

PAPER

Strategies for DOA-DNN Estimation Accuracy Improvement at Low and High SNRs*

Daniel Akira ANDO^{†a)}, *Student Member*, Toshihiko NISHIMURA[†], *Senior Member*, Takanori SATO[†], *Member*, Takeo OHGANE^{†b)}, Yasutaka OGAWA[†], *Fellows*, and Junichiro HAGIWARA^{††}, *Member*

SUMMARY Implementation of several wireless applications such as radar systems and source localization is possible with direction of arrival (DOA) estimation, an array signal processing technique. In the past, we proposed a DOA estimation method using deep neural networks (DNNs), which presented very good performance compared to the traditional root multiple signal classification (root-MUSIC) algorithm when the number of radio wave sources is two. However, once three radio wave sources are considered, the performance of that proposed DNN decays especially at low and high signal-to-noise ratios (SNRs). In this paper, mainly focusing on the case of three sources, we present two additional strategies based on our previous method and capable of dealing with each SNR region. The first, which supports DOA estimation at low SNRs, is a scheme that makes use of principal component analysis (PCA). By representing the DNN input data in a lower dimension with PCA, it is believed that the noise corrupting the data is greatly reduced, which leads to improved performance at such SNRs. The second, which supports DOA estimation at high SNRs, is a scheme where several DNNs specialized in radio waves with close DOA are accordingly selected to produce a more reliable angular spectrum grid in such circumstances. Finally, in order to merge both ideas together, we use our previously proposed SNR estimation technique, with which appropriate selection between the two schemes mentioned above is performed. We have verified the superiority of our methods over root-MUSIC and our previous technique through computer simulation when the number of sources is three. In addition, brief discussion on the performance of these proposed methods for the case of higher number of sources is also given.

key words: antenna array, DOA estimation, deep neural network, principal component analysis

1. Introduction

Direction of arrival (DOA) estimation is a very known array signal processing that is extremely important for many wireless applications. One of the most traditional techniques for DOA estimation is the super-resolution multiple signal classification MUSIC/root-MUSIC algorithm [1], [2]. However, this algorithm being classified as spectral-based, it requires the spectral decomposition of the correlation matrix of the antenna array received signal, which makes its online use

prohibitive as the array dimension increases. Therefore, investigation of new approaches for DOA estimation, such as deep learning, is a trending research topic.

Deep learning applied to wireless communication problems is receiving much attention from the industry and academia, since performance of such data-driven techniques can greatly surpass traditional model-based techniques [3]. Although offline training of deep neural networks (DNNs) can be computationally costly, once training is finalized, DNNs can be easily deployed online to the specific situation for which they were trained. The complexity of such online implementation of DNNs is also thought to be comparatively light due to the fact that most of this computation relies on matrix multiplication. In fact, several studies, such as [4]–[7], have reported great results from the implementation of deep learning in DOA estimation. In [4], a framework for end-to-end channel and DOA estimation in the context of massive multiple-input multiple-output (massive MIMO) is proposed. In [5], a combination of a detection and DOA estimation network, which reduces the training-set size and makes it possible to train several DNNs corresponding to different position sectors, is presented. In [6], a low-complexity DOA estimation technique for hybrid MIMO systems with uniform circular array at a base station is presented. In [7], a DOA estimation system which is robust to array imperfections is explained. Our research group also tackled this problem in [8]–[11], where we have demonstrated great DOA estimation performance.

Principle component analysis (PCA) is an algorithm used to represent the information contained in a higher dimensional data in a lower dimensional space while keeping intact as much of this information as possible. It is heavily used in areas such as data compression, image analysis, visualization, pattern recognition, regressions, etc. It has been verified that PCA is a very effective technique in order to enhance the performance of machine learning models at the same time that it reduces the number of features in the data, which simplifies these models greatly [12]–[15]. Yet, most studies take advantage of PCA in areas such as image classification, such as [13]. In the DOA estimation field, PCA was also used in [14], [15]. In [14], a PCA-like unsupervised neural network is used to reduce the dimensionality of the training dataset generated from a broadband acoustic signal emitted by a low-altitude and high-subsonic flight target. The authors verified that the performance of their 2-dimensional DOA estimation technique surpasses that of root-MUSIC at

Manuscript received December 26, 2023.

Manuscript revised May 9, 2024.

Manuscript publicized July 18, 2024.

[†]Graduate School/Faculty of Information Science & Technology, Hokkaido University, Sapporo-shi, 060-0814 Japan.

^{††}Faculty of Social Informatics, Mukogawa Women's University, Nishinomiya-shi, 663-8137 Japan.

*A part of this paper was presented at International Symposium on Communications and Information Technologies (ISCIT 2023) [19].

a) E-mail: dakiraando@m-icl.ist.hokudai.ac.jp

b) E-mail: ohgane@m-icl.ist.hokudai.ac.jp

DOI: 10.23919/transcom.2023EBP3217

lower SNRs. In [15], a 1-dimensional narrowband DOA estimation with K-nearest neighbors algorithm is proposed, where PCA is applied to the training dataset in order to reduce the computational complexity of this machine learning algorithm and to remove noise from the signal data. The authors also verified that the performance of their method greatly surpasses root-MUSIC at lower SNRs.

The aim of this paper is to improve the DNN’s DOA estimation accuracy, and our contributions are:

- Proposing a method to improve DOA estimation performance at low SNR, which consists of application of PCA to the DNN training, validation and test datasets;
- Proposing a method to improve DOA estimation performance at high SNR, which consists of several separately trained DNNs specialized in radio waves with close DOA;
- Realizing a system binding together the two methods above in order to develop a full technique for DOA estimation at any SNR.

In terms of the first method with PCA, previous studies [14], [15] have not thoroughly evaluated the effects of applying PCA specifically to a DNN input dataset. Therefore, we also give here an extensive and detailed explanation on the effects of PCA at different simulation settings, such as varying a) sizes of the antenna array, b) number of principal components chosen as the new dimension of the input dataset, and c) test SNRs. Note that, in this study, we consider Probabilistic PCA. This is a consequence of the Scikit-learn framework [16], implemented during our numerical simulations, which is based on it. In addition, use of the abbreviation “PCA” will be maintained to keep the notation uncluttered. The second method is based on the observation that, when 3 radio sources are considered, many incorrect estimation cases at higher SNRs are due to radio waves closely impinging onto the antenna array within 20°. Consequently, we offline train different DNNs that are each specialized in close waves impinging at specific regions of the angular spectrum. These new DNNs are then used instead of the conventional DNN, which are expected to produce a more reliable narrow DOA spectrum grid for subsequent DOA detection. Lastly, we use one of our previous strategies for SNR estimation [11] to merge these two proposed methods into a full system capable of operating online while greatly surpassing the performance of our previous technique [9] and root-MUSIC.

Our main study goal in this paper is when there are 3 radio wave sources. As stated in [17], 3 flying objects in the field of view of an antenna array is a possible scenario in air-to-air emitter location or radar systems. Moreover, as verified in [18], at sub-terahertz and line-of-sight indoor office environments of 140 GHz, the average number of subpaths (or multipaths) is significantly small, e.g. mostly ranging between 2 and 5. Therefore, our 3 sources consideration is not only realistic in airborne radar applications based on [17], but also it is a first step towards the goal of radio propagation measurements at sub-terahertz bands [18]. Moreover, it was concluded in [11] that we needed to solve the above men-

tioned issues (i.e. poor estimation performance at lower and higher SNRs), which arise when the number of sources is simply raised from 2 to 3. However, we also give here a brief discussion on the performance of the proposed methods for the case of 4 and 5 number of sources.

The remainder of this paper is organized as follows. The antenna array model is explained in Sect. 2. Our previous works [9], [19] are detailed in Sect. 3. Our proposed techniques based on these works are presented in Sect. 4. Then, in Sect. 5, we validate our proposed methods through computer simulations while using our past technique and root-MUSIC as benchmark. Lastly, in Sect. 6 our work is concluded.

2. Antenna Array Model

Let there be K radio wave sources located in the far-field region of a uniform linear array (ULA) consisting of L omnidirectional antennas with no mutual coupling and spaced at half-wavelength. These sources are emitting narrowband waves whose planar wavefronts impinge onto the ULA at angles $\theta[\text{degrees}] = [\theta_1, \dots, \theta_K]^T$ at least 1° apart, where $[\cdot]^T$ indicates the transpose operator. Then, the baseband received signal $\mathbf{x}(t) \in \mathbb{C}^{L \times 1}$ can be modeled by

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{z}(t), \tag{1}$$

where $\mathbf{s}(t) \in \mathbb{C}^{K \times 1}$ is the vector containing the incident radio waves’ complex amplitudes, $\mathbf{z}(t) \in \mathbb{C}^{L \times 1}$ is the additive white Gaussian noise vector following a circular complex Gaussian distribution $\mathbf{z}(t) \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_L)$ with zero mean and variance σ^2 , where \mathbf{I}_L represents an L -dimensional identity matrix, and $\mathbf{A}(\boldsymbol{\theta})$ is the mode matrix, which accounts for the relative phase delay corresponding to path length difference of the incident waves on each ULA element and is described as

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{bmatrix} 1 & \dots & 1 \\ e^{-j\pi \sin \theta_1} & \dots & e^{-j\pi \sin \theta_K} \\ e^{-j\pi 2 \sin \theta_1} & \dots & e^{-j\pi 2 \sin \theta_K} \\ \vdots & \ddots & \vdots \\ e^{-j\pi(L-1) \sin \theta_1} & \dots & e^{-j\pi(L-1) \sin \theta_K} \end{bmatrix}. \tag{2}$$

Furthermore, the radio waves are assumed to be uncorrelated and received with equal power normalized to one.

For many DOA estimation techniques, the estimated correlation matrix $\hat{\mathbf{R}}_{xx}$ of the received signal is usually used, where this can be calculated by the equation below:

$$\hat{\mathbf{R}}_{xx} = \frac{1}{N_{\text{snap}}} \sum_{n=1}^{N_{\text{snap}}} \mathbf{x}(t_n)\mathbf{x}(t_n)^H, \tag{3}$$

where N_{snap} is the total number of snapshots, $\mathbf{x}(t_n)$ represents the n th snapshot taken from the received signal, and $(\cdot)^H$ is the conjugate transpose operator.

3. Authors' Previous Work

3.1 Input and Output Definitions of DNN

The generation procedure of the DNN datasets is described here. Note that these are the original datasets prior to dimensionality reduction through PCA, where they consist of input $\mathbf{u} = \{\mathbf{u}_1, \dots, \mathbf{u}_N\}$ and target vectors $\mathbf{t} = \{\mathbf{t}_1, \dots, \mathbf{t}_N\}$. Here, N is the number of samples, $\mathbf{u}_i \in \mathbb{R}^{D_{\text{in}} \times 1}$ and $\mathbf{t}_i \in \mathbb{R}^{D_{\text{out}} \times 1}$ for $i = 1, \dots, N$ are the i th samples of the input and target vectors with D_{in} and D_{out} features, respectively.

(a) Input Layer

Due to its Hermitian nature, the estimated correlation matrix $\hat{\mathbf{R}}_{xx}$ can be written as in (4). Then, a vector \mathbf{u}_i proper for being fed as input to the DNN can be generated as follows: first, we arrange the diagonal elements of (4) in the first entries of the input vector; next, we take the real $\Re(\cdot)$ and imaginary $\Im(\cdot)$ parts of each lower triangular element column by column and from left to right, subsequently arranging these in the remaining space of the input vector (See (5)). The upper triangular elements can be ignored due to the fact that they are simply the complex conjugate of the lower triangular elements.

$$\hat{\mathbf{R}}_{xx} = \begin{bmatrix} r_{11} & r_{21}^* & \cdots & r_{L1}^* \\ r_{21} & r_{22} & \cdots & r_{L2}^* \\ \vdots & \vdots & \ddots & \vdots \\ r_{L1} & r_{L2} & \cdots & r_{LL} \end{bmatrix} \quad (4)$$

$$\mathbf{u}_i = [r_{11}, \dots, r_{LL}, \Re(r_{21}), \Im(r_{21}), \dots, \Re(r_{L1}), \Im(r_{L1}), \dots, \Re(r_{L(L-1)}), \Im(r_{L(L-1)})]^T. \quad (5)$$

The resultant input vector $\mathbf{u}_i \in \mathbb{R}^{D_{\text{in}} \times 1}$ has $D_{\text{in}} = L^2$ features, where each of them corresponds to each unit of the DNN input layer.

(b) Output Layer

We design the DNN output layer in such a way that the DNN should produce an angular spectrum discretized in angle bins, where each of these covers a portion of the spectrum. Therefore, each unit of the DNN output layer corresponds to each angle bin. In this study, an angle spectrum ranging from -60° to $+60^\circ$ is considered. When this spectrum is discretized in steps of 1° , the total number of angle bins (and thus the number of features D_{out}) becomes 121.

Since the DNN output units represent the probability of incident radio wave onto the corresponding angle bins, the target vector $\mathbf{t}_i = [t_1, \dots, t_j, \dots, t_{D_{\text{out}}}]$ can be generated following (6) below.

$$t_j = \begin{cases} 1 & \text{if wave is incident onto the } j\text{th bin} \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

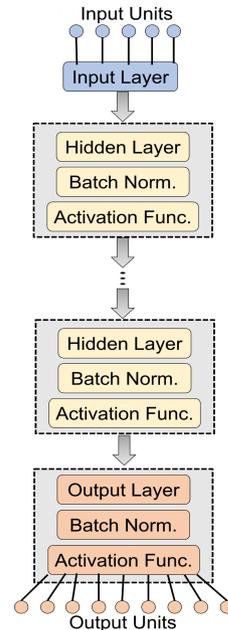


Fig. 1 DNN structure with input, hidden, and output layers. Reprinted from [19] (©2023 IEEE).

where the j th angle bin covers the spectrum region from $j - 61.5^\circ$ to $j - 60.5^\circ$.

3.2 DNN for DOA Estimation

A traditional feed-forward neural network, whose structure consists of an input layer, an output layer and an arbitrary number of hidden layers, is used (Fig. 1). In addition, we insert the batch normalization regularizer [20] in all layers of the DNN in order to improve the overall stability of the learning process. The activation functions for the hidden layers and for the output layer are the rectified linear unit (ReLU) and Sigmoid, respectively. By using Sigmoid as an activation function, we guarantee that the DNN produces output values that can be regarded as probabilities ranging from 0.0 to 1.0. During the learning phase, the DNN weights are updated in accordance to the Adam optimization [21].

In this work, we investigate mainly two performance metrics: the probability of correct DOA estimation and the root mean squared error (RMSE), where the former is verified during the validation and test phases, and the latter only during the test phase. The probability of correct DOA estimation is calculated as the ratio of the number of correct DOA estimation samples over the total number of evaluated samples, where DOA estimation is only counted as correct when the absolute error of all the DOA estimates $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \dots, \hat{\theta}_K]^T$ within a sample is below a certain estimation tolerance error:

$$\text{Correct DOA} \iff |\theta_j - \hat{\theta}_j| \leq \mu, \forall j \in \{1, \dots, K\}, \quad (7)$$

where μ is the estimation tolerance, considered to be 0.5° here (verification whether the estimated DOA is within the

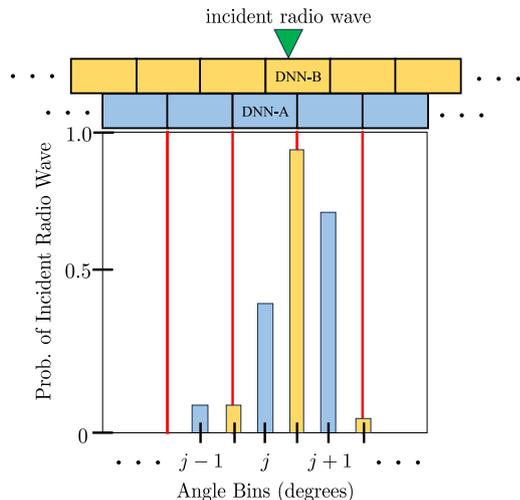


Fig. 2 An example of the Staggered-DNN output, where the radio wave is incident near the right border of the j angle bin of DNN-A. In this example, the $j + 1$ bin is mistakenly detected. Cases such as this are one of the verified causes generally leading to incorrect DOA estimation. However, after combining both DNN-A and DNN-B grids, correct DOA detection becomes possible. Reprinted from [19] (©2023 IEEE).

1° -width angle bin). The RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{KN_t} \sum_{k=1}^K \sum_{n=1}^{N_t} (\hat{\theta}_k^{(n)} - \theta_k^{(n)})^2}, \quad (8)$$

where N_t is the total number of test samples. During the validation phase, the DNN weights corresponding to the highest probability of correct DOA estimation are saved for subsequent use at the test phase, where this is done in an effort to avoid overfitting. However, note that the saved weights are not necessarily optimal in terms of RMSE.

Previously we verified that incidence of radio waves onto the vicinity of the angle bin border generally results in incorrect DOA detection due to wrongful excitement of neighboring bins (Fig. 2), thus causing significant decline in overall performance. In [9] we proposed a strategy to cope with such cases. This relies on the training of one additional support DNN (called DNN-B) whose angle grid is stacked up on top of that of the main DNN (DNN-A), where the DNN-B angle grid is shifted by 0.5° with respect to that of DNN-A, thus ranging from -60.5° to $+60.5^\circ$ (totaling 122 angle bins). This strategy was then named Staggered DNN, and we have demonstrated that the spectrum contribution provided with DNN-B enhances the estimation accuracy around the bin borders of the DNN-A bins. Both DNNs are offline trained separately with the same input dataset, but with accordingly modified target datasets reflecting the corresponding angle bin grid. Then, during the test phase, the spectrum grid produced with both DNNs are merged as it is shown in Fig. 2, resulting in a combined angular spectrum grid. Lastly, a DOA detection algorithm is applied on this resultant grid so as to extract the DOA estimates.

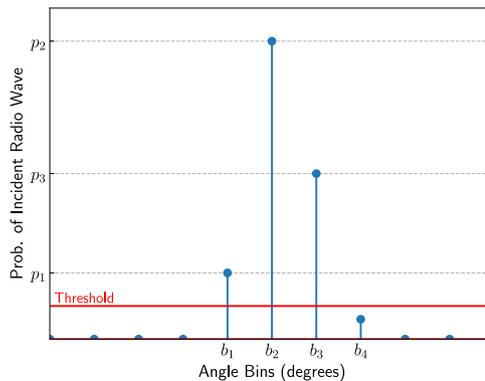


Fig. 3 Illustration of the usage of “Neighbors Weighted Average” on a sample of tested DNN output, assuming only one DOA. For the computation of the DOA estimate, all bins whose probability of incident radio wave is below a certain threshold are ignored (e.g. bin b_4). Reprinted from [19] (©2023 IEEE).

3.3 DOA Detection Algorithm

After calculating the DNN output, or equivalently the angular grid spectrum, it is still necessary to recover the DOA information contained in it. A detection algorithm called “Neighbors Weighted Average” was presented in [19]. Here, we briefly explain it again, and detail the full algorithm in Appendix A.

Various DNN outputs are normally contaminated by spurious bins in the vicinity of those corresponding to true DOA bins. However, by taking advantage of such bins, we managed to develop an algorithm capable of detecting more accurate DOAs than if we had simply chosen the most likely bin by means of, for instance, peak search.

Figure 3 illustrates a DNN output example for the case of only one radio wave. Although no proper optimization procedure has been performed, we have verified that very accurate DOA estimation is possible when the threshold value (straight red line in Fig. 3) is 0.1. Then, the DOA estimate $\hat{\theta}$ can be calculated as:

$$\hat{\theta} = \sum_{i=1}^3 p_i b_i \bigg/ \sum_{i=1}^3 p_i \quad (9)$$

This method has proven to be powerful when there are K clearly distinguished hills of angle bins (for instance, there is only one hill in Fig. 3). On the other hand, the full algorithm described in Appendix is capable of dealing with other cases of less ideal angular spectrum grid.

4. Proposed Strategies for Accuracy Enhancement

4.1 Lower SNRs: Staggered DNN-PCA

The flow chart of the proposed technique for accuracy enhancement at lower SNRs can be seen in Fig. 4. At the training phase, we generate N input samples \mathbf{u}_i for

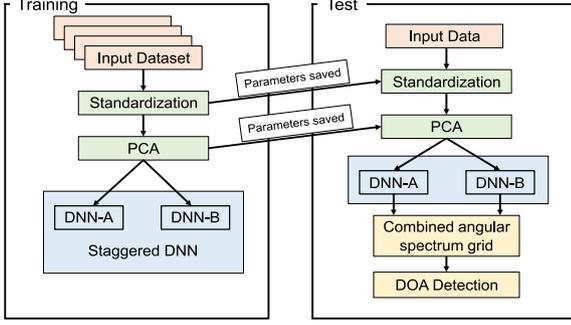


Fig. 4 Flow chart of the proposed Staggered DNN-PCA for lower SNRs, where dimensionality reduction of the DNN input vector is performed with PCA.

$i = 1, \dots, N$ at 30 dB. Then, these are standardized, resulting in $\bar{\mathbf{u}}_i = \Sigma^{-1}(\mathbf{u}_i - \boldsymbol{\mu})$. Here, $\boldsymbol{\mu} \in \mathbb{R}^{D_{in} \times 1}$ and $\Sigma = \text{diag}(\boldsymbol{\sigma}) \in \mathbb{R}^{D_{in} \times D_{in}}$ are the vector and diagonal matrix containing the means and standard deviations, respectively, of each feature of the training dataset, where $\boldsymbol{\sigma} \in \mathbb{R}^{D_{in} \times 1}$ corresponds to the standard deviation vector. By applying PCA to this standardized input dataset we achieve a dimensionality reduction from D_{in} to an arbitrary D_{pca} . In order to derive the PCA parameters (for a more thorough explanation refer to [12]), first the covariance matrix \mathbf{S} of the standardized input dataset must be calculated:

$$\mathbf{S} = \frac{1}{N} \sum_{i=1}^N \bar{\mathbf{u}}_i \bar{\mathbf{u}}_i^T. \quad (10)$$

Then, the eigendecomposition of \mathbf{S} is performed:

$$\mathbf{S} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^{-1}, \quad (11)$$

where $\mathbf{U} \in \mathbb{R}^{D_{in} \times D_{in}}$ and $\boldsymbol{\Lambda} \in \mathbb{R}^{D_{in} \times D_{in}}$ are the matrix containing the eigenvectors of \mathbf{S} and the diagonal matrix containing the corresponding eigenvalues, respectively. After choosing the desired number of dimensions D_{pca} to remain in the new dataset, the parameters \mathbf{M} and \mathbf{W} necessary for dimensionality reduction with PCA can be calculated:

$$\sigma_{pca}^2 = \frac{1}{D_{in} - D_{pca}} \sum_{j=D_{pca}+1}^{D_{in}} \lambda_j, \quad (12)$$

$$\mathbf{W} = \mathbf{U}_{pca} (\boldsymbol{\Lambda}_{pca} - \sigma_{pca}^2 \mathbf{I})^{1/2} \mathbf{R}, \quad (13)$$

$$\mathbf{M} = \sigma_{pca}^2 \mathbf{I} + \mathbf{W}^T \mathbf{W}, \quad (14)$$

where σ_{pca}^2 can be interpreted as the average variance lost per discarded dimension, $\boldsymbol{\Lambda}_{pca} \in \mathbb{R}^{D_{pca} \times D_{pca}}$ is the diagonal matrix of the largest D_{pca} eigenvalues λ_j ($j = 1, \dots, D_{pca}$), $\mathbf{U}_{pca} \in \mathbb{R}^{D_{in} \times D_{pca}}$ is the matrix with the corresponding eigenvectors (or principal components), \mathbf{I} is the $D_{pca} \times D_{pca}$ identity matrix and \mathbf{R} is an arbitrary orthogonal rotation matrix considered to be equal to \mathbf{I} in this study. Finally, the dimension of a sample of the standardized input dataset can be reduced from D_{in} to the chosen D_{pca} by:

$$\mathbf{u}_{pca,i} = \mathbf{M}^{-1} \mathbf{W}^T \bar{\mathbf{u}}_i, \quad (15)$$

where $\mathbf{u}_{pca,i} \in \mathbb{R}^{D_{pca} \times 1}$ is the representation of $\bar{\mathbf{u}}_i$ in a lower dimension, i.e. the projection points of $\bar{\mathbf{u}}_i$ onto the D_{pca} principal components. This new input dataset is then fed to DNN-A and DNN-B for offline training, while the parameters $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$, \mathbf{M} and \mathbf{W} are then saved for later use during the test phase, where new data must go through the same dimensionality reduction process before being fed to the trained Staggered DNN.

Finally, during the test phase, after feeding both DNN-A and DNN-B with a test input sample that went through the PCA process described above, their outputs are combined in order to produce the resultant staggered angular spectrum grid (Sect. 3.2), on which the DOA detection algorithm (Sect. 3.3) is applied.

As it will be shown later in Sect. 5, this strategy provides outstanding accuracy improvement at lower SNRs and number of sources $K = 3$ even when Staggered DNN is trained with a dataset generated at 30 dB. The effectiveness of PCA when K is 4 and 5 is also briefly discussed in Appendix B. As opposed to the noise, which is distributed along all principal components of \mathbf{S} , the bulk of the signal information is believed to be mainly distributed along the first D_{pca} principal components. Therefore, improvement in the data SNR is achieved when the $D_{in} - D_{pca}$ dimensions are discarded, which ultimately results in more precise DOA estimation. On the other hand, accuracy at higher SNRs is deteriorated due to dimensionality reduction due to loss of signal information, which is only lightly corrupted by noise at such SNRs. Quantitative analysis on the effect of PCA on the input vector SNR, which is believed to be related with estimation performance, is left as our future work.

4.2 Higher SNRs: Staggered Narrow Range DNN

After a preliminary assessment, we observed that waves with close DOA are mostly responsible for incorrect DOA estimation especially at higher SNRs. For instance, when $K = 3$ and the SNR is 30 dB, roughly 80% of all cases of unsuccessful estimation (following (7)) are due to waves lying within a range of 20° . Consequently, in order to overcome this issue especially for the case $K = 3^\dagger$, we have designed a new strategy called ‘‘Staggered Narrow Range DNN’’ (Staggered NRDNN) as shown in Fig. 5.

First, DOA estimation is performed in the exact same way that has been described so far (Fig. 5(a)). After applying Neighbors Weighted Average to the Staggered DNN output (Sect. 3.3), the initial DOA estimates $\hat{\boldsymbol{\theta}} = \{\hat{\theta}_1, \dots, \hat{\theta}_i, \dots, \hat{\theta}_K\}$ sorted in ascending order are obtained. If these K estimated DOAs are within the range $\Delta \hat{\theta} = \hat{\theta}_K - \hat{\theta}_1 \leq 20^\circ$, then it is very likely that either one was incorrectly detected. At this point a new and more reliable angular spectrum should be produced. To this end, we train 7 different Staggered NRDNNs (7 NRDNN-As and 7 NRDNN-Bs) offline, each covering a

[†]In Appendix B, we discuss the validity of this technique for higher number of sources K .

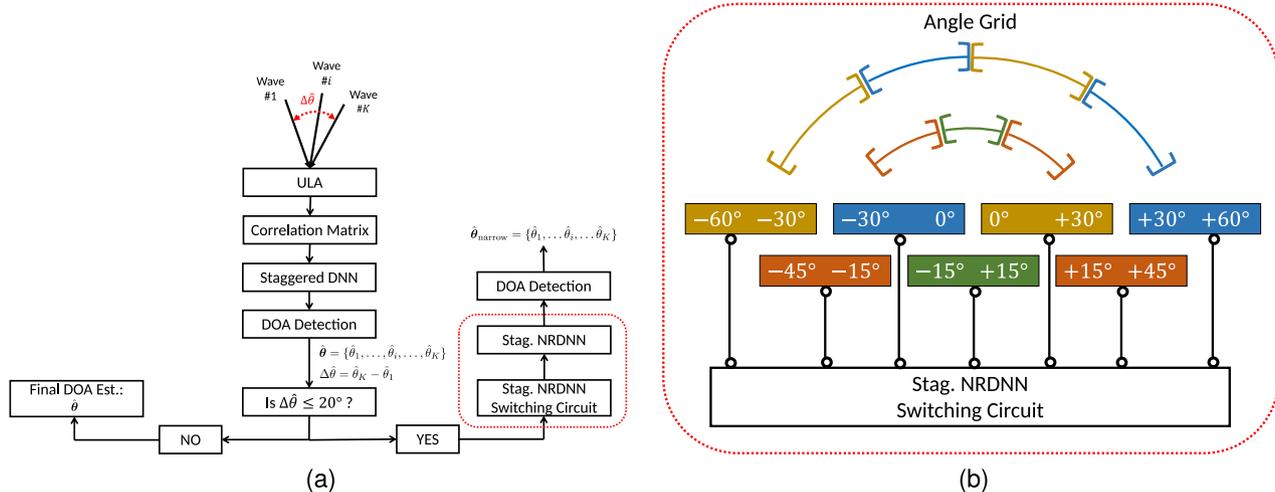


Fig. 5 (a) Flow chart of the proposed Staggered NRDNN strategy for higher SNRs. (b) Visualization of the spectrum grid ranges for which each NRDNN is trained.

predetermined portion of the angle grid (Fig. 5(b)). For instance, there is a Staggered NRDNN covering the angle bins $\{29.5^\circ, 30.0^\circ, 30.5^\circ, \dots, 60.0^\circ, 60.5^\circ\}$. As it will be shown in Sect. 5, this allows us to create Staggered DNNs specialized in different grid regions, thus making it possible to produce more precise angular spectrum for DOA detection. Next, according to the initial DOA estimates, the appropriate Staggered NRDNN covering them is chosen to produce a new spectrum grid, to which again Neighbors Weighted Average is applied to detect the final DOA estimate $\hat{\theta}_{\text{narrow}}$. On the other hand, if the initial DOA estimates do not lie within the 20° range, then they are kept as the final estimates.

Now we describe the training process for these new NRDNNs. The input and target vector remain the same as in (5) and (6), respectively, where no PCA is involved. In contrast to the main DNN-As and DNN-Bs, where the data was generated by randomly selecting the DOAs θ from a uniform distribution between -60.5° and $+60.5^\circ$, the training and validation data of the NRDNNs are generated in such a way that all K DOAs θ lie within a range of $\Delta\theta = \theta_K - \theta_1 \leq 30^\circ$, which in turn is uniformly sampled from $[-60.5^\circ, +60.5^\circ]$. This dataset is then used in all 14 NRDNNs for training and validation. As it will be seen from the simulation results in Sect. 5, good accuracy improvement is achieved even when all NRDNNs are trained with this same dataset. Therefore, there is no need to generate 7 different training datasets corresponding to each specific grid region.

In addition, the precision metric is used during validation in contrast to the probability of correct DOA estimation (Sect. 3.2). The precision of a DNN output is calculated as $n_{TP}/(n_{TP} + n_{FP})$, where n_{TP} and n_{FP} correspond to the number of true positives (DNN angle bin corresponding to a true DOA was excited) and false positives (DNN angle bin where there is no true DOA was excited), respectively. The choice on this metric is due to the impossibility of properly calculating the probability of correct DOA estimation in the considered data generation procedure, where at least 1

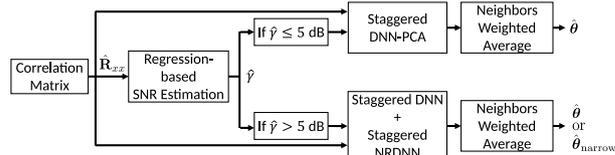


Fig. 6 Flow chart of the proposed full system implemented with Staggered DNN-PCA and Staggered NRDNN with the aid of regression-based SNR estimation [11].

DOA might not lie within the range covered by an NRDNN. Finally, the weights corresponding to the highest precision obtained during the validation of each NRDNN are saved for the test phase.

4.3 Implementation of Full System with the Above Strategies

We have developed two separate strategies for improving the DOA estimation with Staggered DNN at two regions: lower SNRs and higher SNRs. Now it is necessary to bind them together. For this, we use a technique proposed in [11]: Regression-based SNR estimation. The idea is to estimate the SNR $\hat{\gamma}$ from the correlation matrix $\hat{\mathbf{R}}_{xx}$. If $\hat{\gamma} \leq 5$ dB, then Staggered DNN-PCA is applied; otherwise, Staggered NRDNN is used (Fig. 6).

In order to estimate the SNR $\hat{\gamma}$, we observed that this in dB and $-\log(\lambda_s)$ are linearly correlated, where λ_s represents the smallest eigenvalue of the correlation matrix $\hat{\mathbf{R}}_{xx}$. Therefore, we obtain a prediction function for $\hat{\gamma}$ with respect to $-\log(\lambda_s)$ by training a regression model based on the ordinary least squares method. Since both our input ($-\log(\lambda_s)$) and output ($\hat{\gamma}$) data are one-dimensional, the linear function that approximates the desired prediction function has only two coefficients: the y -intercept and the slope. These can be calculated by minimizing the residual sum of squares between the observed and the predicted SNRs. After fitting the

training dataset to this model, the following SNR prediction function when $L = 10$ and $K = 3$ is obtained:

$$\hat{\gamma} \text{ (dB)} = -2.1452 + 9.9923(-\log(\lambda_s)). \quad (16)$$

5. Simulation Results

The performance evaluation of Staggered DNN-PCA (Sect. 4.1), Staggered NRDNN (Sect. 4.2) and the system combining both (Sect. 4.3) is now described here. Training, validation and test datasets must be generated for each Staggered DNN technique and for the regression-based SNR estimation. The parameters for the previously proposed Staggered DNN [9] are shown in Tables 1 and 2. The parameters for the techniques proposed in this paper will be presented in their respective sections. Furthermore, as previously mentioned in Sect. 3, two metrics are used here for performance evaluation: probability of correct DOA estimation and RMSE (the former is calculated in the same fashion as in (7) for root-MUSIC). Their results must be interpreted depending on the type of application. The probability of correct DOA estimation should be of more relevance in such

cases where correct DOA estimation of all incoming waves at a time is vital. On the other hand, in such cases where average precision is more important than occasional detection error, RMSE should be mostly considered. Nevertheless, as the RMSE is very sensitive to outliers, other metrics should be regarded concurrently, such as the absolute error median. As mentioned in Sect. 1, we turn our focus in this section on the simulation results for the case of $K = 3$ radio wave sources. For the cases of 4 and 5 sources, a brief discussion on the simulation results is given in Appendix B as a preliminary evaluation.

5.1 Staggered DNN-PCA

All parameters for Staggered DNN-PCA are kept the same as in Tables 1 and 2, except the number of input layer units, which is equivalent to the number of principal components D_{pca} chosen during the dimensionality reduction process described in Sect. 4.1.

First, in Figs. 7 and 8 we show the probability of correct DOA estimation and RMSE of the proposed Staggered DNN-PCA, respectively, with respect to the number of principal components when the number of antenna elements L is varied from 10 to 15 and when the test dataset was generated at 0, 5, 10 and 20 dB. We compare the performance of the proposed technique with that of root-MUSIC, which is shown as horizontal black straight lines in the said figures.

In Fig. 7, we can promptly see that there is an optimal number of principal components especially with respect to 0 dB, and that this optimal value is different for each L . Moreover, when the SNR is 0 dB and $L \geq 11$, the performance of Staggered DNN-PCA surpasses that of root-MUSIC for certain numbers of principal components; in fact, when $L \geq 12$, an improvement of roughly 10% is achieved in terms of the optimal number of principal components. From the blue and orange curves corresponding to 0 and 5 dB, respectively, we can conclude that estimation precision improvement is possible by reducing the size of the DNN input vector with PCA. We believe that, by applying PCA and only selecting the dimensions corresponding to the largest D_{pca} eigenvalues of the covariance matrix \mathbf{S} (refer to (11)), we manage to strongly reduce the noise corrupting the input vector. On the other hand, when the SNR is 20 dB, not only no visible effect from PCA can be seen, but also the proposed method does not surpass root-MUSIC performance. It is believed that information on the signal only lightly corrupted by noise is lost by applying PCA. Therefore, Staggered DNN-PCA appears to be inefficient at higher SNRs (fact that will be verified later in this section).

In Fig. 8, we can see that the choice on the number of principal components mainly affects the RMSE when the SNR is 0 and 5 dB. It is also visible that too small values of principal components (i.e. $D_{\text{pca}} \leq 6$) impacts the performance at any SNR and L significantly; likewise for too large values ($D_{\text{pca}} \approx 30$) at 0 and 5 dB when L is 10 or 11. This suggests that the new size of the DNN input vector must be chosen after careful analysis as shown in this figure.

Table 1 Parameters for Staggered DNN [9] data generation.

Number of antenna elements L	5–15
Number of incident radio waves K	3
Direction of arrival angles θ	random from −60.5° to +60.5°
Training and validation SNR	30 dB
Test SNR	0–30 dB
Number of snapshots N_{snap}	100
Number of training data samples	250,000
Number of validation data samples	10,000
Number of test data samples N_t	150,000

Table 2 Parameters for Staggered DNN [9] training.

Input layer units D_{in}	L^2
Hidden layers	5
Units per hidden layer	182
Output layer units D_{out}	DNN-A: 121 DNN-B: 122
Hidden layers activation function	ReLU
Output layer activation function	Sigmoid
Loss function	Binary cross-entropy
Optimizer	Adam
Batch size	256
Number of epochs	500

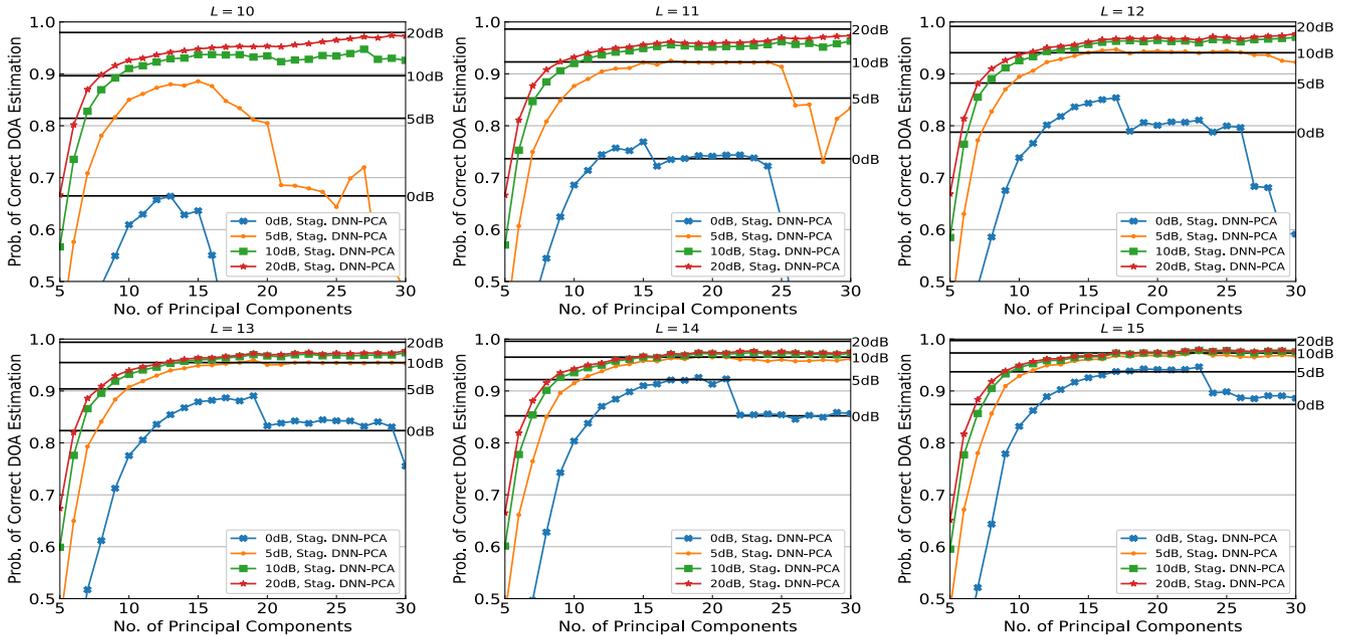


Fig. 7 Comparison of the probability of correct DOA estimation performance of root-MUSIC (solid black lines) and Staggered DNN-PCA for varying number of principal components, or new dimension of input vector after PCA, when the number of antenna elements $L \in [10, 15]$. These methods were tested at 0, 5, 10 and 20 dB.

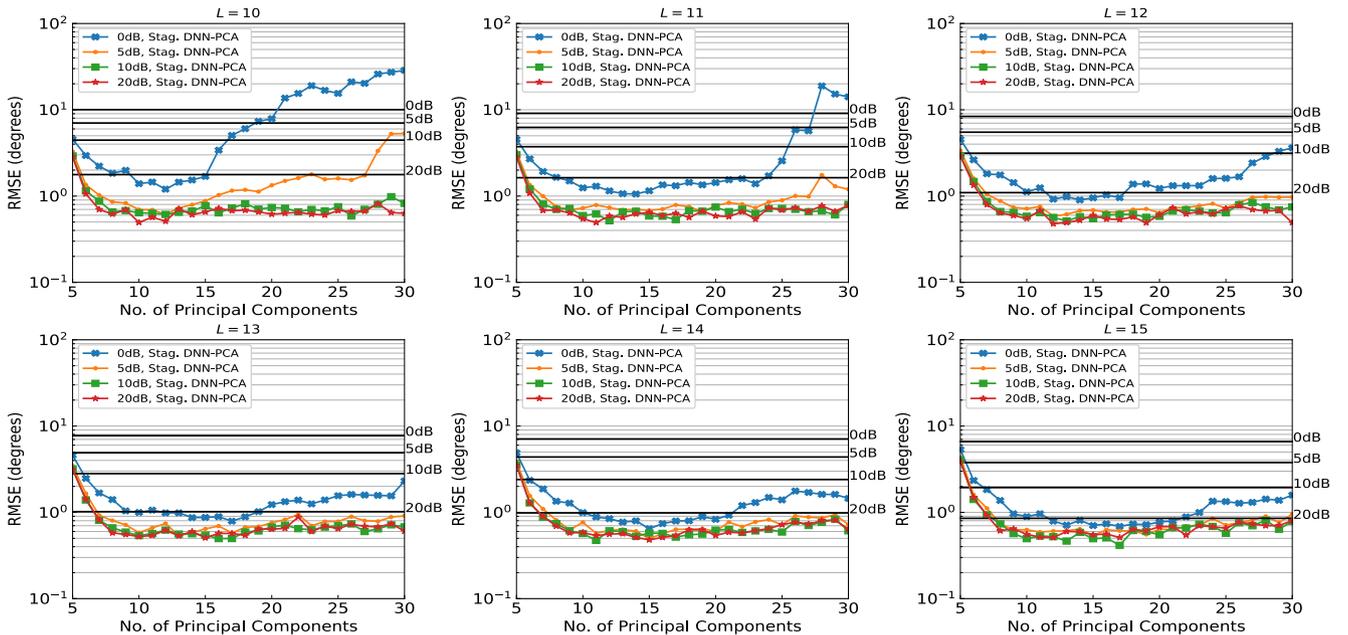


Fig. 8 Comparison of the RMSE performance of root-MUSIC (solid black lines) and Staggered DNN-PCA for varying number of principal components, or new dimension of input vector after PCA, when the number of antenna elements $L \in [10, 15]$. These methods were tested at 0, 5, 10 and 20 dB.

With respect to the RMSE at 0 dB, the optimal number of principal components appears to be slightly different from that corresponding to the probability of correct DOA estimation at each L . This is possibly because the structure of the DNNs (number of hidden layers and units thereof) was optimized in terms of the probability of correct DOA esti-

mation [11], [19], not RMSE. For this reason, we use the optimal value of D_{pca} in terms of the probability of correct DOA estimation (Fig. 7) in the subsequent simulations.

The comparison of the probability of correct DOA estimation and RMSE of Staggered DNN-PCA with those of Staggered DNN and root-MUSIC for varying number of an-

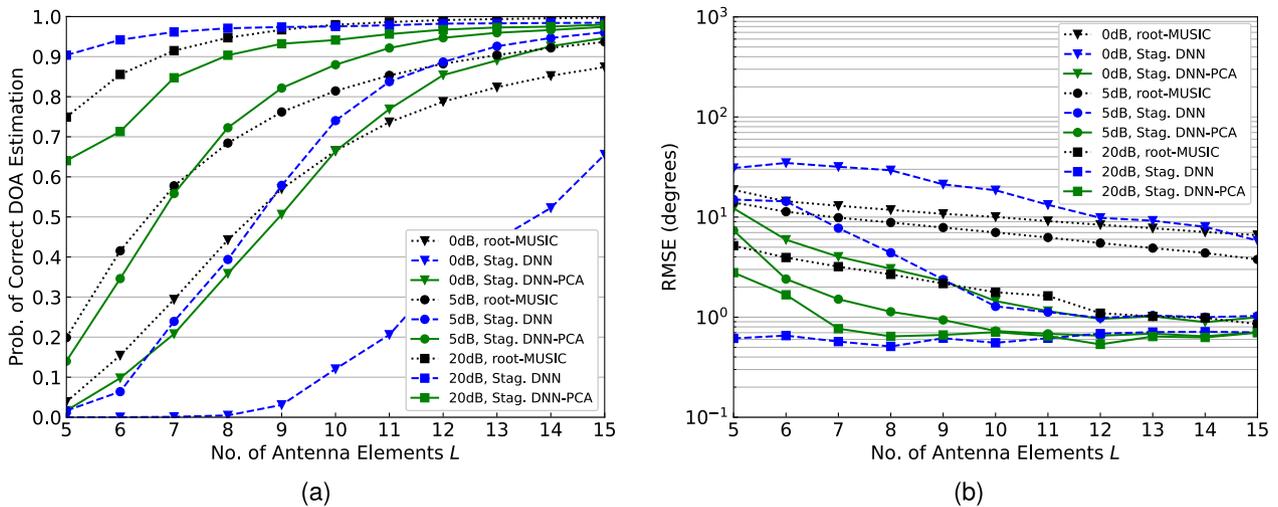


Fig. 9 Performance comparison of 3 DOA techniques while varying the number of antenna elements L : root-MUSIC (dotted black lines), Staggered DNN (dashed blue lines) and Staggered DNN-PCA (solid green lines). These methods were tested at 0, 5 and 20 dB (lower triangle, circle, and square markers, respectively). (a) Probability of correct DOA estimation. (b) RMSE.

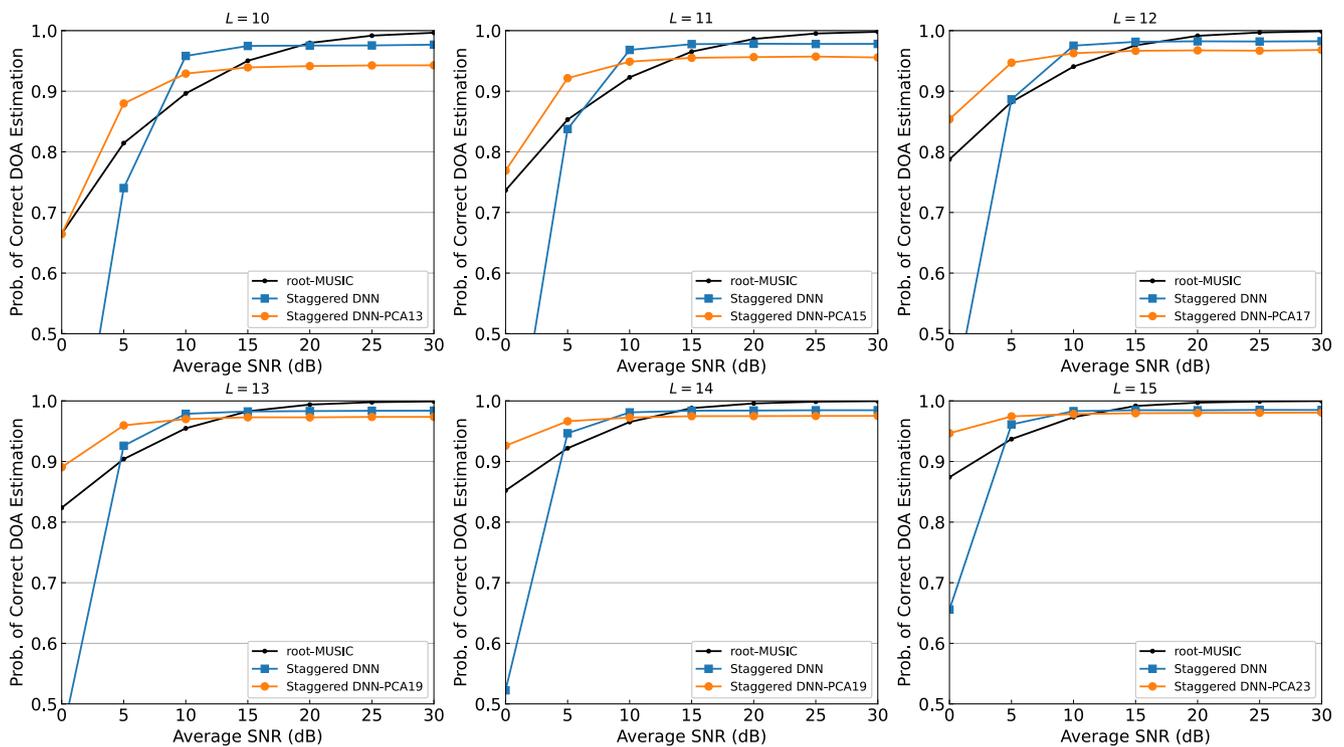


Fig. 10 Comparison of the probability of correct DOA estimation performance of root-MUSIC, Staggered DNN and Staggered DNN-PCAx for varying test SNRs when the number of antenna elements $L \in [10, 15]$. The notation PCAx denotes the number of principal components x chosen.

tenna elements L when the test SNR is 0, 5 and 20 dB is shown in Fig. 9. The number of principal components D_{pca} chosen for each L was the optimal number verified in a graph such as those portrayed in Fig. 7. These values can be found in Table 3, where the dimension reduction $(D_{in} - D_{pca})/D_{in}$ in percentage is also shown. From both figures, the proposed Staggered DNN-PCA is very superior compared to Staggered

DNN when the SNR is 0 and 5 dB. Taking as an example the point where $L = 9$, when the input vector dimension has been reduced by approximately 84% (input vector features from 81 to 13), we managed to improve the probability of correct DOA estimation at 0 dB by 12.5 times (raise from 0.04 to 0.5), and the RMSE by -18 dB (reduction from 20 to 2.5 degrees). On the other hand, again we can see that,

Table 3 Dimension reduction $(D_{in} - D_{pca})/D_{in}$ by PCA.

$D_{in} (= L^2)$	D_{pca}	Dimension Reduction (%)
5^2	8	68.0
6^2	7	80.56
7^2	9	81.63
8^2	11	82.81
9^2	13	83.95
10^2	13	87.0
11^2	15	87.60
12^2	17	88.19
13^2	19	88.76
14^2	19	90.31
15^2	23	89.78

when the SNR is 20 dB, Staggered DNN clearly provides the best performance at any number of antenna elements L . This result indicates once more that PCA does not provide fruitful results at higher SNRs.

Finally, in Fig. 10, we show the probability of correct DOA estimation with respect to the test SNR when L is varied from 10 to 15. When $L \geq 11$, once more it can be seen that not only the proposed Staggered DNN-PCA presents the best performance in all the three methods when the SNR is 0 and 5 dB, but also an input vector dimension reduction of 85% on average (Table 3) is accomplished. In particular the great difference in performance between applying PCA or not should be noted. Even if PCA proves to be ineffective at 10 dB or higher, as the number of antenna elements L increases towards 15, the performance of both Staggered DNNs with and without PCA becomes fairly equal. This could suggest that Staggered DNN-PCA at higher SNRs is more attractive under the condition that L is large. In any case, root-MUSIC still proves to be a stronger algorithm at higher SNRs, given its super-resolution characteristics at high SNR and sufficient large number of snapshots.

5.2 Staggered NRDNN

All parameters for the training of Staggered NRDNN are kept the same as in Tables 1 and 2, except the following:

- Here we only consider $L = 10$;
- The $K = 3$ training and validation DOAs $\theta = \{\theta_1, \theta_2, \theta_3\}$ are generated in a way that they are uniformly distributed within $[\theta_{min}, \theta_{max}]$, where (a) $\theta_{max} - \theta_{min} = 30^\circ$ and (b) this range is randomly sampled from $[-60.5^\circ, +60.5^\circ]$;
- The number of output layer units of NRDNN-A and NRDNN-B are 31 and 32, respectively.

In Fig. 11, an example of one test spectrum grid when the SNR is 20 dB by using Staggered NRDNN is shown. If the spectrum grid from Staggered DNN alone was used (upper plot in Fig. 11), the DOA detection would be incorrect, where the absolute error of θ_2 would be $|\theta_2 - \hat{\theta}_2| = 0.93^\circ$. As explained in Sect. 4.2, where it is very likely that either DOA

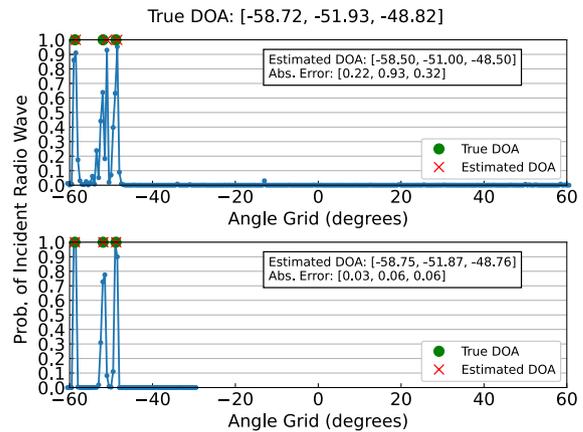


Fig. 11 Example of spectrum grid generation with Staggered DNN (upper figure) and Staggered NRDNN (lower figure) when the SNR is 20 dB and the true DOAs are close within the range of 20° . The green circles and the red X's represent the true DOAs and the estimated DOAs with Neighbors Weighted Average (Sect. 3.3), respectively.

is incorrectly detected when they lie within a range of 20° , and noting that the range of estimated DOAs is less than 20° ($-48.50^\circ - (-58.50^\circ) = 10.00^\circ < 20^\circ$), it can be said that the spectrum grid produced by the appropriate Staggered NRDNN could be more reliable. The selected Staggered NRDNN to produce such spectrum is the one covering the angle bins that fully includes the estimated DOAs, that is, the one that covers the range $[-60^\circ, -30^\circ]$ (first one from the left in Fig. 5(b)). As a result, we obtain the spectrum grid shown in the lower part of Fig. 11. Not only it is a cleaner spectrum, but also it manages to detect all 3 DOAs more precisely, as it can be seen from the considerable drop in the absolute error in the same figure. Therefore, as it will be shown in the next section, this strategy can indeed increase the performance of Staggered DNN at higher SNRs.

5.3 Full System

In Fig. 12, for different test SNRs, the probability of correct DOA estimation and RMSE of the proposed combination of Staggered DNN-PCA and Staggered NRDNN is presented. We compare it again with Staggered DNN and root-MUSIC. The regression-based SNR estimation has been trained and used in the same way as explained in [11]. Here we only consider the case where $L = 10$. The test DOAs were generated as in Table 1.

In terms of probability of correct DOA estimation (Fig. 12(a)), our Staggered NRDNN strategy proposed to cope with close waves shows very good results, especially when the SNR is 20 and 25 dB, where the proposed technique surpasses root-MUSIC performance, while Staggered DNN alone cannot do the same. When the SNR is 30 dB, indeed root-MUSIC still shows better estimation performance; however, our proposed method still manages to perform well. We believe that its use is more attractive than root-MUSIC due to lesser online computational cost, once all necessary DNNs have been offline trained. From the RMSE

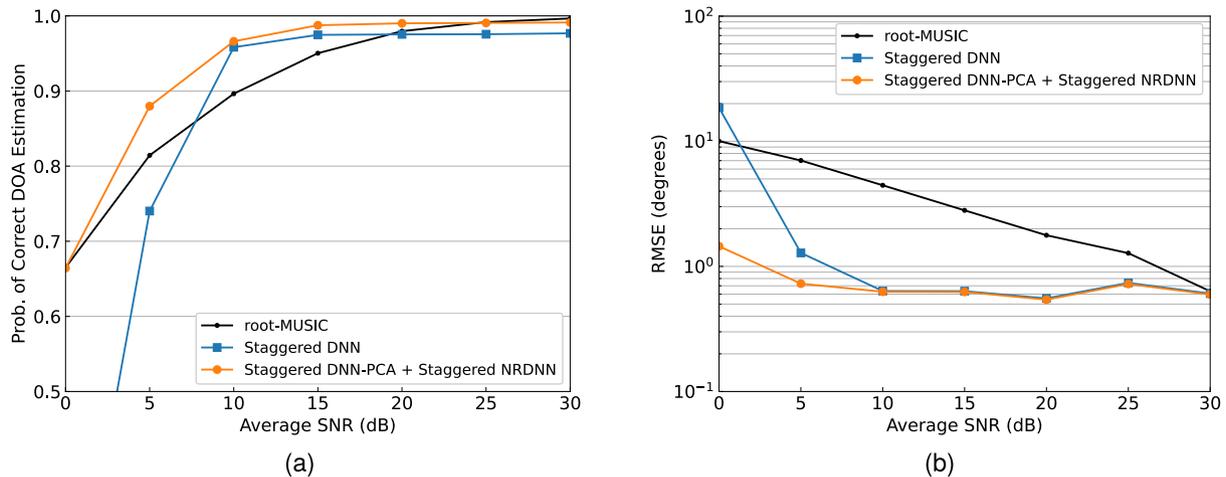


Fig. 12 Performance comparison of root-MUSIC, Staggered DNN and full system (Staggered DNN-PCA + Staggered NRDNN) for varying test SNRs when $L = 10$. The number of principal components chosen for Staggered DNN-PCA was the corresponding optimal value of 13. (a) Probability of correct DOA estimation. (b) RMSE.

in Fig. 12(b), no considerable change is visible when using Staggered NRDNN at 10 dB and over, but this was expected, since the RMSE is an averaging metric and close waves are statistically less common in accordance to our simulation settings.

6. Final Remarks

In this study, we have developed strategies for improving the DOA estimation performance of our previously proposed DNN-based method. Very good results overall surpassing root-MUSIC were achieved in the past when only two radio wave sources were considered. However, as reported in [11], in addition to the fact that accuracy at low SNRs is overwhelmingly poor unless multiple DNNs trained under these conditions are provided, in the event of three radio waves, estimation performance also drops considerably, especially at high SNRs. Consequently, the need to develop schemes that handle these deficiencies was apparent.

Therefore, in this paper we have proposed two separate strategies, each of which tackles these issues at low and high SNRs independently. At low SNRs, we have demonstrated that estimation accuracy is tremendously improved by representing the DNN input vector in a lower dimension (reduction of approximately 85%) by means of PCA, even though this data is generated at a much higher SNR. Additionally, by reducing the size of the input layer, we concurrently manage to reduce the computational cost of DNN. At high SNRs, after noticing that the majority of incorrect estimation cases are due to close waves, we have developed a method where different DNNs specialized in close waves are used instead of the conventional DNN, resulting in a more reliable narrow DOA spectrum grid for subsequent DOA detection. Finally, in order to combine both strategies in a way that such DNN could potentially be deployed in a real scenario, we have used a previously proposed idea [11] of estimating the SNR of the

incoming radio waves, so that the appropriate strategy can be switched depending on this SNR. We have obtained great results with this proposed system for the case of three sources with the promise that it could be used instead of root-MUSIC in a bid to acquire better DOA estimation performance while reducing computational cost.

However, further investigation is still necessary before implementation in real scenarios. Despite the brief discussion of the applicability of the proposed methods for higher number of radio wave sources given in the Appendix, more detailed results are still needed. Moreover, study on a less complex SNR estimation module and performance comparison at a) different number of snapshots, b) coherent radio waves, c) uneven power among the incoming waves is necessary.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 23H01406.

References

- [1] R.O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol.AP-34, no.3, pp.276–280, March 1986. DOI: 10.1109/TAP.1986.1143830.
- [2] B.D. Rao and K.V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Process.*, vol.37, no.12, pp.1939–1949, Dec. 1989. DOI: 10.1109/29.45540.
- [3] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol.3, no.4, pp.563–575, Dec. 2017. DOI: 10.1109/TCCN.2017.2758370.
- [4] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol.67, no.9, pp.8549–8560, Sept. 2018. DOI: 10.1109/TVT.2018.2851783.
- [5] M. Chen, Y. Gong, and X. Mao, "Deep neural network for estimation of direction of arrival with antenna array," *IEEE Access*, vol.8, pp.140688–140698, Aug. 2020. DOI: 10.1109/

ACCESS.2020.3012582.

[6] D. Hu, Y. Zhang, L. He, and J. Wu, "Low-complexity deep-learning-based DOA estimation for hybrid massive MIMO systems with uniform circular arrays," *IEEE Wireless Commun. Lett.*, vol.9, no.1, pp. 83–86, Jan. 2020. DOI: 10.1109/LWC.2019.2942595.

[7] Z.-M. Liu, C. Zhang, and P.S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Antennas Propag.*, vol.66, no.12, pp.7315–7327, Dec. 2018. DOI: 10.1109/TAP.2018.2874430.

[8] Y. Kase, T. Nishimura, T. Ohgane, Y. Ogawa, D. Kitayama, and Y. Kishiyama, "Fundamental trial on DOA estimation with deep learning," *IEICE Trans. Commun.*, vol.E103-B, no.10, pp.1127–1135, Oct. 2020. DOI: 10.1587/transcom.2019EBP3260.

[9] Y. Kase, T. Nishimura, T. Ohgane, Y. Ogawa, T. Sato, and Y. Kishiyama, "Accuracy improvement in DOA estimation with deep learning," *IEICE Trans. Commun.*, vol.E105-B, no.5, pp.588–599, May 2022. DOI: 10.1587/transcom.2021EBT0001.

[10] D.A. Ando, T. Nishimura, T. Sato, T. Ohgane, Y. Ogawa, and J. Hagiwara, "A proposal of an end-to-end DoA estimation system aided by deep learning," *Proc. WPMC 2022*, pp.98–103, Oct. 2022. DOI: 10.1109/WPMCS5625.2022.10014749.

[11] D.A. Ando, Y. Kase, T. Nishimura, T. Sato, T. Ohgane, Y. Ogawa, and J. Hagiwara, "Deep neural networks based end-to-end DOA estimation system," *IEICE Trans. Commun.*, vol.E106-B, no.12, pp.1350–1362, Dec. 2023. DOI: 10.1587/transcom.2023CEP0006

[12] M.E. Tipping and C.M. Bishop, "Mixtures of probabilistic principal component analyzer," *Neural Comput* 1999, vol.11, no.2, pp.443–482, Feb. 1999. DOI: 10.1162/089976699300016728

[13] J.T.C. Ming, N.M. Noor, O.M. Rijal, R.M. Kassim, and A. Yunus, "Lung disease classification using different deep learning architectures and principal component analysis," *Proc. ICBAPS 2018*, pp.187–190, July 2018. DOI: 10.1109/ICBAPS.2018.8527385

[14] Z. Chen and L. Pei, "A PCA-BP fast estimation method for broadband two-dimensional DOA of high subsonic flight targets based on the acoustic vector sensor array," *Proc. CISP-BMEI 2021*, pp.1–6, Oct. 2021. DOI: 10.1109/CISP-BMEI53629.2021.9624358

[15] Y. Liu, H. Chen, and B. Wang, "DOA estimation of underwater acoustic signals based on PCA-kNN algorithm," *Proc. CIBDA 2020*, pp.486–490, April 2020. DOI: 10.1109/CIBDA50819.2020.00115

[16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *JMLR* vol.12, no.85, pp.2825–2830, 2011.

[17] G.W. Stimson, *Introduction to Airborne Radar*, 2nd ed., p.183, SciTech Publishing, Mendham, NJ, US, 1998.

[18] S. Ju, Y. Xing, O. Kanhere, and T.S. Rappaport, "Millimeter wave and sub-terahertz spatial statistical channel model for an indoor office building," *IEEE J. Sel. Areas Commun.*, vol.39, no.6, pp.1561–1575, June 2021. DOI: 10.1109/JSAC.2021.3071844.

[19] D.A. Ando, T. Nishimura, T. Sato, T. Ohgane, Y. Ogawa, and J. Hagiwara, "Performance analysis of DNN-PCA for DOA estimation with three radio wave sources," *Proc. ISCIT 2023*, pp.436–441, Oct. 2023.

[20] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv:1502.03167v3*, March 2015.

[21] D.P. Kingma and J.L. Ba, "Adam: A method for stochastic optimization," *arXiv:1412.6980v9*, Jan. 2017.

Appendix A: Detailed Description of Neighbors Weighted Average

The fully detailed algorithm of "Neighbors Weighted Average" (Sect. 3.3) is presented in Algorithm 1. Its goal is to detect the DOA information from the Staggered DNN

Algorithm 1 "Neighbors Weighted Average"

```

1: Parameters:  $K, \hat{\mathbf{t}}, \epsilon \leftarrow 0.1, \zeta \leftarrow 3$ 
2: Result:  $\hat{\theta}$ 
3: while True do
4:    $\hat{K} \leftarrow \text{countSources}(\epsilon, \hat{\mathbf{t}});$ 
5:   if  $\hat{K} = K$  then
6:      $\hat{\theta} \leftarrow \text{NWA}(\epsilon, \hat{\mathbf{t}});$ 
7:     break;
8:   else if  $\hat{K} > K$  then
9:      $\epsilon \leftarrow \epsilon + 0.1;$ 
10:    if  $\epsilon > 0.4$  then
11:       $\hat{\theta} \leftarrow \text{peakSearch}(K, \hat{\mathbf{t}});$ 
12:      break;
13:    end if
14:  else if  $\hat{K} < K$  then
15:    while True do
16:       $\hat{\theta} \leftarrow \text{NR-NWA}(\epsilon, \zeta, \hat{\mathbf{t}});$ 
17:      if  $\text{length}(\hat{\theta}) = K$  then
18:        break;
19:      else
20:         $\zeta \leftarrow \zeta - 1;$ 
21:        if  $\zeta = 0$  then
22:           $\hat{\theta} \leftarrow \text{peakSearch}(K, \hat{\mathbf{t}});$ 
23:          break;
24:        end if
25:      end if
26:    end while
27:  break;
28: end if
29: end while

```

output, i.e. the angular spectrum grid estimated by the Staggered DNN, for as many output situations as possible, since spurious bins can hinder proper detection.

The algorithm takes as parameters:

- The number of radio wave sources K , which is considered to be known;
- The output of Staggered DNN $\hat{\mathbf{t}}$ (or estimated spectrum grid);
- The probability threshold ϵ with starting value of 0.1;
- The limit on the number of bins ζ for computation of the weighted average of close DOAs, with starting value of 3.

As previously explained in Sect. 3.3, the ϵ and ζ values are not necessarily optimized; yet, we have achieved great results.

Not always a clean spectrum grid is estimated, with K clearly formed hills. For this reason, the first step (line 4) is to count the number of hills \hat{K} present within the estimated spectrum $\hat{\mathbf{t}}$. The probability threshold ϵ is necessary in this moment. The next step relies on \hat{K} :

1. Is $\hat{K} = K$? (lines 5–7)
2. Is $\hat{K} > K$? (lines 8–13)
3. Is $\hat{K} < K$? (lines 14–23)

If case 1, then the weighted average in (9) is simply applied, where all n_{bin} bins within a hill are included in this computation, as shown below:

$$\hat{\theta} = \sum_{i=1}^{n_{\text{bin}}} p_i b_i \Big/ \sum_{i=1}^{n_{\text{bin}}} p_i . \tag{A.1}$$

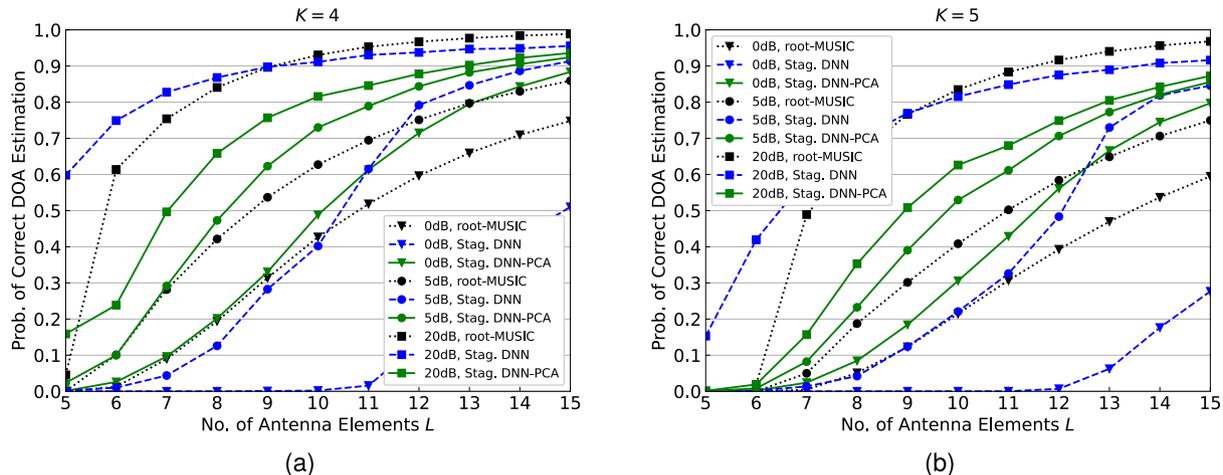


Fig. A.1 Comparison of the performance in terms of probability of correct DOA estimation of 3 DOA techniques while varying the number of antenna elements L : root-MUSIC (dotted black lines), Staggered DNN (dashed blue lines) and Staggered DNN-PCA (solid green lines). These methods were tested at 0, 5 and 20 dB (lower triangle, circle, and square markers, respectively). (a) $K = 4$. (b) $K = 5$.

Case 2 occurs when there are spurious bins higher than the probability threshold ϵ , resulting in spurious hills that should not be included in the DOA detection. A solution for this problem is to increment the value of ϵ step by step (line 9) until such spurious hills are damped and K clear hills are present for DOA detection. Yet, we set 0.4 as a limit to ϵ . If such a limit is reached and still there are no exact K hills, then we apply a traditional peak search algorithm to the Staggered DNN output $\hat{\mathbf{t}}$ (line 11), where the bins corresponding to the K largest peaks are chosen as the DOA estimates $\hat{\theta}$.

Case 3 most probable occurrence is in the event of close radio waves, which results in overlapping hills. In this case, we need to set a bin limit ζ in order to calculate the weighted average (line 16):

$$\hat{\theta} = \sum_{i=1}^{\zeta} p_i b_i \bigg/ \sum_{i=1}^{\zeta} p_i. \quad (\text{A.2})$$

We have verified that $\zeta = 3$ is a good choice. Nevertheless, there are some situations where the number of bins ζ corresponding to one or more DOAs is less than 3. For this reason, we gradually decrement the value of ζ . If this value reaches 0, then again we apply peak search to the Staggered DNN output $\hat{\mathbf{t}}$ (line 22), where the bins corresponding to the K largest peaks are chosen as the DOA estimates $\hat{\theta}$.

Appendix B: Performance of Proposed Methods for Higher Number of Sources

The scope of this study includes primarily the performance analysis of the proposed methods Staggered DNN-PCA in Sect. 4.1 and Staggered NRDNN in Sect. 4.2 for the case of only 3 radio wave sources (i.e. $K = 3$). However, verification of their applicability for higher values of K is also necessary as a fundamental step for future deployment in real-scenario applications. Therefore, in this appendix, we

give a brief analysis of the DOA estimation performance of both proposed methods when K is 4 or 5. In practical scenarios with alternating numbers of sources, we envision a dynamic system which consists of several Staggered DNN-PCAs and NRDNNs, each corresponding to each K , that are concurrently deployed according to this K .

B.1 Staggered DNN-PCA

Figure A.1 shows the extension of the results in Fig. 9(a) when the number of sources K is 4 (Fig. A.1(a)) and 5 (Fig. A.1(b)). Here, the optimal number of principal components for each L was found in the same manner as it was done for Fig. 7. All other parameters were kept unchanged. Comparing Fig. A.1 with Fig. 9(a), the performance of Staggered DNN-PCA appears to be the best at any value of L at 0 and 5 dB; especially when $K = 5$ at 0 dB (see the straight green line with triangle marker in Fig. A.1(b)), Staggered DNN-PCA excels over root-MUSIC at all L , which contrasts with the case $K = 3$ in Fig. 9(a), where the probability of correct DOA estimation of Staggered DNN-PCA only surpasses that of root-MUSIC when $L \geq 10$.

Figure A.2 shows the performance of Staggered DNN-PCA at different test SNRs. This is the extension of the results in Fig. 10 for the case $L = 10$. Although the performance of all DOA estimation methods, including root-MUSIC, is degraded as K increases, Staggered DNN-PCA still shows the best probability of correct DOA estimation at 0 and 5 dB. On the other hand, comparing with the performance of Staggered DNN at SNRs of 10 dB or greater, the performance of Staggered DNN-PCA worsens considerably as K increases. We stated in Sect. 5.1 that information on the signal, which is only lightly corrupted by noise at higher SNRs, is lost by applying PCA. Moreover, as K increases, inaccurate DNN outputs are more often produced due to radio waves with close DOA (more to be discussed in the next

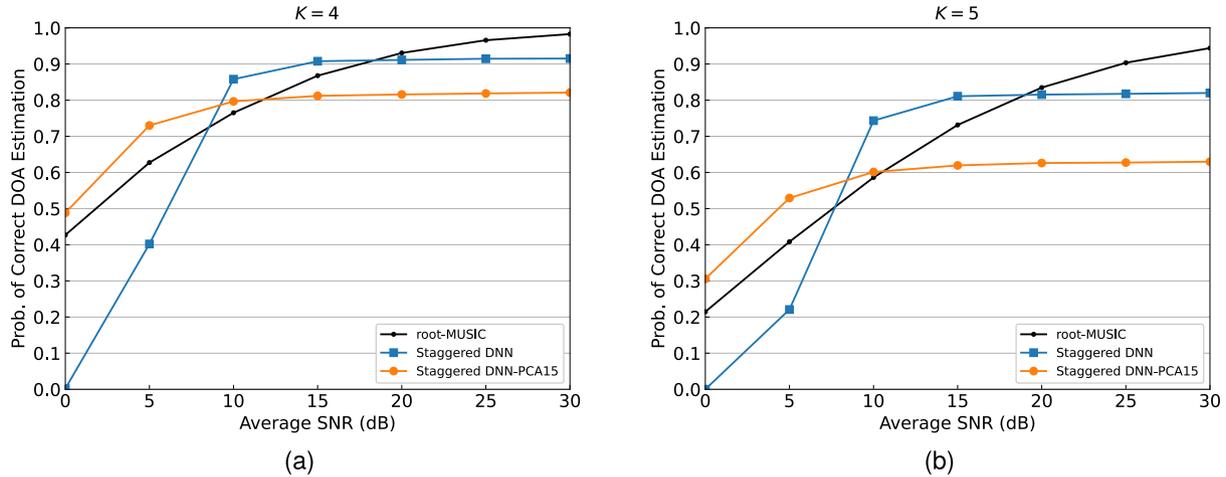


Fig. A-2 Comparison of the probability of correct DOA estimation performance of root-MUSIC, Staggered DNN and Staggered DNN-PCA x for varying test SNRs when the number of antenna elements $L = 10$. The notation PCA x denotes the number of principal components x chosen. (a) $K = 4$. (b) $K = 5$.

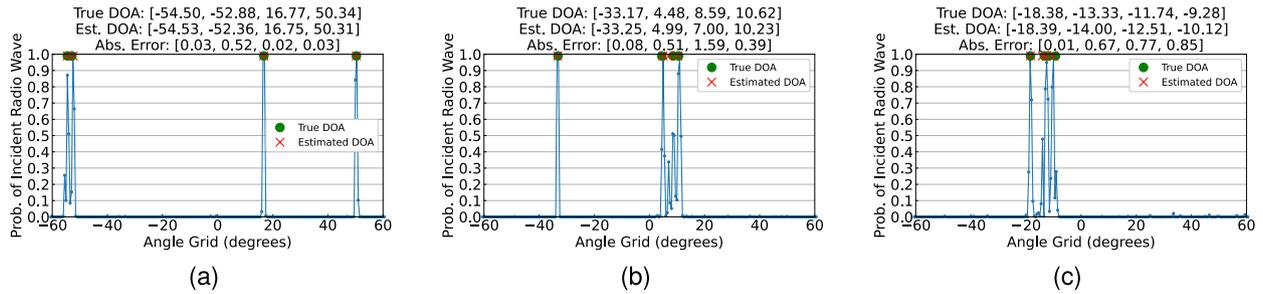


Fig. A-3 Examples of Staggered DNN output when DOA was incorrectly detected at 20 dB for the case $K = 4$. The green circles and the red X's represent the true DOAs and the estimated DOAs, respectively.

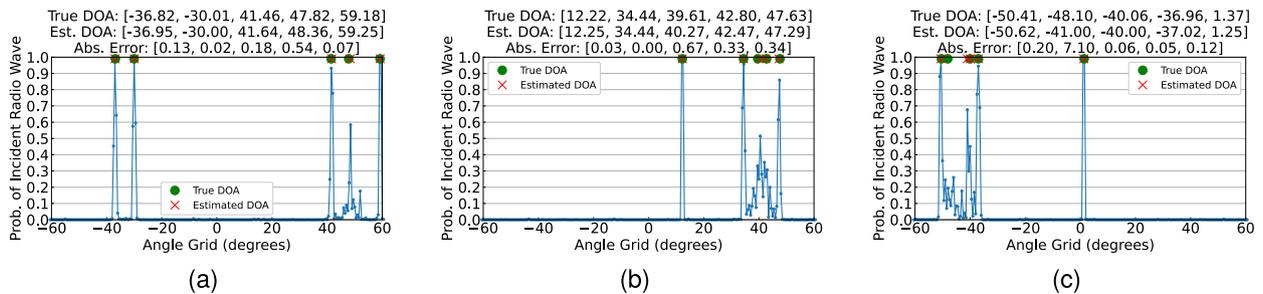


Fig. A-4 Examples of Staggered DNN output when DOA was incorrectly detected at 20 dB for the case $K = 5$. The green circles and the red X's represent the true DOAs and the estimated DOAs, respectively.

section). The combination of both factors are believed to be the reason for significant degradation at higher SNRs as K increases.

In conclusion, applying PCA is still very effective for DOA estimation improvement at lower SNRs when the number of radio wave sources K is 4 and 5.

B.2 Staggered NRDNN

It was explained in Sect. 4.2 and verified in Sect. 5.2 that the

proposed Staggered NRDNN is effective in producing more accurate angular spectra; thus, improving DOA estimation performance. However, this approach was designed based on the observation that $K = 3$ radio waves with DOA range of 20° are the main cause of incorrect DOA estimation at higher SNRs with Staggered DNN. Observing the results presented in Figs. A-3 and A-4, which show examples of the output (i.e. angular spectrum) of Staggered DNN when DOA estimation was incorrect at 20 dB for the case $K = 4$ and $K = 5$, respectively, we concluded that, as K increases,

the possible patterns of angular spectrum corresponding to incorrect DOA also increases. For instance, as opposed to the case $K = 3$, we can observe from Figs. A·3 and A·4 such patterns of angular spectrum where not only all K peaks, but also $K - K'$ peaks are in close range. Here, K' is the number of radio waves apart from the close range of 20° . In fact, we verified that, by applying Staggered NRDNN as described in Sect. 4.2 to the cases $K = 4$ and $K = 5$, no significant improvement in DOA performance was achieved. However this was an expected result, since this method was designed for $K = 3$. We believe that the redesign and retrain of this method based on different K can result in more accurate DOA estimation. Since this investigation is out of the scope of this paper, we leave it as future work.



Daniel Akira Ando received the B.E. degree in communication networks engineering from University of Brasilia, Brazil, in 2018 and the M.E. degree in media networks engineering from Hokkaido University, Japan, in 2021. He is currently pursuing the Ph.D. degree at Hokkaido University, Japan. His research interests are in MIMO signal processing for wireless communications. He received the IEICE RCS Young Researcher Award in 2020.



Toshihiko Nishimura received the B.S. and M.S. degrees in physics and Ph.D. degree in electronics engineering from Hokkaido University, Sapporo, Japan, in 1992, 1994, and 1997, respectively. Since 1998, he has been with Hokkaido University, where he is currently a Professor. His current research interests are in MIMO systems using smart antenna techniques. He received the Young Researchers' Award of IEICE in 2000, the Best Paper Award from IEICE in 2007, and TELECOM System Technology Award from the

Telecommunications Advancement Foundation of Japan in 2008, the best magazine paper award from IEICE Communications Society in 2011, and the Best Tutorial Paper Award from the IEICE Communications Society in 2018. He is a member of the IEEE.



Takanori Sato was born in Hokkaido, Japan, in 1992. He received his Ph.D. degree in the field of media and network technologies from Hokkaido University, Japan, in 2018. He was a Research Fellow of Japan Society for the Promotion of Science (JSPS) from 2017 to 2019. In 2019, he moved to University of Hyogo as an assistant professor. He is currently an associate professor in Hokkaido University. His research interests include the theoretical and numerical studies of optical fibers and photonic circuits using

the coupled mode theory and the finite element method. He is a member of the Japan Society of Applied Physics (JSAP), Institute of Electrical and Electronics Engineers (IEEE), and the Optical Society of America (OSA).



Takeo Ohgane received the B.E., M.E., and Ph.D. degrees in electronics engineering from Hokkaido University, Sapporo, Japan, in 1984, 1986, and 1994, respectively. From 1986 to 1992, he was with Communications Research Laboratory, Ministry of Posts and Telecommunications. From 1992 to 1995, he was on assignment at ATR Optical and Radio Communications Research Laboratory. Since 1995, he has been with Hokkaido University, where he is currently a Professor. During 2005–2006, he was

at Centre for Communications Research, University of Bristol, U.K., as a Visiting Fellow. His research interests are in MIMO signal processing for wireless communications. He received the IEEE AP-S Tokyo Chapter Young Engineer Award in 1993, the Young Researchers' Award of IEICE in 1990, the Best Paper Award from IEICE in 2007, TELECOM System Technology Award from the Telecommunications Advancement Foundation of Japan in 2008, the Best Magazine Paper Award from IEICE Communications Society in 2011, and the Best Tutorial Paper Award from IEICE Communications Society in 2018. He is a member of the IEEE.



Yasutaka Ogawa received the B.E., M.E., and Ph.D. degrees from Hokkaido University, Sapporo, Japan, in 1973, 1975, and 1978, respectively. Since 1979, he has been with Hokkaido University, where he is currently a Professor Emeritus. During 1992–1993, he was with ElectroScience Laboratory, the Ohio State University, as a Visiting Scholar, on leave from Hokkaido University. His professional expertise encompasses super-resolution estimation techniques, applications of adaptive antennas for mobile

communication, multiple-input multiple-output (MIMO) techniques, and measurement techniques. He proposed a basic and important technique for time-domain super-resolution estimation for electromagnetic wave measurement such as antenna gain measurement, scattering/diffraction measurement, and radar imaging. Also, his expertise and commitment to advancing the development of adaptive antennas contributed to the realization of space division multiple accesses (SDMA) in the Personal Handy-phone System (PHS). He received the Yasujiro Niwa Outstanding Paper Award in 1978, the Young Researchers' Award of IEICE in 1982, the Best Paper Award from IEICE in 2007, TELECOM system technology award from the Telecommunications Advancement Foundation of Japan in 2008, the Best Magazine Paper Award from IEICE Communications Society in 2011, the Achievement Award from IEICE in 2014, and the Best Tutorial Paper Award from IEICE Communications Society in 2018. He also received the Hokkaido University Commendation for excellent teaching in 2012. He is a Life Fellow of the IEEE.



Junichiro Hagiwara received the B.E., M.E., and Ph.D. degrees from Hokkaido University, Sapporo, Japan, in 1990, 1992, and 2016, respectively. He joined the Nippon Telegraph and Telephone Corporation in April 1992 and transferred to NTT Mobile Communications Network, Inc. (currently NTT DOCOMO, INC.) in July 1992. Later, he became involved in the research and development of mobile communication systems. His current research interests are in the application of stochastic theory to the

communication domain. He was a visiting professor at Hokkaido University from 2018 to 2023. He is a member of the IEEE.