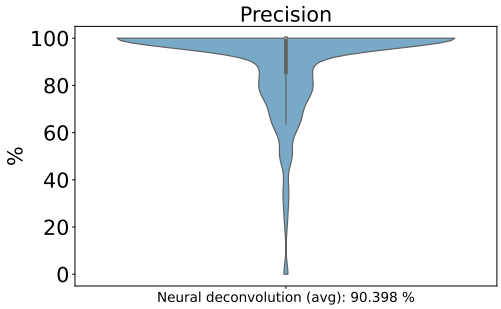

Figure 7: Precision of the method.

the source is identified by the method, its location is very accurate.

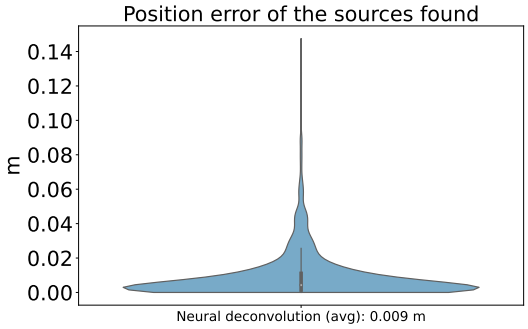

Figure 6: Distance error of the sources found.

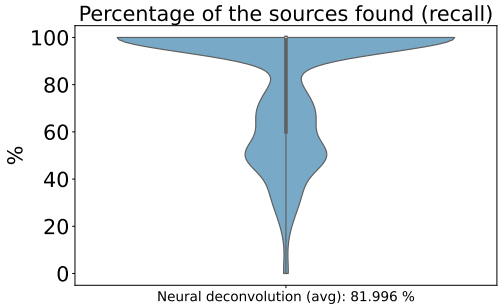

Figure 8: Recall of the method.

The next question is how many of the sources are identified (recall) and how many of them are correct (precision). The precision is shown in Fig. (7), and the recall is shown in Fig. (8). We verify that the recall of the method is about 82% and the precision is 90%. Both values are very good, indicating that, besides some outliers, the method can localize the sources of our test distribution pretty well.

The confusion matrix of the real and predicted source types can be seen in Fig. (9), in which we defined the selected type as the source type with the strongest amplitude given by the DAMAS-MS. There we see that the dipoles can be identified with almost perfect accuracy. For monopoles, the model can identify the correct radiation in most cases but is still wrong about 25% of the time. Yet, for the cases in which the method select the wrong source type, the amplitude errors of monopole sources are not considerably high, being, on average, around 3.5 dB as we show in Table 1. This indicates that for cases in which the method has a wrong identification, the source type has a smaller importance.

	Mon.	Dip (horiz.)	Dip (vert.)
Mon.	0.44	3.57	3.58
Dip (horz.)	-	0.18	0.00
Dip (vert.)	-	-	0.14

Table 1: Average strength errors based on real source types (rows) and estimated ones (columns).

At last, in Fig. (10) we see that the calculation time is very low, about 7 ms on average, so the method can still run in real-time at least for monopoles and dipoles (less than 30 ms [8]). The calculation time has some outliers where the NNLS solver took too long to converge, since the current problem is considerably more difficult than the traditional monopole-only inversion. For an application

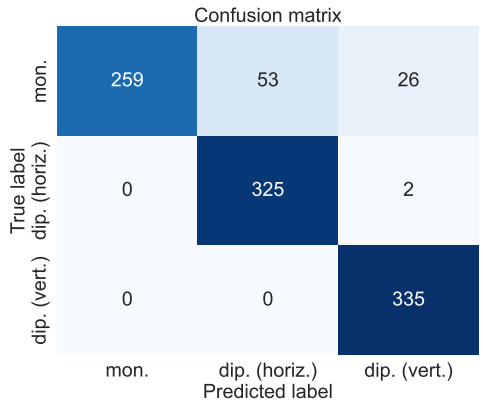

Figure 9: Confusion matrix of result. We see that the DAMAS-MS with neural grid compression method can distinguish dipoles very well, while being slightly confused by monopole sources.

within a product that has a maximum allowed latency, a good trade-off would be limiting the number of iterations of the solver.

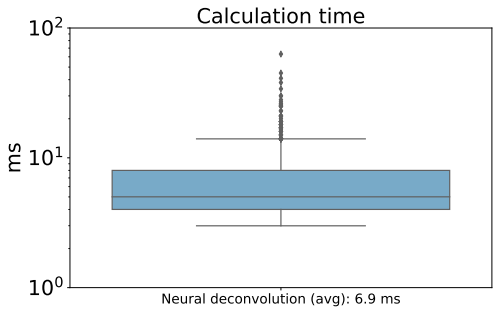

Figure 10: Calculation time of the method.

To illustrate how an individual result looks, we show an example in Fig. (11) along with the network mask in Fig. (12) and the deconvolution result in Fig. (13). We verify that the masks contain likely positions of multiple type of sources, not just for the current source type. The likely reason for this is that the monopole projection is ambiguous and thus the neural network is not capable of perfectly identifying only the correct type of source. However, this has not prevented the DAMAS-MS from finding the correct source positions and strengths in this example, which would be impossible with the traditional monopole-only DAMAS. In future publications, we may investigate new approaches to improve the network identification depending on how problematic this is for real sources and real-time implementation. A straight-forward, but computationally intensive, approach to improve the result is to use all projections as inputs for the network.

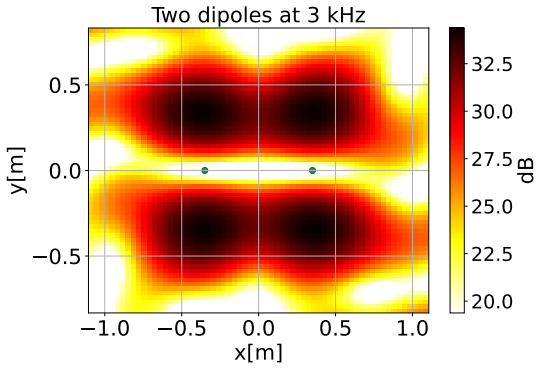

Figure 11: Monopole projected beampattern of two dipoles each with 60 dB strength.

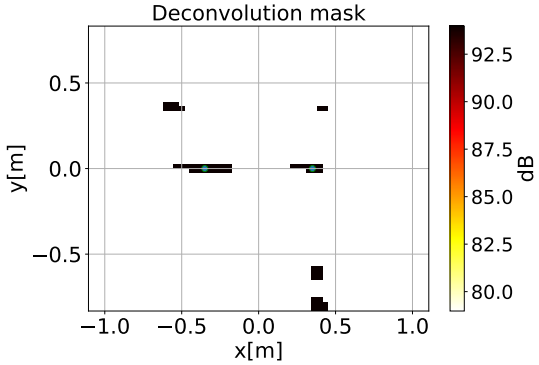

Figure 12: Neural network mask of pattern in Fig. (11).

5. CONCLUSION

We have derived an extended version of the DAMAS method, DAMAS-MS, which accommodates multiple source types per each direction and employs a neural grid compression to develop an efficient neural deconvolution algorithm for estimating the strength of various multipole components without prior knowledge of their direction, while retaining real-time feasibility. We validated

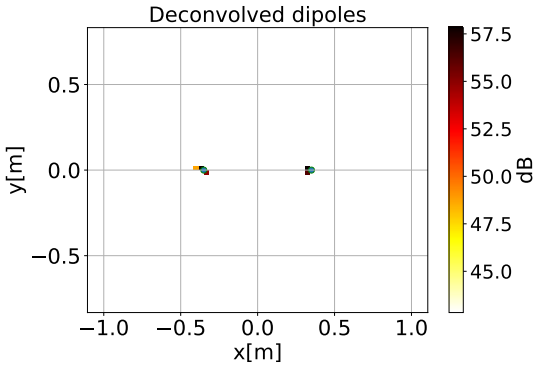

Figure 13: Deconvolved map with neural deconvolution of pattern in Fig. (11).

our method through simulating various beampatterns with different source types, yielding good results in terms of source location and strength estimation. The model was also able to identify the type of source pretty well in the case of dipoles, while the monopoles were somewhat ambiguous for the method. However, the confusion of monopoles with dipoles does not seem to be so critical for the cases wrongly identified by the model, since their average strength error still remained as low as ca. 3.5 dB. Notably, the method remains feasible for real-time implementation, at least for monopoles and dipoles. Future research will focus on validating the method with higher-order sources and investigating the generalization of the method for real sources, which is presented in its monopole-only variant [8].

6. REFERENCES

[1] R. Merino-Martínez, M. Snellen, and D. G. Simons, "Functional beamforming applied to imaging of flyover noise on landing aircraft," *J. Aircr.*, vol. 53, pp. 1830–1843, 2016.

[2] H. Camargo, P. Ravetta, R. Burdisso, and A. Smith, "A comparison of beamforming processing techniques for low frequency noise source identification in mining equipment," in *Proceedings of the ASME 2009 International Mechanical Engineering Congress and Exposition*, pp. 205–211, 2009.

[3] P.-A. Grumiaux, S. Kitić, L. Girin, and A. Guérin, "A survey of sound source localization with deep learning methods," *J. Acoust. Soc. Am.*, vol. 152, pp. 107–151, 2022.

[4] R. Merino-Martínez, P. Sijtsma, M. Snellen, T. Ahlefeldt, J. Antoni, C. J. Bahr, D. Blacodon, D. Ernst, A. Finez, S. Funke, T. F. Geyer, S. Haxter, G. Herold, X. Huang, W. M. Humphreys, Q. Leclère, A. Malgoezar, U. Michel, T. Padois, A. Pereira, C. Picard, E. Sarradj, H. Siller, D. G. Simons, and C. Spehr, "A review of acoustic imaging methods using phased microphone arrays," *CEAS Aeronaut. J.*, vol. 10, pp. 197–230, 2019.

[5] Y. Liu, A. R. Quayle, and A. P. Dowling, "Beamforming correction for dipole measurement using two-dimensional microphone arrays," *The Journal of the Acoustical Society of America*, vol. 124, no. 182, 2008.

[6] J. Gao, H. Wu, and W. Jiang, "Dipole-based beamforming method for locating dipole sources with unknown orientations in three-dimensional domains," *The Journal of the Acoustical Society of America*, vol. 147, no. 125, 2020.

[7] T. F. Brooks and W. M. Humphreys, "A deconvolution approach for the mapping of acoustic sources (damas) determined from phased microphone arrays," *J. Sound Vib.*, vol. 294, pp. 856–879, 2006.

[8] T. Lobato, R. Sottek, and M. Vorländer, "Deconvolution with neural grid compression: A method to accurately and quickly process beamforming results," *The Journal of the Acoustical Society of America*, vol. 153, no. 4, pp. 2073–2089, 2023.

[9] L. Santana, "Fundamentals of acoustic beamforming," in *NATO Educational Notes EN-AVT-287*, vol. 4, Washington, DC: NATO, 2017.

[10] R. Sottek, "A hearing model approach to time-varying loudness," *Acta Acustica united with Acustica*, vol. 102, pp. 725–744, July 2016.

[11] "ECMA 418-2 2nd edition: Psychoacoustic metrics for itt equipment, part 2: Models based on human perception," 2022.

[12] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, (Red Hook, NY), pp. 8026–8037, Curran Associates Inc., 2019.