ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10097014

NEURAL OPTIMIZATION OF GEOMETRY AND FIXED BEAMFORMER FOR LINEAR MICROPHONE ARRAYS

Longfei Yan[1,2]*,Weilong Huang*[3]*,W. Bastiaan Kleijn*[1]*,Thushara D. Abhayapala*[2]

[1]School of Engineering and Computer Science, Victoria University of Wellington,
[2]Audio & Acoustic Signal Processing Group, Australian National University,
[3]DingTalk Hummingbird Audio Lab, Alibaba Group

ABSTRACT

Fixed beamforming based on uniform linear microphone arrays often suffers from non-optimal performance for broadband signals. This paper addresses the issue by jointly optimizing the array geometry and spatial filters through a neural network based model. The model, composed of two feed forward neural networks, is optimized in an end-to-end manner. It satisfies the distortionless constraint in the look direction. Experimental results show that the proposed model outperforms the previous state-of-the-art fixed beamformer with overall better scores. Moreover, the proposed model can control the tradeoff between Directivity Factor (DF) and White Noise Gain (WNG) in a flexible way.

***Index Terms*—** linear microphone arrays, directivity factor, white noise gain, neural network, array geometry

1. INTRODUCTION

Broadband beamformers are widely applied in many areas such as radar, sonar, microphone arrays and radio astronomy [1–5]. Conventionally, sensors with uniform spacing are deployed in broadband beamformers. However, they have limitations in processing broadband signals. For example, uniformly spaced microphone arrays with limited array aperture and microphones may not function well in all the frequency bands ranging from a few hundred to a few thousand Hertz [6]. Spatial aliasing can happen if the inter-microphone spacing is larger than half of the wavelength for high frequency bands. Meanwhile, a compactly arranged microphone array can be sensitive to white noise and microphone imperfections for low frequency bands [7–9]. Therefore, uniform arrays often do not yield optimal performance for broadband signals.

Much research has shown that non-uniform arrays have superior performance than their uniform counterparts [10–12]. With non-uniform arrangement of microphones, the dilemma of spacing for both high and low frequency bands can be mitigated. In [6], a robust superdirective broadside beamformer was proposed through both stochastic and analytic optimization. This beamformer optimized array filters along with the non-uniform geometry. Its optimization objectives are either maximum directivity or frequency invariance. In [13], the array geometry optimization is reformulated from mixed-integer programming to convex optimization by relaxing some of the constraints and pre-select a set of potential microphone positions. The optimal microphone positions can be chosen from the pre-selected positions. However, mixed-integer problems are NP-hard. As a result, the computational complexity of the algorithm in [13] grows exponentially. It requires careful adjustment between solution optimality and computation feasibility.

Fixed beamformers have been employed for broadband beamforming as they are data independent and can be used in different acoustic environments. During the last decade, fixed beamformers like Differential Microphone Arrays (DMAs) have been popular in broadband beamforming due to their relative frequency-independent behaviour and high directivity gain [7]. Compared with other fixed beamformers such as Delay-and-Sum (DS) and superdirective beamformers, DMAs have more balanced DF and WNG values. They are inspired by the spatial derivative of the acoustic pressure field [14] and thus have compact apertures.

DMAs of different geometry have been investigated, such as linear [7,15], circular [16,17] and concentric circular [18,19]. Recently, DMA beamformers based on Particle Swarm Optimization (PSO) techniques [20] have successfully optimized the geometry for linear microphone arrays [21, 22]. We can observe that it is common to have microphone spacing larger than 1 cm even among DMAs. The DMA beamformer proposed in [22] focuses on frequency-invariant beampatterns by combining subarrays. In [21], a DMA beamformer proposed exhibited a better DF and WNG tradeoff than traditional approaches. However, the distortionless constraint is not guaranteed in its look direction. This can cause power distortion across different frequency bins, which undermines the quality of the broadband signal perceived. The white noise amplification phenomenon in low frequency bins also persists in these PSO-based DMA beamformers.

To address the above-mentioned issues in broadband beamforming, we propose a novel model called Neural Optimization of Non-Uniform Linear Array (NONULA) for fixed beamformers. The placement of microphones are not constrained to pre-defined locations. Maximum array size and minimum inter-microphone spacing are considered. Our contributions include: *(i)* To the best of our knowledge, NONULA is the first model that optimizes linear array geometry with neural networks. *(ii)* NONULA is the first end-to-end neural network-based model for simultaneous array geometry and filter optimization. *(iii)* NONULA can obtain overall better DF and WNG performance than previous models. *(iv)* NONULA facilitates designers with a choice of balance between DF and WNG in a flexible manner.

The rest of the paper is organized into five more sections. In Section 2, the employed signal model of beamforming is explained. In Section 3, various performance measures utilized in our study are presented. In Section 4, we first describe the existing PSO models that can optimize geometry for linear microphone arrays. Then we propose the NONULA model in details. Simulation results of previous models and our proposed NONULA model are discussed in Section 5. Section 6 summarizes the characteristics of the proposed NONULA model.

2. SIGNAL MODEL

We consider a non-uniform linear array comprised of M omnidirectional microphones. Denote δ_k as the distance from the kth microphone to the first microphone and $\delta_1 = 0$. The steering vector can be expressed as [1]

$$\mathbf{d}(\omega, \cos\theta) = [1 \; e^{-\jmath\omega\tau_2 \cos\theta} \; \cdots \; e^{-\jmath\omega\tau_M \cos\theta}]^T, \qquad (1)$$

where $\jmath = \sqrt{-1}$, $\omega = 2\pi f$, $\tau_k = \delta_k/c$, θ is the azimuth angle of the source signal, f is the temporal frequency and c is 340 m/s. We assume the source signal is from the endfire direction at $\theta = 0$.

Under anechoic conditions with far field assumption, the received signal vector $\mathbf{y}(\omega)$ by the microphone array is expressed as

$$\begin{aligned}\mathbf{y}(\omega) &= [Y_1(\omega) \; Y_2(\omega) \; \cdots \; Y_M(\omega)]^T \\ &= \mathbf{d}(\omega, \cos\theta_\ell)\mathbf{X}(\omega) + \mathbf{v}(\omega),\end{aligned} \qquad (2)$$

where $Y_k(\omega)$ is the signal received at the kth microphone, θ_ℓ represents the look direction, $\mathbf{X}(\omega)$ is the source signal and $\mathbf{v}(\omega)$ is the noise vector.

To estimate $\mathbf{X}(\omega)$ from the observed $\mathbf{y}(\omega)$, a complex linear filter $\mathbf{h}(\omega)$ is applied in beamforming. We have

$$\begin{aligned}\hat{\mathbf{X}}(\omega) &= \mathbf{h}^H(\omega)\mathbf{y}(\omega) \\ &= \mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos\theta_\ell)\mathbf{X}(\omega) + \mathbf{h}^H(\omega)\mathbf{v}(\omega),\end{aligned} \qquad (3)$$

where $\hat{\mathbf{X}}(\omega)$ is the estimate of $\mathbf{X}(\omega)$ from beamforming. The distortionless constraint requires that

$$\mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos\theta_\ell) = 1 \quad \forall\omega \qquad (4)$$

3. PERFORMANCE MEASURES

Beampattern quantifies the input-output behaviour of a microphone array given a source signal from the direction θ [23]. It is defined as

$$\mathcal{B}[\mathbf{h}(\omega), \theta] = \mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos\theta). \qquad (5)$$

The robustness of a microphone array to sensor imperfections such as sensor noise and positional errors can be quantified by White Noise Gain (WNG) [24]. It is expressed as

$$\text{WNG}[\mathbf{h}(\omega)] = \frac{|\mathcal{B}[\mathbf{h}(\omega), \theta_\ell]|^2}{\mathbf{h}^H(\omega)\mathbf{h}(\omega)}. \qquad (6)$$

Another performance measure commonly used together with WNG is Directivity Factor (DF). It evaluates the directivity of a microphone array in the presence of isotropic noise field. The formula of DF is

$$\text{DF}[\mathbf{h}(\omega)] = \frac{|\mathcal{B}[\mathbf{h}(\omega), \theta_\ell]|^2}{\mathbf{h}^H(\omega)\boldsymbol{\Gamma}_{0,\pi}(\omega)\mathbf{h}(\omega)}, \qquad (7)$$

where $\boldsymbol{\Gamma}_{0,\pi}(\omega)$ is a $M \times M$ matrix. The elements in $\boldsymbol{\Gamma}_{0,\pi}(\omega)$ are given by

$$[\boldsymbol{\Gamma}_{0,\pi}(\omega)]_{ij} = \text{sinc}[\omega(\delta_i - \delta_j)/c], \qquad (8)$$

where $\text{sinc}(x) = \sin x/x$.

When multiple measures are employed simultaneously for a physical system, naturally multi-objective optimization scenario arise. To compare the performance of different beamforming techniques in a multi-objective manner, we can adopt a weighted sum approach with respect to both WNG and DF [25]:

$$J_s(\omega) = \text{DF}[\mathbf{h}(\omega)] + \text{WNG}[\mathbf{h}(\omega)] \cdot r_w, \qquad (9)$$

where $J_s(\omega)$ is the multi-objective score for ω and r_w is the weighting coefficient for WNG.

4. MODELS FOR GEOMETRY OPTIMIZATION

4.1. PSO Models

The ideal beampattern of an Nth-order DMA is:

$$\mathcal{B}_N(\theta) = \sum_{n=0}^{N} a_{N,n} \cos^n \theta, \qquad (10)$$

where N is the order of derivative and $a_{N,n}$ is the nth coefficient of the beampattern.

Conventionally, null-constrained approach is an effective way to design DMA beampatterns [7]. An Nth-order DMA can have N distinct null directions, where the DMA beampattern is 0 in those directions. The matrix $\mathbf{D}(\omega)$ can be constructed with null constraints:

$$\mathbf{D}(\omega) = \begin{bmatrix} \mathbf{d}^H(\omega, \cos\theta_\ell) \\ \mathbf{d}^H(\omega, \cos\theta_1) \\ \vdots \\ \mathbf{d}^H(\omega, \cos\theta_N), \end{bmatrix} \qquad (11)$$

where $\theta_1, \cdots, \theta_N$ are N distinct null directions. This yields

$$\mathbf{D}(\omega)\mathbf{h}(\omega) = \mathbf{i}, \qquad (12)$$

where $\mathbf{i} = [1 \; 0 \; \cdots \; 0]^T$ is a binary vector of length $N+1$ that has zero entries everywhere except for the first entry.

Given $M > N + 1$, a minimum-norm solution of (12) can be obtained:

$$\mathbf{h}_{\text{MN}}(\omega) = \mathbf{D}^H(\omega)[\mathbf{D}(\omega)\mathbf{D}^H(\omega)]^{-1}\mathbf{i}, \qquad (13)$$

which maximizes WNG for DMA. This is also called the Maximum WNG (MWNG) differential beamformer [26].

To optimize the geometry of DMA, PSO techniques [21, 22] have been used to decide the geometry vector $\boldsymbol{\delta} = [\delta_1 \; \cdots \; \delta_M]$. In [21], two tradeoff parameters δ_w and δ_d are also employed to further refine the filters $\mathbf{h}(\omega)$:

$$\mathbf{h}(\omega) = \boldsymbol{\Gamma}_w^{-1}(\omega)\mathbf{D}^H(\omega)[\mathbf{D}(\omega)\boldsymbol{\Gamma}_w^{-1}(\omega)\mathbf{D}^H(\omega) + \delta_d\mathbf{I}]^{-1}\mathbf{i}, \quad (14)$$

where $\boldsymbol{\Gamma}_w(\omega) = [\boldsymbol{\Gamma}_{0,\pi}(\omega) + \delta_w\mathbf{I}]$. However, δ_d breaks the distortionless constraint when it is not zero.

The PSO algorithm in [21] starts with randomly initializing many candidate solutions for $\boldsymbol{\delta}$, δ_w and δ_d. They together form particles. A fitness function F is defined to evaluate and compare particles. In each iteration, the best particle, i.e., the best candidate solution, is picked from all the particles. Its information is utilized to guide the update of other particles. This update process repeats in every iteration until the specified number of iteration is reached. One of our baselines is a modified implementation of this PSO algorithm.

4.2. NONULA

The NONULA model proposed by us consists of two feed forward neural networks: SpacingNet and FilterNet. Fig. 1 demonstrates the workflow of NONULA. Denote inter-microphone spacing as $\boldsymbol{\eta} = [\eta_1 \; \eta_2 \; \cdots \; \eta_{M-1}]$, where $\eta_k = \delta_{k+1} - \delta_k$. Initially, the spacing input $\boldsymbol{\eta}_{\text{init}}$ is fed into SpacingNet. The output of SpacingNet is then used to calculate the steering vector $\mathbf{d}(\omega, \cos\theta_\ell)$. Subsequently, both $\mathbf{d}(\omega, \cos\theta_\ell)$ and the filter input are fed into FilterNet. If the termination condition is not met, SpacingNet and FilterNet will repeatedly optimize the spacing and filter output. In this process, the spacing and filter input will remain the same to make the

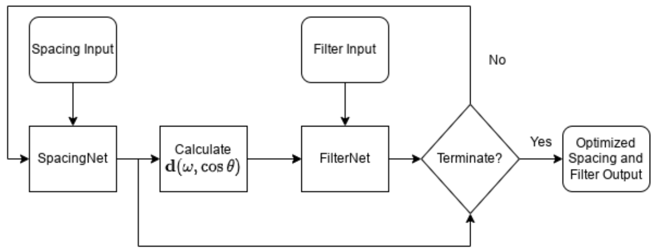

Fig. 1. Workflow diagram of NONULA

convergence of optimization easier. The weights of SpacingNet and FilterNet and $\mathbf{d}(\omega, \cos\theta_\ell)$ are updated in each iteration.

SpacingNet and FilterNet both have 5 layers: the input layer, three hidden layers and the output layer. SpacingNet uses ReLu as the activation function in hidden layers, whereas FilterNet uses complex ReLu [27, 28]. ReLu is the default activation function in many neural networks by virtue of easy training and good performance. Complex ReLu is necessary for facilitating the complex representation in the filter. The size of input and output layers is $M - 1$ for SpacingNet and M for FilterNet. SpacingNet also has the softmax activation in the final layer. Multiplying softmax outputs by the total array length allocates spacing between microphones.

Denote the minimum spacing as $\eta_{\min}$ and the maximum array aperture as $L_{\max}$. Denote the optimized spacing and filter output of SpacingNet and FilterNet as $\mathbf{s}$ and $\mathbf{H}$ respectively. To satisfy the constraint of $\eta_{\min}$ and $L_{\max}$, we allocate $L = L_{\max} - \eta_{\min} \times (M - 1)$ as the maximum size for spacing allocation. The actual spacing is $\boldsymbol{\eta} = \mathbf{s} + \eta_{\min} \cdot \mathbf{1}$, where $\mathbf{1}$ is a vector of 1's having the same length as $\mathbf{s}$. To satisfy the distortionless constraint, $\mathbf{H}$ is obtained after dividing $\mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos\theta_\ell)$ in every frequency bin for normalization.

The loss function J of NONULA is the mean of negative multi-objective score plus the penalty of under-achieved DF and WNG factors across all the interested frequency bins:

$$J = \frac{1}{N_\omega} \sum_{\omega} \Big[- J_s(\omega) + (\sigma_r(\mathrm{DF}_{tgt}(\omega) - \mathrm{DF}[\mathbf{h}(\omega)]) + (\sigma_r(\mathrm{WNG}_{tgt}(\omega) - \mathrm{WNG}[\mathbf{h}(\omega)])\Big], \tag{15}$$

where N_ω is the number of frequency bins, $\sigma_r(\cdot)$ represents the ReLu activation function, $\mathrm{DF}_{tgt}(\omega)$ and $\mathrm{WNG}_{tgt}(\omega)$ are the target DF and WNG for frequency ω. The NONULA model is trained by minimizing (15) with gradient descent.

5. EXPERIMENTAL RESULTS

We use two baseline models in our experiments. The first one is a PSO-based endfire non-uniform linear array [21]. To have a fair comparison, we implemented this model without distortion in the look direction [29]. Henceforward we refer to this implementation as 'PSO DMA'. The other baseline models is the uniform DMA [26]. In particular, the uniform DMA is derived from (13) and it employs the second-order supercardioid beampattern. The two null directions of are 106° and 153°. The uniform DMA filters are also utilized as the initial filter input for NONULA. The initial spacing input is the uniform spacing given L. $\eta_{\min}$ is set to 1 cm.

To compare NONULA's scores with PSO DMA, we conduct two series of experiments. $\mathrm{DF}_{tgt}(\omega)$ is set to 13 dB to encourage

Table 1. Mean scores of NONULA and PSO DMA when $r_w = 1$

L \ M	8		10		12	
	NN	PSO	NN	PSO	NN	PSO
13	16.51	12.50	17.46	13.45	18.21	14.25
18	16.68	13.44	18.63	14.98	19.50	16.25
23	16.77	13.95	18.60	15.89	20.14	17.18
28	17.01	14.25	18.77	16.15	20.26	17.81
33	16.87	14.44	18.88	16.67	20.30	18.11

Table 2. Mean scores of NONULA and PSO DMA when $r_w = 0.5$

L \ M	8		10		12	
	NN	PSO	NN	PSO	NN	PSO
13	12.60	11.31	13.12	11.86	13.49	12.30
18	12.83	11.65	14.33	12.95	14.89	13.70
23	12.81	11.96	14.26	13.28	15.40	14.28
28	12.94	11.99	14.43	13.47	15.60	14.63
33	13.03	12.18	14.41	13.66	15.70	14.82

NONULA to outperform the uniform DMA. When $r_w = 1$, we set $\mathrm{WNG}_{tgt}(\omega)$ to 0 dB. This allows NONULA to focus on both WNG and DF. When $r_w = 0.5$, we set $\mathrm{WNG}_{tgt}(\omega)$ to -10 dB so that NONULA can have a DF-centered design. NONULA and PSO DMA are optimized with the same settings in frequency bands from 100 Hz to 8 kHz.

The detailed comparison between NONULA and PSO DMA are demonstrated in Table 1 and 2. There are five choices for array length L: 13, 18, 23, 28 and 33 cm. There are also three choices for the number of microphones M: 8, 10 and 12. The table entries are scores calculated from (9) and averaged across frequency bins. Each column in tables is split into two sub-columns. The left sub-column shows the mean scores for NONULA, whereas the right sub-column for PSO DMA. We can observe that the mean scores of NONULA are always higher than its corresponding mean scores of PSO DMA. In the table, NONULA is dubbed as NN and PSO DMA is dubbed as PSO for brevity.

The optimized array geometry from NONULA and PSO DMA with two different settings are illustrated in Fig. 2 and 4. In Fig. 2, the geometry optimized by NONULA has a more complicated layout, whereas the geometry designed by PSO DMA forms only three sub-arrays. This could be due to the fact that neural networks have more power for optimization. With less microphones and smaller array length, the complexity levels of geometry from NONULA and PSO DMA are similar in Fig. 4. Notably, the subarrays in both NONULA and PSO DMA have internal spacing around 0.02 m. This corresponds with the half wavelength λ of the highest frequency we are interested in, which is $\lambda = 340/(8000 \times 2) = 0.02125$ m. By having the spacing of sub-arrays less than λ, the whole microphone array can effectively avoid spatial aliasing in high frequency bins. Moreover, forming nested arrays with subarrays helps handle broadband signals in both high and low frequency bins.

To have a better understanding of how NONULA's behaviour changes frequency-wise with respect to other techniques, we plot Fig. 3 and 5 regarding the performance of DF and WNG. We observe that NONULA has the best overall DF values while maintaining a good level of WNG values. For the DF-centered design in Fig. 5, NONULA has a smarter tradeoff between DF and WNG

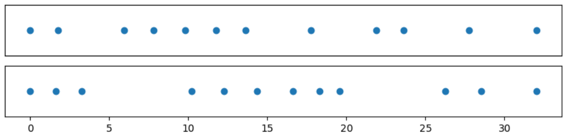

Fig. 2. Optimized array geometry of NONULA (top) and PSO DMA (bottom), when $r_w = 1$, $M = 12$ and $L = 0.33$.

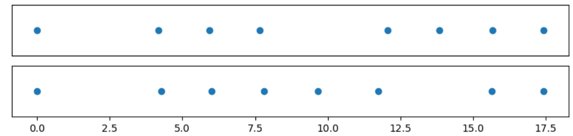

Fig. 4. Optimized array geometry of NONULA (top) and PSO DMA (bottom), when $r_w = 0.5$, $M = 8$ and $L = 0.18$.

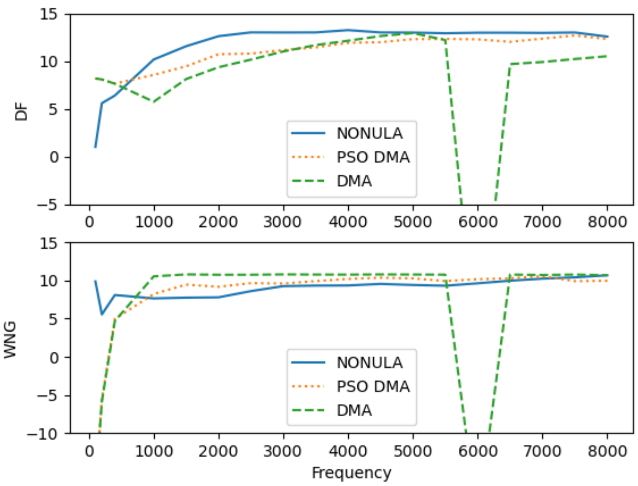

Fig. 3. Performance of NONULA compared with other techniques when $r_w = 1$, $M = 12$ and $L = 0.33$.

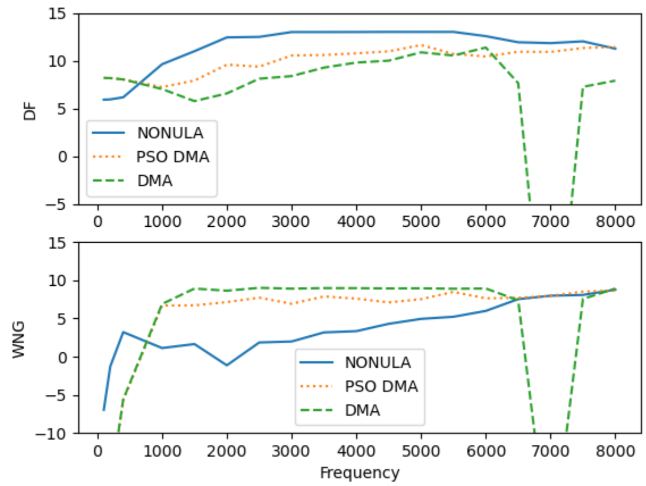

Fig. 5. Performance of NONULA compared with other techniques when $r_w = 0.5$, $M = 8$ and $L = 0.18$. NONULA has a smarter tradeoff between DF and WNG when it is DF-centered.

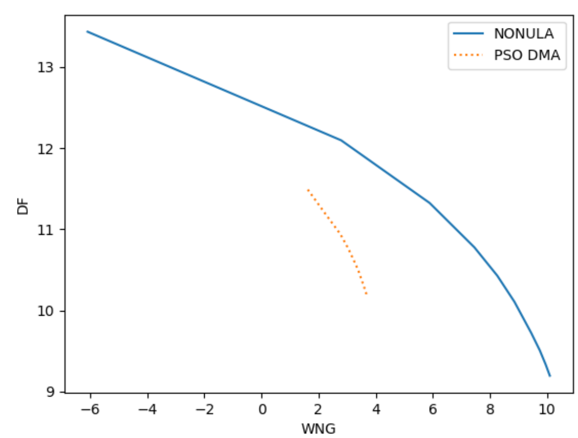

Fig. 6. Pareto fronts of NONULA and PSO DMA when $M = 12$ and $L = 0.33$. Each pair of DFs and WNGs is averaged across frequency bins. Models are trained at the Mel scale.

and achieves better DF performance. This characteristic is desirable when customized performance is required. PSO DMA cannot enhance its WNG values in low frequency bins to the desired level, whereas NONULA never goes below the target WNG value. The uniform DMA also suffers from low WNG values in low frequency bins and it is unstable at some frequency bins.

Since both NONULA and PSO DMA perform multi-objective optimization, we plot Pareto fronts in Fig. 6 to better visualize their performance. We use the negative scores as the training loss without penalties. Both NONULA and PSO DMA are retrained at the Mel scale, which better reflects perceptual distance in human hearing. The Pareto front consists of the set of non-dominated DF and WNG pairs. Each pair represents mean DF and WNG values of all frequency bins with a fixed WNG weight r_w. Different pairs are obtained by varying r_w from 0.1 to 1. Fig. 6 shows that the NONULA Pareto front is above the PSO DMA one. This indicates that a better solution always exists in NONULA when fixing one DF or WNG value from PSO DMA. Moreover, the NONULA curve spans with larger range. This illustrates that NONULA is more flexible in the tradeoff between DF and WNG. The performance of PSO DMA is limited to a small range of values.

6. CONCLUSION

This paper proposed a neural network based model that could optimize both linear array geometry and fixed beamforming in an end-to-end fashion. Compared with DMA-based techniques, our approach exhibited superior robustness in the low frequency region. The proposed model outperformed PSO DMA consistently in various settings with overall better scores while offered a more flexible tradeoff between DF and WNG. This shows that neural networks are more powerful tools for linear microphone array optimization.

7. REFERENCES

[1] Barry D Van Veen and Kevin M Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE assp magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[2] Harry L Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*, John Wiley & Sons, 2004.

[3] Wei Liu and Stephan Weiss, *Wideband beamforming: concepts and techniques*, John Wiley & Sons, 2010.

[4] Matt O'Connor, W Bastiaan Kleijn, and Thushara Abhayapala, "Distributed sparse mvdr beamforming using the bialternating direction method of multipliers," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 106–110.

[5] Fei Ma, Thushara D Abhayapala, and Prasanga N Samarasinghe, "A time-domain nearfield frequency-invariant beamforming method," *arXiv preprint arXiv:2105.08219*, 2021.

[6] Marco Crocco and Andrea Trucco, "Stochastic and analytic optimization of sparse aperiodic arrays and broadband beamformers with robust superdirective patterns," *IEEE Transactions on audio, speech, and language processing*, vol. 20, no. 9, pp. 2433–2447, 2012.

[7] Jacob Benesty and Jingdong Chen, *Study and design of differential microphone arrays*, vol. 6, Springer Science & Business Media, 2012.

[8] Jacob Benesty, Jingdong Chen, and Chao Pan, *Fundamentals of differential beamforming*, Springer, 2016.

[9] Longfei Yan, Weilong Huang, W. Bastiaan Kleijn, and Thushara D. Abhayapala, "Phase error analysis for first-order linear differential microphone arrays," in *2022 17th International Workshop on Acoustic Signal Enhancement (IWAENC)*. IEEE, 2022.

[10] Hans Schjær-Jacobsen and Kaj Madsen, "Synthesis of nonuniformly spaced arrays using a general nonlinear minimax optimization method," *IEEE Transactions on Antennas and Propagation*, vol. 24, no. 4, pp. 501–506, 1976.

[11] SS Shanan and CA Pomalaza-Raez, "The use of nonuniform element spacing in array processing algorithms," *The Journal of the Acoustical Society of America*, vol. 86, no. 4, pp. 1416–1418, 1989.

[12] Jingjing Yu and Kevin D Donohue, "Performance for randomly described arrays," in *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. IEEE, 2011, pp. 269–272.

[13] Yuval Konforti, Israel Cohen, and Baruch Berdugo, "Array geometry optimization for region-of-interest broadband beamforming," in *2022 17th International Workshop on Acoustic Signal Enhancement (IWAENC)*. IEEE, 2022.

[14] Gary W Elko, "Superdirectional microphone arrays," in *Acoustic signal processing for telecommunication*, pp. 181–237. Springer, 2000.

[15] Gary W Elko, "Differential microphone arrays," in *Audio signal processing for next-generation multimedia communication systems*, pp. 11–65. Springer, 2004.

[16] Jens Meyer, "Beamforming for a circular microphone array mounted on spherically shaped objects," *The Journal of the Acoustical Society of America*, vol. 109, no. 1, pp. 185–193, 2001.

[17] Gongping Huang, Jacob Benesty, and Jingdong Chen, "On the design of frequency-invariant beampatterns with uniform circular microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 5, pp. 1140–1153, 2017.

[18] Gongping Huang, Jacob Benesty, and Jingdong Chen, "Design of robust concentric circular differential microphone arrays," *The Journal of the Acoustical Society of America*, vol. 141, no. 5, pp. 3236–3249, 2017.

[19] Gongping Huang, Jingdong Chen, and Jacob Benesty, "Insights into frequency-invariant beamforming with concentric circular microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 12, pp. 2305–2318, 2018.

[20] James Kennedy and Russell Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*. IEEE, 1995, vol. 4, pp. 1942–1948.

[21] Sohel J Patel, Steven L Grant, Maciej Zawodniok, and Jacob Benesty, "On the design of optimal linear microphone array geometries," in *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*. IEEE, 2018, pp. 501–505.

[22] Jilu Jin, Gongping Huang, Jingdong Chen, and Jacob Benesty, "Design of optimal linear differential microphone arrays based array geometry optimization," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 5741–5745.

[23] Jacob Benesty, Jingdong Chen, and Yiteng Huang, *Microphone array signal processing*, vol. 1, Springer Science & Business Media, 2008.

[24] EN Gilbert and SP Morgan, "Optimum design of directive antenna arrays subject to random variations," *Bell System Technical Journal*, vol. 34, no. 3, pp. 637–663, 1955.

[25] Giorgio Chiandussi, Marco Codegone, Simone Ferrero, and Federico Erminio Varesio, "Comparison of multi-objective optimization methodologies for engineering applications," *Computers & Mathematics with Applications*, vol. 63, no. 5, pp. 912–942, 2012.

[26] Jingdong Chen, Jacob Benesty, and Chao Pan, "On the design and implementation of linear differential microphone arrays," *The Journal of the Acoustical Society of America*, vol. 136, no. 6, pp. 3097–3113, 2014.

[27] Chiheb Trabelsi, Olexa Bilaniuk, Ying Zhang, Dmitriy Serdyuk, Sandeep Subramanian, Joao Felipe Santos, Soroush Mehri, Negar Rostamzadeh, Yoshua Bengio, and Christopher J Pal, "Deep complex networks," in *International Conference on Learning Representations*, 2018.

[28] Maxime W Matthès, Yaron Bromberg, Julien de Rosny, and Sébastien M Popoff, "Learning and avoiding disorder in multimode fibers," *Physical Review X*, vol. 11, no. 2, pp. 021060, 2021.

[29] Lester James V. Miranda, "PySwarms, a research-toolkit for Particle Swarm Optimization in Python," *Journal of Open Source Software*, vol. 3, 2018.