# DA-MUSIC

# DA-MUSIC: Data-Driven DoA Estimation via Deep Augmented MUSIC Algorithm

Julian P. Merkofer, *Student Member, IEEE*, Guy Revach, *Graduate Student Member, IEEE*, Nir Shlezinger, *Senior Member, IEEE*, Tirza Routtenberg, *Senior Member, IEEE*, and Ruud J. G. van Sloun, *Member, IEEE*

***Abstract*—Direction of arrival (DoA) estimation of multiple signals is pivotal in sensor array signal processing. A popular multi-signal DoA estimation method is the multiple signal classification (MUSIC) algorithm, which enables high-performance super-resolution DoA recovery while being highly applicable in practice. MUSIC is a model-based algorithm, relying on an accurate mathematical description of the relationship between the signals and the measurements and assumptions on the signals themselves (non-coherent, narrowband sources). As such, it is sensitive to model imperfections. In this work, we propose to overcome these limitations of MUSIC by augmenting the algorithm with specifically designed neural architectures. Our proposed deep augmented MUSIC (DA-MUSIC) algorithm is thus a hybrid model-based/data-driven DoA estimator, which leverages data to improve performance and robustness while preserving the interpretable flow of the classic method. DA-MUSIC is shown to learn to overcome limitations of the purely model-based method, such as its inability to successfully localize coherent sources as well as estimate the number of coherent signal sources present. We further demonstrate the superior resolution of the DA-MUSIC algorithm in synthetic narrowband and broadband scenarios as well as with real-world data of DoA estimation from seismic signals.**

***Index Terms*—DoA estimation, MUSIC, model-based deep learning.**

## I. Introduction

SOURCE separation, localization, and tracking are crucial tasks in sensor array processing. In particular, Direction of arrival (DoA) estimation of multiple, broadband, and possibly coherent signal sources plays a key role in a wide range of applications, including radar, communications, image analysis, and speech enhancement [2], [3], [4]. Over the last decades, a multitude of different Direction of arrival (DoA) estimation algorithms have been proposed, and the problem is an active area of research [5], [6].

A leading scheme employed in many DoA applications is the popular multiple signal classification (MUSIC) algorithm [7], which can provide asymptotically unbiased estimates of the number of incident wavefronts present, their approximate frequencies, and their DoAs. multiple signal classification (MUSIC) and other classic approaches, e.g., conventional beamforming [8] and MVDR beamforming [9], are based on knowledge of the underlying statistical model; this dependency induces several key limitations. Among the drawbacks of the model-based (MB) approaches is the inherent limitation of the signal model from which they are derived, which considers narrowband signals. This results, for example, in the inability to consistently estimate the DoA of correlated (coherent) signals, as well as a failure to resolve closely spaced signals with an insufficient number of samples or a low signal-to-noise ratio (SNR) [5], [7].

To extend narrowband models to broadband DoA estimation, various extensions and alternative approaches have been explored [10], because when the impinging signals are broadband, multiple frequency ranges can carry different information regarding the DoA angles at hand. Generally, these extensions from narrowband to broadband can be subdivided into coherent and incoherent processing methods. In coherent processing methods, the covariance matrices of the observations in different frequency bins are coherently combined using certain transformation matrices. Most coherent techniques are based on the coherent signal subspace (CSS) concept [11], which focuses the transformations of the covariances at different frequencies into a surrogate narrowband model, yet utilizes different transformation matrices. Incoherent broadband processing methods combine DoA estimates obtained separately for each frequency bin [12].

The recent success of data-driven (DD) deep learning (DL) across a wide range of disciplines gave rise to neural network (NN)-aided DoA estimators [13]. The works [14], [15], [16], [17] implemented model-agnostic DoA estimation using a dense NN, a convolutional neural network (CNN), a

Manuscript received 11 January 2023; revised 14 August 2023; accepted 7 September 2023. Date of publication 28 September 2023; date of current version 13 February 2024. This work was supported by the Spectralligence, EUREKA IA Call, ITEA4 under Project 20209. The work of Tirza Routtenberg was supported by the PAZY Foundation, Israel. The review of this article was coordinated by Prof. Zhiguo Shi. An earlier version of this work was presented at the 2022 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) as the paper [DOI: 10.1109/ICASSP43922.2022.9746637]. *(Corresponding author: Julian P. Merkofer.)*

Julian P. Merkofer was with the ETH Zürich, 8092 Zürich, Switzerland. He is now with the Electrical Engineering Department, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands (e-mail: j.p.merkofer@tue.nl).

Guy Revach is with the D-ITET, ETH Zürich, 8092 Zürich, Switzerland (e-mail: grevach@ethz.ch).

Nir Shlezinger is with the School of ECE, Ben-Gurion University of the Negev, Beer Sheva 8499000, Israel (e-mail: nirshl@bgu.ac.il).

Tirza Routtenberg is with the School of ECE, Ben-Gurion University of the Negev, Beer Sheva 8499000, Israel, and also with the ECE Department, Princeton University, Princeton, NJ 08540 USA (e-mail: tirzar@bgu.ac.il).

Ruud J. G. van Sloun is with the Electrical Engineering Department, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands (e-mail: r.j.g.v.sloun@tue.nl).

Digital Object Identifier 10.1109/TVT.2023.3320360

sparse-connected CNN, and a U-Net architecture, respectively. Other methods [18], [19] utilize the knowledge of the spatial covariance matrices, where [18] trained a CNN based on a classification task (allowing it to also operate with unknown number of sources) and [19] analyzed purely data-driven (DD) estimators and DD methods in combination with maximum likelihood estimation (MLE). The works [20], [21] also combined DL with MLE by using multiple dense NNs and a ResNet, respectively, to estimate a subset of candidate angles. [22] extends the architecture of [21] to estimate the number of sources with the ResNet. While such black-box NNs enable handling array imperfections due to their model-agnostic nature, they involve highly parameterized models that may be computationally intensive and lack the interpretability of model-based (MB) methods.

Alternatively, NNs were used to robustify the MB MUSIC as a form of hybrid MB/DD system [23], [24], [25]. Specifically, the work [26] and [27] proposed to estimate the discretized MUSIC spectrum from the (spatially smoothed) covariance matrix of the measurements through the utilization of multiple convolutional NNs. While these methods are more robust to model inaccuracies compared with the original MUSIC algorithm, utilizing the MUSIC spectrum as a label for training causes them to experience the same drawbacks as their MB counterpart. Another DD approach proposed in [28] considered systems with subarray sampling and uses NNs to obtain a single estimated covariance matrix from incoherent subarray measurements. This NN-aided estimate is utilized for DoA recovery via the subspace-based MUSIC algorithm. The method addresses the fundamental dependency of MUSIC on the estimated covariance matrix, thereby robustifying the MUSIC algorithm, yet it does not fully exploit the NNs' ability to improve MUSIC, as the NN is trained using the true covariance matrix as a label, without considering its downstream task. Furthermore, the work [29] proposed a hybrid MB/DD approach that includes an eigenvalue decomposition (EVD) of full-row Toeplitz matrices reconstruction (FTMR) matrices [30] and classification deep NNs for estimation of a MUSIC-like spectrum and detection of number of sources present via eigenvectors and eigenvalues

The works [31], [32] applied CNNs for broadband DoA estimation of a single source. The two approaches, however, utilize very different preprocessing methods and are trained as classification and regression tasks, respectively. While the classification approach limits itself to a fixed resolution, the regressor of [32] outputs the DoA angles directly. Therefore, it can achieve an arbitrary resolution, yet it is limited to the spatial covariance matrix as input. Another approach [33] first decomposed the broadband signals into non-overlapping narrowband components and used support vector regression to obtain the DoA.

Accordingly, there is a concrete need for DoA estimation algorithms that can resolve multiple signals that are possibly broadband and coherent in a low complexity and interpretable manner. The aforementioned individual limitations of both existing MB and DD DoA estimation algorithms thus motivate the derivation of a NN-aided DoA estimator capable of leveraging data to enable MUSIC to successfully operate in such scenarios, which is the focus of this work.

We introduced deep augmented multiple signal classification (DA-MUSIC) in [1], a hybrid MB/DD implementation of MUSIC, which exploits the structure of the classic subspace algorithm while augmenting it with NNs to learn to enhance its performance. The proposed architecture overcomes the fundamental limitations of MUSIC, enabling it to accurately detect the locations of coherent sources. Our design builds upon the insight that the sensitivity to model mismatch and inability to handle coherent and broadband sources of MUSIC lie in its estimation of the noise and signal subspaces through an EVD of the empirical covariance matrix. Accordingly, the proposed DA-MUSIC improves this crucial step by obtaining a surrogate, pseudo covariance matrix through a recurrent neural network (RNN) from the measurements, which is learned along with a NN that acts as a peak finder. This work presents crucial extensions to DA-MUSIC allowing it to operate with an unknown and varying number of signal sources by successfully estimating the number of sources for non-coherent and coherent signals. An additional neural augmentation allows the categorization into noise and signal eigenvectors to be obtained in a learned fashion by training a classifier. Numerical results show an estimation accuracy of 98% in comparison to MUSIC with 89% accuracy in the same non-coherent scenario. We further demonstrate DA-MUSIC's capabilities in various SNR domains, in three different broadband scenarios, as well as with real-world seismic signals. DA-MUSIC not only manages to focus frequency components but is also able to concurrently focus an interdependent elevation angle to receive a stable azimuth estimation.

The rest of the article is organized as follows: Section II describes the assumed system model and surveys some of the related work of DoA estimation, as well as the technical details of the MB MUSIC algorithm; Section III introduces the DA-MUSIC algorithm; Section IV presents the results of the simulations; and Section V provides concluding remarks.

Throughout the article, we use boldface lower-case letters for vectors; e.g., $\mathbf{x}$, and boldface uppercase letters for matrices, e.g., for $\mathbf{X}$. The $i$th entry of a vector $\mathbf{x}$ is denoted by $x_i$ and entries separated by commas represent column vectors. We use calligraphic letters to denote the discrete-time Fourier transform, e.g., $\mathcal{X}(\omega)$ is the frequency representation of a signal $x(t)$. We use $(\cdot)^\top$ and $(\cdot)^{\mathrm{H}}$ to denote the transpose and Hermitian operators respectively, and $\|\cdot\|$ and $\|\cdot\|_{\mathrm{F}}$ to denote the $\ell_2$ and Forbenius norms respectively. Further, $\mathcal{U}$ and $\mathcal{CN}$ represent the uniform and the complex normal distributions. The symbol $\mathbb{I}$ denotes the identity matrix.

## II. System Model and Preliminaries

In this section, we detail the system model for which we develop the proposed DA-MUSIC algorithm in Section III. To that aim, we first discuss the signal model in Section II-A. Then, we formulate the DoA estimation problem in Section II-B, and survey some of the related literature in Section II-C. Since our proposed solution builds upon the MUSIC algorithm, we recall

MUSIC in Section II-D and briefly discuss the CSS method for broadband extension in Section II-E.

### A. Signal Model

We distinguish the considered signal model into two cases: the conventional narrowband setting and its more general broadband setup. For more details regarding the two models, we refer the reader to the textbooks [10], [34], and [35].

*1) Narrowband:* The most common setup in the array signal processing literature concerning DoA estimation is the narrowband setting. Here, the time it takes for the waves to propagate the array is assumed to be negligible such that the occurring delays are sufficiently small. Therefore, the signal measurement model for an arbitrary array structure consisting of $M$ sensor elements measuring $D$ impinging narrowband signals takes the following form:

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

In (1), the measurements $\mathbf{x}(t) \in \mathbb{C}^M$ at time instance $t$ depend on the signals $\mathbf{s}(t) \in \mathbb{C}^D$, which originate from the unknown angles $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_D]$, while $\mathbf{v}(t) \in \mathbb{C}^M$ is additive white Gaussian noise. The matrix $\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{C}^{M \times D}$ contains the steering vectors $\{\mathbf{a}(\theta_d)\}$, i.e.,

$$\mathbf{A}(\boldsymbol{\theta}) = \Big[\mathbf{a}(\theta_1) \ldots \mathbf{a}(\theta_D)\Big], \tag{2}$$

where $\{\theta_d\}$ are the source directions. For example, the steering vector $\mathbf{a}(\psi)$ for a uniform linear array (ULA) with an element spacing of $\Delta m = \ell/2$, where $\ell$ is the wavelength of the signals, is defined for direction $\psi$ as

$$\mathbf{a}(\psi) = \Big[1,\ e^{-j\pi \sin\psi},\ \ldots,\ e^{-j\pi(M-1)\sin\psi}\Big]. \tag{3}$$

Consequently, the steering vectors (and, in turn, the matrix $\mathbf{A}(\boldsymbol{\theta})$) specify the underlying array geometry. The collection of the measurements at the array elements over multiple time instances is defined as

$$\mathbf{X} = \Big[\mathbf{x}(1)\ \ldots\ \mathbf{x}(T)\Big], \tag{4}$$

where $T$ is referred to as the number of snapshots.

*2) Broadband:* In practice, many signals are broadband, and the delay caused by propagating the array aperture needs to be incorporated into the signal model. The following notation models the measurements received at array element $m \in \{1, \ldots, M\}$

$$x_m(t) = \sum_{d=1}^{D} s_d(t - \tau_{md}) + v_m(t), \tag{5}$$

where $\tau_{md}$ represents the time delay of the $m$th array element measuring source $d \in \{1, \ldots, D\}$ compared to the measurement at the reference element $m = 1$. Transformed to the frequency domain, the relationship in (5) at frequency $\omega$ becomes

$$\mathcal{X}_m(\omega) = \sum_{d=1}^{D} e^{-j\omega\tau_{md}}\, \mathcal{S}_d(\omega) + \mathcal{V}_m(\omega), \tag{6}$$

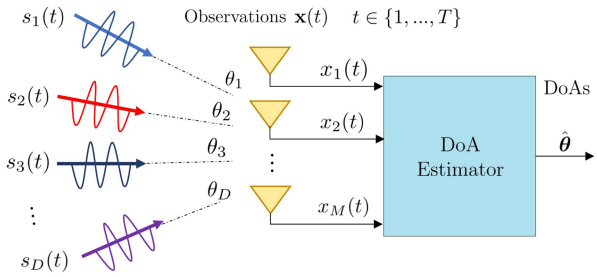


Fig. 1. DoA estimation illustration.

which can, in turn, be written in vector-matrix form analogous to the narrowband system model (1) as follows:

$$\mathcal{X}(\omega) = \mathbf{A}(\omega, \boldsymbol{\theta})\, \mathcal{S}(\omega) + \mathcal{V}(\omega). \tag{7}$$

For a ULA, the broadband steering vectors are defined for frequency $\omega$ and angle of interest $\psi$ as

$$\mathbf{a}(\omega, \psi) = \Big[1,\ e^{-j\omega\frac{\Delta m}{c}\sin\psi},\ \ldots,\ e^{-j\omega(M-1)\frac{\Delta m}{c}\sin\psi}\Big], \tag{8}$$

where $c$ is the propagation velocity.

### B. Problem Formulation

DoA estimation is concerned with localizing signal sources by determining their angles of incidence utilizing the measurements of an array aperture [34]. Formally, this corresponds to estimating the angles $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_D]$ from the measurements $\mathbf{x}(t) = [x_1(t), \ldots, x_M(t)]$ measured at multiple time instances $t \in \{1, \ldots, T\}$. An illustration of such a system is depicted in Fig. 1. The following additional assumptions are imposed upon the observation model as well as the derived synthetic data. The signals are uncorrelated to the noise (though signals may possibly be correlated with each other) and generated in the far-field region. And there is uniform propagation in all directions in an isotropic and non-dispersive medium. We further assume that we have knowledge (though possibly mismatched) of the underlying array geometry, implying that we can compute an approximation of $\mathbf{a}(\psi)$ or $\mathbf{a}(\omega, \psi)$. We consider a data-driven setting where we have access to training data. This data is a set of $U$ pairs, $\{(\mathbf{X}_u, \boldsymbol{\theta}_u)\}_{u=1}^{U}$, each comprising the observations and DoA angles from where the signals originated. In many scenarios; e.g., in wireless communications, a specific training set can be developed before deployment. If the ground-truth DoAs are not obtainable at all, the data-driven DoA estimator must be trained by utilizing synthetic data that closely describes the real signals. Our goal is thus to leverage the available domain knowledge and data to design a system for recovering the DoAs $\boldsymbol{\theta}$ from a corresponding observation matrix $\mathbf{X}$ whose columns are $T$ snapshots of the measured waveforms at the $M$ sensors.

### C. Related Literature

We next provide an overview of the relevant DoA estimation methods for the presented numerical evaluations of Section IV, followed by reviewing relevant extensions. An extensive overview of various MB DoA estimators can be found

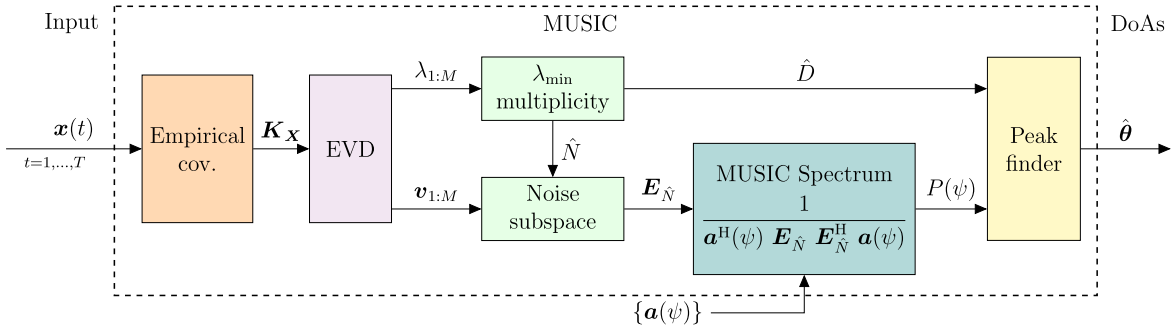


Fig. 2. Block diagram of the MUSIC algorithm.

in [5], [6], and a recent literature review of DD DoA estimation approaches can be found in [36].

*1) MB Narrowband Estimators:* The conventional beamformer (i.e. maximizing the steered response) is a basic approach to DoA estimation [37]. An alternative family of DoA estimators is based on subspace methods, which aim at recovering the DoAs by identifying the noise and signal subspaces. The MUSIC algorithm [7] is a highly popular subspace-based method and has been researched extensively. Being the focus of this article, more information and a detailed description of the algorithm can be found in Section II-D. Extensions of MUSIC include Root-MUSIC [38], a polynomial-rooting version, as well as spatially smoothed MUSIC, which removes the correlation between the incident signals by dividing the receiver array into overlapping subarrays [39]. In practice, however, the number of coherent sources is mostly unknown, and therefore the decorrelation effect of spatial smoothing is not obvious [40]. Another popular subspace-based method for DoA estimation is estimation of signal parameters via rotational invariance techniques (ESPRIT) [41] and its variations [42]. These methods rely heavily on the accuracy of the underlying model assumptions, are generally sensitive to array aperture perturbations, and are inherently derived for narrowband signals.

*2) MB Broadband Estimators:* Broadband DoA estimation algorithms can be categorized into two groups: incoherent methods and coherent methods. Generally, incoherent methods use independent frequency bins (IFBs) to process the DoA information at every frequency separately and then combine the results of these narrowband DoA estimations. There are many different variations in the implementation of this approach [43], [44], [45], which typically vary in the computation of the covariance matrices. Unfortunately, the computational complexity increases with each frequency bin.

Conversely, coherent methods combine the covariance matrices estimated for IFBs in order to apply narrowband techniques directly over a single, so-called focused covariance matrix. An overview of coherent techniques can be found in [10], [46]. A leading approach for coherent broadband DoA estimation is based on the CSS method [11], with example estimators given in [47], [48], [49]. Many variations of the CSS method are concerned with the crucial aspect of the focusing strategy, proposing different techniques to combine the covariances at each IFB to obtain a faithful narrowband formulation. See, for example, [50], [51] and the more recent summary [52]. These methods, however, experience significant bias with larger angular sectors of interest and can impose additional model assumptions such as noise statistics.

*3) DD Estimators:* Inspired by the dramatic success of DL in computer vision and natural language processing, recent years have witnessed a growing interest in the application of NNs for DD DoA estimation. A recent review of DL DoA estimation approaches can be found in [36] and essential methods were covered in Section I. DL architectures presented in the numerical evaluations or otherwise closely related to this work are summarized and discussed below.

The work [18] implemented a CNN architecture for DD DoA estimation based on a multi-label classification task. Particularly, the method takes real, imaginary, and phase information of the sample covariance as a three-channel input and predicts a probability grid of directions. Given the binary cross-entropy loss for training, the method is capable of operating with a potentially unknown and varying number of sources. However, with a growing number of sources, the number of classes grows exponentially.

Similarly to the above, [26] proposed CNNs for DoA estimation by taking the same channels of the sample covariance matrix as input. However, the CNNs are trained as a regression task estimating a discretized segment of the MUSIC spectrum. Thereby, they not only inherit the drawbacks and imitations of MUSIC but are, besides increased robustness, dependant on the underlying model assumptions.

### D. MUSIC Algorithm

As our proposed DA-MUSIC algorithm originates from the MUSIC algorithm, we next present this method in detail. MUSIC, originally proposed by Schmidt in [7], considers the narrowband signal model with incoherent sources (1), where the signals in $\mathbf{s}(t)$ are mutually independent. Fig. 2 visualizes a simple outline of the MUSIC structure as a block diagram. The approach takes the empirical covariance matrix of the received measurements $\mathbf{X}$, then conducts an EVD, followed by categorizing the eigenvectors into signal and noise subspaces. The orthogonality between the two subspaces allows the formulation of a spatial spectrum, which contains peaks at DoA angles.

*1) Formal Derivation:* With the assumption of the signal and the noise being uncorrelated, the covariance matrix of $\mathbf{x}(t)$ is

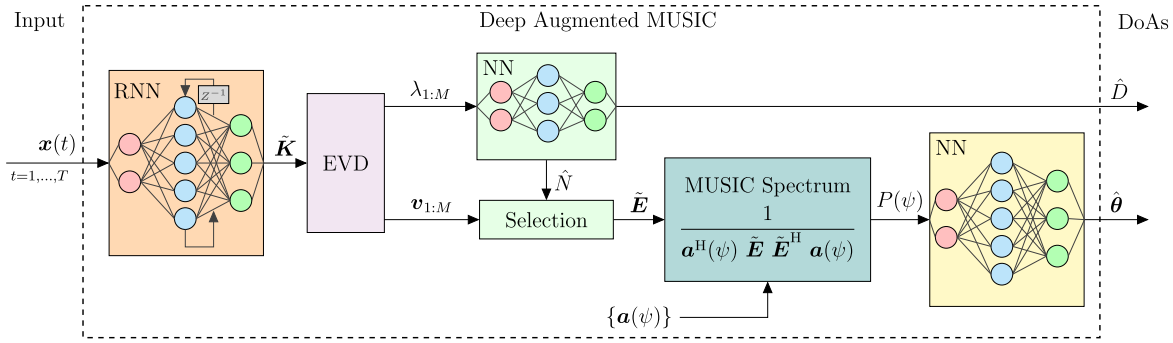


Fig. 3. Block diagram of the DA-MUSIC algorithm.

given by

$$\mathbf{K_X} = \mathbf{A}(\boldsymbol{\theta})\mathbf{K_S}\mathbf{A}^{\mathrm{H}}(\boldsymbol{\theta}) + \lambda\mathbf{K}_{\mathbf{X}}^{0}, \tag{9}$$

with $\mathbf{K_S}$ being the covariance of the incident signals $\mathbf{s}(t)$. The matrix $\mathbf{A}(\boldsymbol{\theta})\mathbf{K_S}\mathbf{A}^{\mathrm{H}}(\boldsymbol{\theta})$ is singular and has a rank of less than $M$ when the number of array elements $M$ is strictly larger than the number of signals $D$. Therefore, $\lambda$ is an eigenvalue of $\mathbf{K_X}$ (in the metric of $\mathbf{K}_{\mathbf{X}}^{0}$, which takes the form $\lambda\mathbf{K}_{\mathbf{X}}^{0} = \sigma^2\mathbb{I}$ for additive white Gaussian noise (AWGN) with variance $\sigma^2$). Further, $\mathbf{A}(\boldsymbol{\theta})\mathbf{K_S}\mathbf{A}^{\mathrm{H}}(\boldsymbol{\theta})$ has to be non-negative definite, because $\mathbf{A}(\boldsymbol{\theta})$ has full rank and $\mathbf{K_S}$ is positive definite, and consequently, $\lambda$ in (9) is the minimal eigenvalue of $\mathbf{K_X}$, denoted $\lambda_{\min}$. The multiplicity of $\lambda_{\min}$ corresponds to the number of incident wavefronts, and equals $N = M - D$.

MUSIC builds upon this representation of the covariance of the signals. The algorithm takes as input $T$ snapshots of the waveforms at $M$ array elements, represented as $\mathbf{X}$ in (4), and uses them to obtain an empirical estimate of $\mathbf{K_X}$ via $\hat{\mathbf{K}}_{\mathbf{X}} = \frac{1}{T}\mathbf{X}\mathbf{X}^{\mathrm{H}}$. Then, the number of incident signals $D$ is estimated via

$$\hat{D} = M - \hat{N}, \tag{10}$$

where $\hat{N}$ is the estimated multiplicity of the minimal eigenvalue of $\hat{\mathbf{K}}_{\mathbf{X}}$. The eigenvectors corresponding to the $\hat{N}$ smallest eigenvalues form the noise subspace $\mathbf{E}_{\hat{N}}$, which is orthogonal to the $D$ dimensional signal subspace spanned by the incident signal mode vectors. Consequently, MUSIC estimates the DoAs by computing the spatial spectrum

$$P(\psi) = \frac{1}{\mathbf{a}^{\mathrm{H}}(\psi)\mathbf{E}_{\hat{N}}\mathbf{E}_{\hat{N}}^{\mathrm{H}}\mathbf{a}(\psi)}, \tag{11}$$

and the $\hat{D}$ dominant peaks of $P(\psi)$ are set as the estimated DoA angles $\hat{\boldsymbol{\theta}}$.

MUSIC is a popular and highly applicable subspace-based method that is reasonably efficient and statistically consistent [5], [53]. When the signal model is adequately accurate, it can achieve super-resolution and deliver a highly accurate estimate of the number of signal sources present. Nevertheless, the algorithm is sensitive towards the accuracy of the empirical estimate of $\mathbf{K_X}$, and cannot reliably estimate the DoA angles as well as the number of sources of coherent signals. The reason for this is that highly correlated signals cause zero entries within the covariance matrix and can, therefore, become indistinguishable from noise [54]. Furthermore, the MUSIC algorithm is inherently a narrowband approach due to the assumptions imposed on the system model. Nonetheless, it can be extended to broadband, i.e., signal models as in (7), and so we next describe the coherent method to achieve this, which is also adopted in our proposed DA-MUSIC.

### E. Coherent Broadband

The main concept behind coherent broadband DoA estimation is to transform the different frequency covariance matrices into a single covariance matrix at a focusing frequency. Accordingly, coherent methods find appropriate transformations for each frequency, transform the covariances, and obtain a focused covariance matrix by some form of averaging. In particular, the CSS method [11] aims at combining the spatial signal subspaces to align the signal subspaces associated with the DoA along all frequency bins. To formulate this, let $\mathbf{K}(\omega)$ be the covariance of the frequency domain observations (7), and divide the spectrum into $B$ IFBs with central frequencies $\{\omega_b\}_{b=1}^{B}$. The focused covariance matrix, which is used as the input covariance for narrowband DoA recovery, is estimated as

$$\mathbf{K} = \sum_{b=1}^{B} \alpha_b\, \mathbf{T}_b\, \mathbf{K}(\omega_b)\, \mathbf{T}_b^{\mathrm{H}}, \tag{12}$$

where $\mathbf{T}_b$ is the focusing matrix and frequency bins are prioritized by the weighting $\alpha_b$. The focusing matrices can be determined by attempting to focus the spectral components at some frequency $\omega_r$ and some focusing angles $\psi$ by solving

$$\mathbf{T}_b = \arg\min_{\mathbf{T}} \|\mathbf{A}(\omega_r, \psi) - \mathbf{T}\,\mathbf{A}(\omega_b, \psi)\|_{\mathrm{F}}. \tag{13}$$

Coherent techniques have been shown to achieve a better estimation accuracy and a smaller computational complexity as well as a lower resolution threshold than non-coherent methods [52]. However, the CSS methods require initial values for the focusing matrix $\mathbf{T}$, the reference frequency $\omega_r$, and the relevant focusing angles $\psi$ to find the focusing matrices with (13), and are typically sensitive towards these initial values. Additionally, it is not guaranteed that the alignment of the signal and noise subspaces exists to form a viable general covariance matrix without disarranging the noise subspace [51].

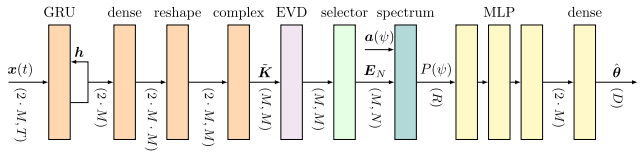


Fig. 4. Detailed network structure of the DA-MUSIC algorithm.

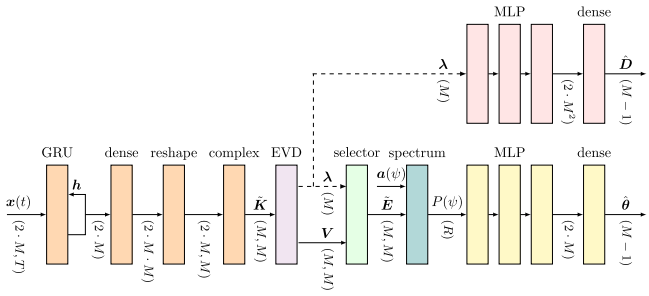


Fig. 5. Detailed network structure of the DA-MUSIC algorithm with a separately trained (internal) classifier.

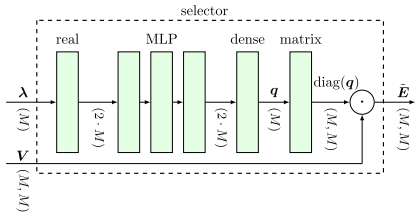


Fig. 6. Detailed outline of the DA-MUSIC subspace selection augmentation.

## III. Deep Augmented MUSIC

DA-MUSIC is a hybrid MB/DD DoA estimation algorithm derived from the classic MUSIC algorithm by replacing crucial and model mismatch sensitive elements of the model-based structure with specific NNs. Fig. 3 outlines the resulting structure of the DA-MUSIC architecture and highlights the remaining similarities to the original MUSIC algorithm, depicted in Fig. 2.

Principally, DA-MUSIC builds upon the understanding that the core challenges associated with the classic MUSIC algorithm can be tackled by providing a surrogate covariance matrix. In particular, as the CSS method provides a surrogate covariance $\mathbf{K}$ that transforms a broadband signal model into a narrowband one, a similar approach can be employed to handle coherent sources and array mismatches. Therefore, to improve the categorization of the noise and signal subspaces, the correlation of the received measurements is learned from temporal data by employing a dedicated RNN, which is augmented into the overall flow of MUSIC. We next elaborate on the architecture in Section III-A, after which we present the training method in Section III-B and provide a discussion in Section III-C.

### A. Architecture

The proposed DA-MUSIC algorithm preserves the structure of the MB MUSIC while replacing certain critical aspects with NNs. Our neural augmentations aim to improve the crucial steps of estimating the noise and signal subspaces from the empirical covariance and the translation of the spatial spectrum into DoAs via peak finding. By doing so, DA-MUSIC is not constrained by the additional model assumptions imposed in the derivation of MUSIC, and can, as we will show, e.g. learn to successfully localize coherent signals. To present the architecture of DA-MUSIC, we commence with the simple case where the number of sources $D$ is known, and then show how its estimation is incorporated. Details of the NNs used in our experimental study are reported in Section IV.

*1) Known Number of Sources:* Fig. 4 depicts a detailed outline of the individual elements of the DA-MUSIC architecture. The respective output dimensions of the corresponding components are given in the bracket notation. First, the input signal $\mathbf{x}(t)$ is transformed into the pseudo covariance matrix $\tilde{\mathbf{K}}$ using a RNN implemented through a gated recurrent unit (GRU). The final state of the GRU is passed to a dense layer enabling a reshaping to the desired dimension of the pseudo covariance matrix $\tilde{\mathbf{K}}$ as well as the subsequent transformation of the complex space. Then, through the continued use of the EVD, the algorithm categorizes the subspaces using the eigenvectors. Inserting the steering vectors $\mathbf{a}(\psi)$ allows to compute an estimate of the spatial spectrum in (11), denoted $P(\psi)$, identically to MB MUSIC, by using the noise eigenvectors selected from $\tilde{\mathbf{K}}$.

Next, DA-MUSIC attains the DoAs from the spatial spectrum $P(\psi)$ using an additional NN, comprised of a multi-layer perceptron (MLP) of three fully connected dense layers followed by a single dense layer with linear activation. The input to the MLP are $R$ samples of $P(\psi)$ taken uniformly in $[0, 2\pi)$. The output of the MLP is the set of estimated DoA angles $\hat{\boldsymbol{\theta}}$. Through the linear output activation the network obtains a larger range of freedom while the periodicity of the angles can be used to map the output to the desired angle range of interest. The benefits of using a NN-based peak-finder compared to a model-based one are two-fold. First, learning the translation of the pseudo-spectrum into DoAs from data enables achieving improved resolution compared to conventional peak finding, since $\hat{\theta}_d \in [0, 2\pi)$ instead of being dependent on the number of angles $\psi$ used to evaluate the spectrum. Furthermore, peak finding is generally non-differentiable; thus, replacing it with a NN facilitates training DA-MUSIC end-to-end. The resulting architecture enables the application of gradient-based optimization, by propagating through the NNs as well as the EVD operation, as done in [55]. Doing so allows us to jointly tune the noise subspace recovery along with the translation of the MUSIC spectrum into DoAs by comparing its estimated DoAs with the true DoAs, as we detail in Section III-B.

*2) Varying Number of Sources:* Delivering unbiased estimates of the number of signal sources as well as the ability to successfully localize these sources makes MUSIC highly applicable. The above-discussed DA-MUSIC architecture can be extended to operate with a potentially unknown and varying number of signal sources, despite the determinant nature of the NNs. This is achieved by addressing three key aspects of the algorithm: the estimation of the number of sources; the selection

process of the noise subspace from the eigenvectors; and the adaptation of the output strategy to overcome the varying number of DoA angles.

*Estimating number of sources:* Our design allows DA-MUSIC to learn abstract mappings as pseudo covariance features $\tilde{\mathbf{K}}$, which is geared towards end-to-end training and is not restricted to a natural ordering of the model-based covariances. Consequently, instead of estimating the number of sources by inspecting the magnitude of its eigenvalues, we opt for a data-driven approach. We augment the process of estimating the number of sources with a classifier, implemented through a MLP, as a classification task. Since subspace methods with $M$ inputs can resolve at most $M-1$ signals, the classifier has $M-1$ classes. Fig. 5 depicts the DA-MUSIC architecture with an added classifier taking the eigenvalues as an input and outputting a probability vector for the estimated number of sources with a softmax activation.

*Computing noise subspace:* The noise subspace selector needs to know the number of sources $D \in \{1,\dots,M-1\}$ to classify the eigenvectors into signal and noise subspaces; i.e., by choosing the eigenvectors corresponding to the $N = M-D$ smallest eigenvalues. When the number of sources is not known, we propose to weight each eigenvector by an estimate of it belonging to the noise subspace. We do this by introducing an additional neural augmentation, depicted in Fig. 6, which uses a MLP to cluster the eigenvalues in a learned fashion.

In particular, the MLP maps the estimated $M$ eigenvalues into a vector $\mathbf{q} = [q_1,\dots,q_M]$ whose entries hold the individual probabilities of choosing the corresponding eigenvector as a noise eigenvector, implemented using a sigmoid activation. The selection is performed by computing $\tilde{\mathbf{E}} = \mathbf{V}\mathrm{diag}(\mathbf{q}) = [\mathrm{q_1}\mathbf{v}_1 \dots \mathrm{q_M}\mathbf{v}_\mathrm{M}]$, allowing to learn a suitable noise subspace. Note that this setting specializes in the conventional approach of assigning based on the multiplicity of the minimal eigenvalues; in the conventional approach, the entries of $\mathbf{q}$ are either ones or zeros and $\tilde{\mathbf{E}}$ coincides with the corresponding subspace. Consequently, the proposed approach provides additional flexibility in selecting the noise subspace and facilitates coping with settings where distinguishing between the eigenvectors is challenging due, for example, to low SNRs.

*Outputting varying number of DoAs:* The most common strategy to overcome the dynamic output dimensions of NNs is to scale the output dimension to the maximum occurring value. In our case, since MUSIC cannot resolve more than $M-1$ sources, the dimension of the last dense layer of DA-MUSIC is set to $M-1$. The only additional modification required with this strategy compared with known $D$ is the slight alteration to the loss function discussed in Section III-B below. An advantage of this strategy is that the approach allows to extract up to $M-1$ DoA, where $\theta_1$ is most likely a true DoA angle, $\theta_2$ has a slightly lower probability to be a DoA angle, etc. The final estimation is thus carried out by first taking $\hat{D}$ from the module that estimates the number of sources, and then using the first $\hat{D}$ outputs as the recovered DoAs.

### B. Training Procedure

DA-MUSIC is trained end-to-end in a supervised setting as a multiple regression problem. As detailed in Section II-B, the training set is comprised of $U$ tuples of sequences of measurements and their corresponding DoAs; i.e., the $u^{\text{th}}$ tuples includes the $T_u$ measurements $\mathbf{X}_u$ and their corresponding $D_u$ DoA angles $\boldsymbol{\theta}_u$. We first describe how this data is used for training when $D$ is known and fixed; i.e., $D_u \equiv D$. This acts as a preliminary step for discussing how the training procedure is carried out in the general case where the model does not have knowledge of $D$.

*1) Known Number of Sources:* Given a sequence of measurements $\mathbf{X}$ as input, the model predicts the estimated DoA angles $\hat{\boldsymbol{\theta}}$ that are compared to the true DoA angles $\boldsymbol{\theta}$. Gradient-based optimization is possible because every element of the architecture is differentiable, allowing backpropagation through the complete structure. Derived from the root mean squared periodic error (RMSPE) [56], [57] the following loss function additionally compares all permutations of the predicted angles with the true angles to capture all possible assignments of the estimated DoA to the true DoA. Thereby the minimal permutation RMSPE includes the permutation invariance of the DoAs and is obtained as:

$$\mathrm{RMSPE}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}) = \min_{\mathbf{P}\in\mathcal{P}_D} \left( \frac{1}{D} \left\| \mathrm{mod}_\beta(\boldsymbol{\theta} - \mathbf{P}\hat{\boldsymbol{\theta}}) \right\|^2 \right)^{\frac{1}{2}}, \quad (14)$$

where $\mathcal{P}_D$ is the set of all $D \times D$ permutations and $\mathrm{mod}_\beta$ denotes the element-wise modulus operation regarding the angle range of interest, e.g. $\beta = \pi$ for $\psi \in [-\pi/2, \pi/2)$ or $\beta = 2\pi$ for $\psi \in [0, 2\pi)$.

*2) Unknown Number of Sources:* As discussed in the previous subsection, DA-MUSIC is designed to resolve a varying and unknown number of sources by introducing an additional NN classifier for the number of sources. To formulate the training procedure of the overall system, we use $\mathbf{w}_c$ to denote the trainable parameters of the classifier, while $\mathbf{w}_d$ represents the parameters of the remaining trainable modules of DA-MUSIC (covariance estimator, peak finder, and subspace selector). The loss used to train $\mathbf{w}_d$ is the RMSPE loss of (14), which is altered during training to account for varying sources by computing,

$$\mathcal{L}_{\mathrm{RMSPE}}\left(\boldsymbol{\theta}_u, \hat{\boldsymbol{\theta}}(\mathbf{X}_u)\right) = \mathrm{RMSPE}\left(\theta_{1:D_u}, \hat{\theta}_{1:D_u}(\mathbf{X}_u)\right). \quad (15)$$

In (15), $\hat{\boldsymbol{\theta}}(\mathbf{X}_u)$ denotes the $M-1$ outputs of DA-MUSIC applied to $\mathbf{X}_u$, while $\theta_{1:D_u} = (\theta_1,\dots,\theta_{D_u})$. This means that only the first $D_u$ angles of the estimated DoA $\hat{\boldsymbol{\theta}}$ are compared with the true DoA $\boldsymbol{\theta}$ while the remaining angles of $\hat{\boldsymbol{\theta}}$ are completely ignored.

The separate classifier is trained using the categorical cross-entropy of the classes as a loss function ensuring optimal training. In particular, letting $\boldsymbol{\lambda}_u$ be the input to the MLP classifier when applying DA-MUSIC to $\mathbf{X}_u$ and letting $\hat{\boldsymbol{D}}(\lambda_u)$ be the softmax output of the MLP applied to $\boldsymbol{\lambda}_u$ (with $\hat{D}_i(\lambda_u)$ being its $i$th entry), the loss used for training $\mathbf{w}_c$ is given by

$$\mathcal{L}_{\mathrm{CE}}\left(D_u, \hat{\boldsymbol{D}}(\boldsymbol{\lambda}_u)\right) = -\log \hat{D}_{D_u}(\lambda_u). \quad (16)$$

**Algorithm 1:** Training DA-MUSIC.

**Data:** Data set $\{(\mathbf{X}_u, \boldsymbol{\theta}_u)\}_{u=1}^{U}$, learning rate $\mu = 0.001$, decay rates $b_1 = 0.9, b_2 = 0.999$, $\epsilon = 10^{-8}$

Initialize weights $\mathbf{w}_d, \mathbf{w}_c$;
Initialize moment vectors $\boldsymbol{\nu}_d, \boldsymbol{\upsilon}_d, \boldsymbol{\nu}_c, \boldsymbol{\upsilon}_c$;
**for** epoch $= 1, 2, \ldots$ **do**
  **for** each batch **do**
    Apply DA-MUSIC to $\{\mathbf{X}_u\}$ for $u \in$ batch;
    Compute gradients $\mathbf{g}_d$ via
    $\mathbf{g}_d \leftarrow \nabla_{\mathbf{w}_d} \sum_{u \in \text{batch}} \mathcal{L}_{\text{RMSPE}}(\boldsymbol{\theta}_u, \hat{\boldsymbol{\theta}}(\mathbf{X}_u))$;
    Compute gradients $\mathbf{g}_c$ via
    $\mathbf{g}_c \leftarrow \nabla_{\mathbf{w}_c} \sum_{u \in \text{batch}} \mathcal{L}_{\text{CE}}(D_u, \hat{\boldsymbol{D}}(\boldsymbol{\lambda}_u))$;
    Update biased first moment via
    $\boldsymbol{\nu}_d \leftarrow b_1 \boldsymbol{\nu}_d + (1 - b_1)\mathbf{g}_d$;
    $\boldsymbol{\nu}_c \leftarrow b_1 \boldsymbol{\nu}_c + (1 - b_1)\mathbf{g}_c$;
    Update biased second raw moment via
    $\boldsymbol{\upsilon}_d \leftarrow b_2 \boldsymbol{\nu}_d + (1 - b_2)\mathbf{g}_d^2$;
    $\boldsymbol{\upsilon}_c \leftarrow b_2 \boldsymbol{\nu}_c + (1 - b_2)\mathbf{g}_c^2$;
    Compute bias corrected moments $\hat{\boldsymbol{\nu}}_d, \hat{\boldsymbol{\nu}}_c, \hat{\boldsymbol{\upsilon}}_d, \hat{\boldsymbol{\upsilon}}_c$;
    Update $\mathbf{w}_d$ via $\mathbf{w}_d \leftarrow \mathbf{w}_d - \mu\, \hat{\boldsymbol{\nu}}_d / (\sqrt{\hat{\boldsymbol{\upsilon}}_d} + \epsilon)$;
    Update $\mathbf{w}_c$ via $\mathbf{w}_c \leftarrow \mathbf{w}_c - \mu\, \hat{\boldsymbol{\nu}}_c / (\sqrt{\hat{\boldsymbol{\upsilon}}_c} + \epsilon)$;
  **end**
**end**

It is noted that since (16) is used for training $\mathbf{w}_c$, then one should block the gradient during backpropagation from passing from the classifier to the EVD, as indicated by the dashed connections in Fig. 5. This ensures that the MLP learns from the eigenvalues themselves and not by influencing and disrupting the GRU. Further, the regressor is completely independent of the classifier, which allows DA-MUSIC to operate with different classifiers if needed or with any other desired scheme delivering an estimate of $D$. The resulting training procedure (employing mini-batch gradient descent with ADAM [58]) is summarized as Algorithm 1.

### C. Discussion

The design of the architecture of DA-MUSIC is derived from the model-based MUSIC structure. This allows for exploiting the successful aspects of the algorithm while improving certain critical elements and alleviating important drawbacks. Replacing the empirical covariance estimation with a RNN is the key neural augmentation of DA-MUSIC, enabling the system to learn the pseudo covariance from the measurements themselves such that the resulting surrogate model facilitates subspace-based DoA recovery. Thereby, the performance of DA-MUSIC, for example, is not affected by coherent signals and other related issues discussed in Section II-D. Furthermore, learning end-to-end allows DA-MUSIC to operate with broadband signals, as it effectively learns to produce a focused pseudo covariance, similarly to the CSS methods discussed in Section II-C. It is noted that our augmentation approach depends on the array geometry, as the steering vectors $\mathbf{a}(\cdot)$ are used for computing the MUSIC spectrum. Nonetheless, as we numerically demonstrate in Section IV, DA-MUSIC learns to overcome mismatches in the array geometry from the data without any alterations or renewed training.

TABLE I
DEFAULT SIMULATION PARAMETERS

| Parameter Description | Notation | Default Value |
|---|---|---|
| Number of training samples | $U$ | $9 \times 10^4$ |
| Number of testing samples | | $1 \times 10^4$ |
| Array geometry | | ULA |
| Number of array elements | $M$ | 8 |
| Element spacing | $\Delta m$ | $\ell_{\min}/2$ |
| SNR | | 10 dB |
| Snapshots | $T$ | 200 |
| Grid points of continuum | $R$ | 360 |
| Min. frequency | $f_{\min}$ | 0 [Hz] |
| Max. frequency | $f_{\max}$ | 999 [Hz] |
| Sampling frequency | $f_s$ | $2(f_{\max} + 1)$ |
| Time length | $T_{samp}$ | 1 s |
| Fast Fourier transform (FFT) points | $N_f$ | $T_{samp} \cdot f_s$ |

Specifically, the RNN utilized by DA-MUSIC is able to learn an appropriate focusing matrix while concurrently correlating the measurements comparably to (12). The usage of a RNN applied to the time-domain signal and only passing the last state to the next layer allows DA-MUSIC to operate with different signal durations and possibly cope in real-time with dynamic variations in the DoAs, though investigation of the latter is left for future work. Our experiments indicate that a deeper RNN aids the correlation aspect while a wider RNN allows a more complicated mapping with this correlation.

Another important component of the architecture of DA-MUSIC is its incorporation of the EVD as a means for division into subspaces. Though computationally expensive, the internal EVD allows DA-MUSIC to not only classify signal and noise subspaces with the eigenvalues but also significantly simplifies the estimation of the number of signal sources present. To enable training end-to-end from the errors in (14), a NN-based peak finder is used in the form of a MLP. Model-based peak-finding is generally non-differentiable; therefore, replacing it with a NN enables gradient-based optimization through the entire DA-MUSIC structure. Furthermore, improved resolution can be achieved by extracting the DoA from the spatial spectrum in a learned manner (i.e., $\hat{\theta}_d \in [0, 2\pi)$ and it is not dependent on the number of angles $\psi$ used to evaluate the spectrum $P(\psi)$). Moreover, while DA-MUSIC is formulated for ULAs, it can be applied with alternative configurations for which one can compute the steering matrix, as we also exemplify in our numerical study in Section IV-C.

## IV. Numerical Evaluations

In this section, we present our numerical evaluations of the proposed DA-MUSIC algorithm[1]. Our experimental study is comprised of evaluations in a narrowband synthetic setting (Section IV-A); a broad synthetic setup (Section IV-B); and experiments with real-world data corresponding to azimuth estimation in seismic arrays (Section IV-C).

[1]The source code used in our experiments can be found at https://github.com/DA-MUSIC/TVT23.

TABLE II
RMSPE OF DIFFERENT DOA ESTIMATION ALGORITHMS WITH CONSTANT AND KNOWN $D$ FOR $T = 200$ SNAPSHOTS

| RMSPE [rad] | DA-MUSIC | CNN | Deep-MUSIC | Classic MUSIC | Random |
|---|---|---|---|---|---|
| non-coherent | | | | | |
| $D = 2$ | **0.0117** | 0.0237 | 0.0329 | 0.0336 | 0.6809 |
| $D = 3$ | **0.0315** | 0.0464 | 0.1600 | 0.0841 | 0.6034 |
| $D = 4$ | **0.0563** | 0.0841 | 0.2656 | 0.1459 | 0.5421 |
| $D = 5$ | **0.0751** | 0.1319 | 0.2701 | 0.2008 | 0.4963 |
| coherent | | | | | |
| $D = 2$ | **0.0140** | 0.0274 | 0.5781 | 0.2350 | 0.6854 |
| $D = 3$ | **0.0407** | 0.0528 | 0.4023 | 0.4819 | 0.6060 |
| $D = 4$ | **0.0519** | 0.0859 | 0.2915 | 0.4522 | 0.5407 |
| $D = 5$ | **0.0658** | 0.1254 | 0.3054 | 0.4401 | 0.5000 |

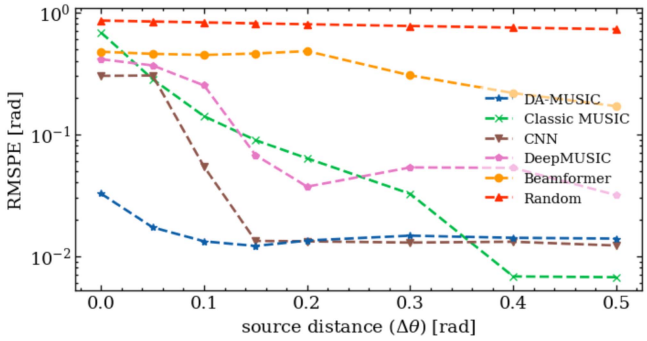


Fig. 7. RMSPE of DoA estimation of $D = 2$ closely spaced sources to analyze the resolution of the methods.

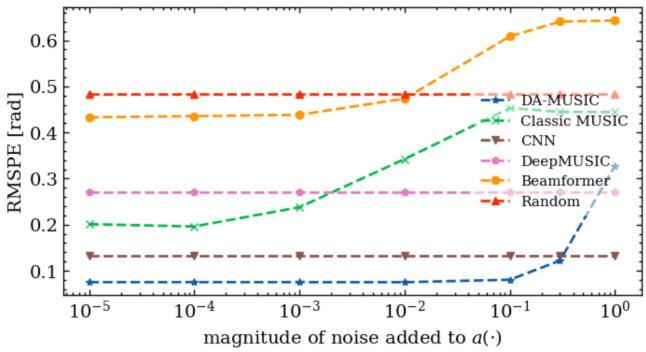


Fig. 8. RMSPE of DoA estimation with various levels of mismatch in the array geometry for $D = 5$ sources.

### A. Synthetic Narrowband

The numerical evaluations of synthetic data presented below are obtained by simulating the measurements $\mathbf{x}(t)$ according to the narrowband system model (1). In particular, we simulate a ULA with $M = 8$ array elements that measure impinging waveforms originating from the DoA angles $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_D]$, which are separately drawn from the uniform distribution $\mathcal{U}(-\pi/2, \pi/2)$. The signals $\mathbf{s}(t) = [s_1(t) \ldots s_D(t)]^\top$ are each drawn randomly from the complex Gaussian distribution $\mathcal{CN}(0, 1)$ for all $t$ modeling random amplitudes and phases. The noises measured at the $M$ array elements $\mathbf{v}(t) = [v_1(t) \ldots v_M(t)]^\top$ are also drawn from $\mathcal{CN}(0, 1)$ for all $t$, followed by appropriate scaling to meet the constant SNR. In the coherent cases, all signals have identical amplitudes and phases, and when not stated otherwise, the simulation parameters are set according to Table I.

*1) Known Number of Sources:* We first evaluate the RMSPE in [rad] achieved when the number of sources $D$ is known in the synthetic narrowband scenario described above. The results, reported in Table II, compare the performance of the following DoA estimators for a different number of sources $D$:

- The DA-MUSIC architecture is implemented according to Fig. 3 and trained separately for each case $D$.
- The classic MB MUSIC algorithm, implemented as described in Section II-D, utilizes the external knowledge of $D$.
- The CNN architecture of [18], with a reduced stride in the first layer to account for $M = 8$, and an increased grid-size for the last layer to achieve $R = 360$.
- The DD DeepMUSIC proposed in [26], while incorporating minor alterations that were necessary to accommodate for the difference in the setup, includes tuning of individual hyperparameters to assure successful training of the CNNs.

The above algorithms are compared with random guessing of the DoA angles. The results in Table II show that the proposed DA-MUSIC notably outperforms all considered benchmarks, notably surpassing the MB MUSIC algorithm not only for coherent sources but also for non-coherent ones, which is the scenario for which the MB algorithm is designed.

To compare the resolution of the algorithms, Fig. 7 shows the RMSPE for localizing $D = 2$ non-coherent signals, which are located close together at a $\Delta\theta$ distance from each other. The MB MUSIC algorithm is shown to collapse when the angular difference approaches $\Delta\theta \approx 0.1$ radians, while DA-MUSIC demonstrates a constant low error for all $\Delta\theta$, indicating its improved resolution.

Next, we evaluate DA-MUSIC in the presence of a mismatch in the array geometry. Fig. 8 depicts the RMSPE achieved when each element of the steering vector $\mathbf{a}(\cdot)$ is corrupted with zero-mean Gaussian noise, leading to a mismatch from the values used to compute the spatial spectrum. Indicating improved robustness, DA-MUSIC is shown to overcome such mismatches in the array geometry from the data. The CNN, DeepMUSIC, and the Random algorithm remain unaffected as they are independent of $\mathbf{a}(\cdot)$.

Fig. 9 depicts the performances differences when localizing $D = 5$ non-coherent signals with $T = 200$ snapshots available for different SNR levels in the range of $[-20, 20]$ dB. DA-MUSIC shows a constant low SNR for positive dB settings and without any fluctuations which slowly decreases with increasing SNR.

Fig. 10 depicts the performance degradation of the estimators with fewer snapshots available. The DD estimators are only trained for the case $T = 200$ as indicated by the circle around the $T = 200$ marker, yet manage to operate with shorter sequences during inference due to the recurrent unit or by taking the covariance matrix as input.

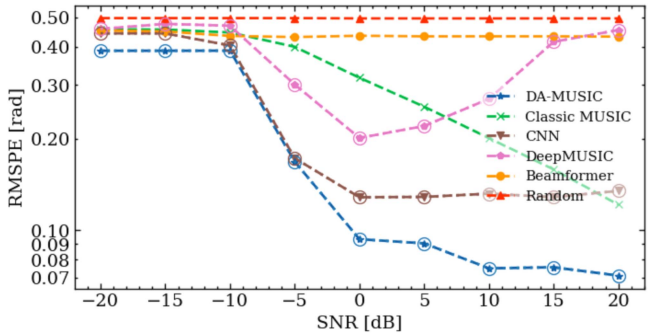


Fig. 9. RMSPE of DoA estimation of $D = 5$ signals with different SNRs.

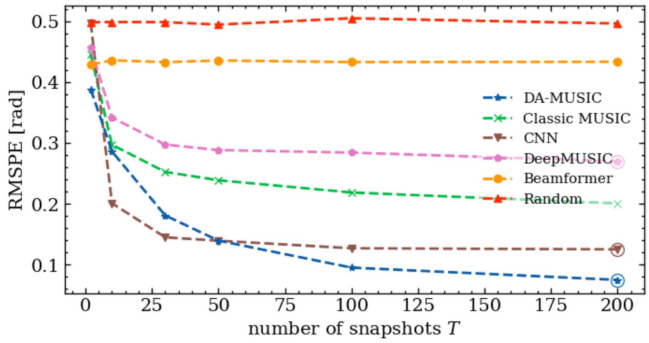


Fig. 10. RMSPE of DoA estimation with $D = 5$ signals with different number of snapshots $T$.

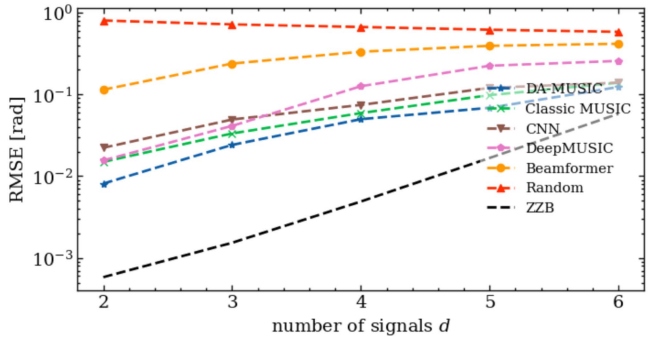


Fig. 11. RMSE obtained for multi-source DoA estimation (in the angle range $[-\pi/3, \pi/3)$ and a source spacing of at least one degree).

We conclude the evaluation of DA-MUSIC with a known number of sources by comparing the performances of the estimators with the common RMSE measure in Fig. 11. Additionally, we provide a comparison with the Ziv-Zakai bound for multi-source DoA estimation [59], [60]. The plot shows performances for different number of sources in a scenario with $T = 200$ snapshots, SNR $= 10$ dB, DoA angle range in $[-\pi/3, \pi/3)$, and a source spacing of at least one degree. These results demonstrate that among all considered DoA estimators, DA-MUSIC is consistently closest to the Ziv-Zakai bound (denoted ZZB in Fig. 11), though all estimators are within some gap of this lower bound.

TABLE III
RMSPE OF DIFFERENT DOA ESTIMATION ALGORITHMS WITH VARYING AND UNKNOWN $D$ FOR $T = 200$ SNAPSHOTS

| RMSPE [rad] | DA-MUSIC | CNN | Classic MUSIC | Beamformer | Random |
|---|---|---|---|---|---|
| non-coherent | | | | | |
| $D = 2$ | 0.0430 | **0.0403** | 0.0428 | 0.1900 | 0.8318 |
| $D = 3$ | **0.0705** | 0.0842 | 0.0917 | 0.2906 | 0.6981 |
| $D = 4$ | **0.0894** | 0.1463 | 0.1489 | 0.3757 | 0.6021 |
| $D = 5$ | **0.1222** | 0.1847 | 0.1856 | 0.4029 | 0.5357 |
| coherent | | | | | |
| $D = 2$ | 0.0383 | **0.0376** | 0.6139 | 0.1670 | 0.8243 |
| $D = 3$ | **0.0688** | 0.0729 | 0.5743 | 0.2843 | 0.6962 |
| $D = 4$ | **0.0869** | 0.1181 | 0.5258 | 0.3883 | 0.6051 |
| $D = 5$ | **0.1046** | 0.1652 | 0.4744 | 0.4179 | 0.5435 |

*2) Unknown and Varying Number of Sources:* Table III shows the results obtained in the exact same narrowband scenario introduced in Section IV-A above but with an unknown and varying number of sources $D \in \{2, \ldots, 5\}$. Here, during inference, the DoA estimation algorithms do not have any knowledge of the varying number of sources present. To compute the RMSPE in such settings (i.e., if the DoA estimators output the wrong number of sources $\hat{D}$), the DoAs are either truncated (least dominant peaks for the MB algorithms and the CNN and highest indexed DoA angles for DA-MUSIC) or padded with random DoA angels until $|\hat{\boldsymbol{\theta}}| = |\boldsymbol{\theta}| = D$. We compare the following DoA estimators:

- The DA-MUSIC architecture is implemented according to Fig. 5 with the classifier which predicts the number of sources being trained along with the overall DoA estimation method.
- The DA-MUSIC (RTC) variation is also implemented according to Fig. 5, yet with a retroactively trained classifier (RTC), i.e., we first train DA-MUSIC for a specific scenario, and then we fix the DoA estimator modules and train only the classification network for alternate scenarios.
- The Classic MUSIC algorithm is implemented as before but determines the number of sources by estimating the multiplicity of the smallest eigenvalue utilizing a predetermined threshold.
- The CNN of [18] as introduced above.
- A conventional beamformer [8], utilizing a peak-finder to estimate the number of sources by determining the number of dominant peaks.
- The Random algorithm corresponds to the base performance when choosing DoA angles at random.

Figs. 12 and 13 show the accuracy in identifying the number of sources versus the number of snapshots $T$ obtained by the mentioned algorithms during the estimation of $\hat{D}$ for non-coherent and coherent signals, respectively. Again, the DD estimators are only trained for the case $T = 200$ as indicated by the circle. Unfortunately, the performance of the internal classifier of DA-MUSIC is dependent on the number of snapshots, and to be able to maintain a more constant accuracy it must be trained for each case. Consequently, DA-MUSIC (RTC), which trains its classifier retroactively, requires separate training for each

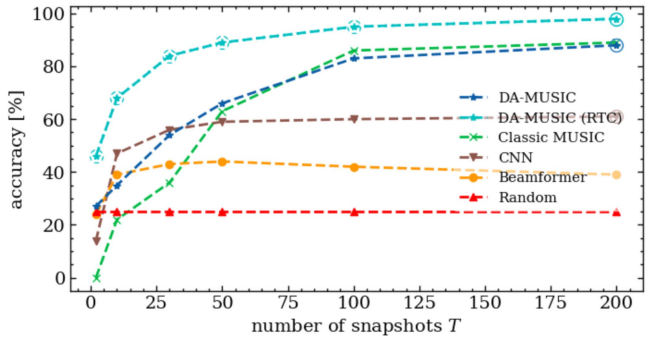


Fig. 12. Accuracy of estimating $\hat{D}$ for various $T$ with $D \in \{2, \ldots, 5\}$ non-coherent narrowband signals.

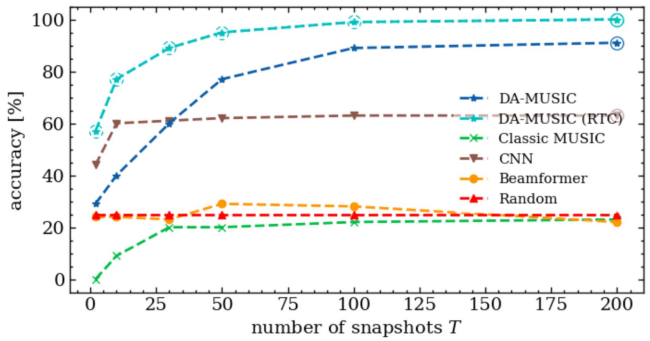


Fig. 13. Accuracy of estimating $\hat{D}$ for various $T$ with $D \in \{2, \ldots, 5\}$ coherent narrowband signals.

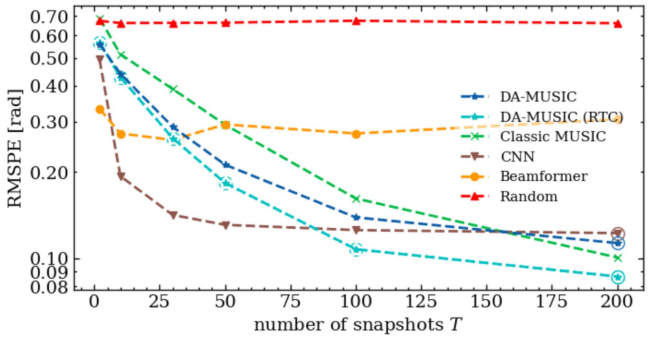


Fig. 14. RMSPE for various $T$ with varying and unknown number of $D \in \{2, \ldots, 5\}$ non-coherent narrowband signals.

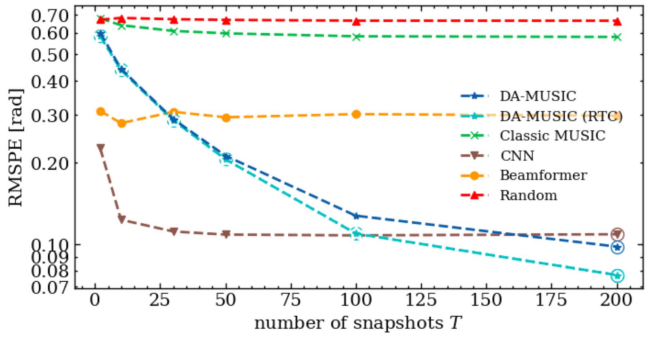


Fig. 15. RMSPE for various $T$ with varying and unknown number of $D \in \{2, \ldots, 5\}$ coherent narrowband signals.

number of snapshots, yet manages to achieve the most accurate predictions.

The corresponding performances in localizing the varying number of sources are depicted in Figs. 14 and 15 for non-coherent and coherent signals, respectively. The shown RMSPE is an average over all considered $D \in \{2, \ldots, 5\}$. The fundamental limitation of the MB MUSIC structure to estimate the number of signal sources for coherent signals also severely impacts the localization abilities of the algorithm.

### B. Synthetic Broadband

We proceed to evaluate DA-MUSIC in a broadband setting. The previously introduced synthetic environment requires certain alterations and assumptions to account for broadband signals. The sensor elements of the receiving array are adequately spaced, and the element spacing is therefore assumed to be

$$\Delta m = \frac{\ell_{\min}}{2} = \frac{1}{2}\,\frac{c}{f_{\max}}, \tag{17}$$

where $\ell_{\min}$ is the minimal wavelength corresponding to the maximal occurring frequency $f_{\max}$ and the frequency spectrum of interest is considered to be within $[f_{\min}, f_{\max}]$. The measurements are simulated utilizing the broadband system model (7), where the elements of $\mathcal{S}(\omega)$ and $\mathcal{W}(\omega)$ are the $N_f$-point FFTs of the elements of $\mathbf{s}(t)$ and $\mathbf{w}(t)$. The parameter values of the different broadband scenarios are specified in Table IV if not specified otherwise. We simulate the following DoA estimators:

- The DA-MUSIC architecture is implemented according to Fig. 3, but the GRU parameters are scaled (having 10 times more parameters available) to enable optimal learning despite the more complex broadband scenarios.
- The classic MUSIC algorithm is implemented in its narrowband format as described in Section II-D and utilizes steering vectors calibrated to the exact array element spacing using $\ell_{\min}/2$.
- Broadband MUSIC corresponds to an incoherent broadband extension of MUSIC [46] and is implemented using 10 [Hz] per IFB; i.e., $|\omega_b - \omega_{b-1}| = 10$ [Hz] for all $b \in \{1, \ldots, B\}$.
- The DoA estimators are again compared to choosing DoA angles at random.

We consider the following three different signal models of the broadband signals $s_d(t)$ for $d \in \{1, \ldots, D\}$:

*1) Broadband Scenario 1:* A broadband signal is obtained as narrowband signals modulated on different carrier frequencies, i.e.,

$$s_d(t) = \bar{s}_d \, \exp\left(2\pi j f_{c,d} t\right), \tag{18}$$

where for each $d \in \{1, \ldots, D\}$, both $\bar{s}_d$ and $f_{c,d}$ are randomly drawn from $\mathcal{CN}(0, 1)$ and $\mathcal{U}(f_{\min}, f_{\max})$ respectively.

*2) Broadband Scenario 2:* Broadband signals are obtained via orthogonal frequency division multiplexing (OFDM) [61], which are modulated on the same carrier frequency. The signals

TABLE IV
SIMULATION PARAMETERS OF THE BROADBAND SCENARIOS

| | Parameter Description | Notation | Value |
|---|---|---|---|
| Scenario 1 | Modulation frequency | $f_{c,d}$ | $0-999$ [Hz] |
| Scenario 2 | Number of subcarriers | $K$ | 1000 |
| | Signal bandwidth | $\Delta f_d$ | 1000 [Hz] |
| Scenario 3 | Modulation frequency | $f_{c,d}$ | $0-899$ [Hz] |
| | Number of subcarriers | $K$ | 10 |
| | Signal bandwidth | $\Delta f_d$ | 100 [Hz] |

are considered in baseband and take the following form

$$s_{d,\text{OFDM}} = \frac{1}{K}\sum_{k=0}^{K-1} \bar{s}_k \ \exp\left(2\pi jk\Delta f_d t/K\right), \tag{19}$$

where $\bar{s}_k \sim \mathcal{CN}(0,1)$ is randomly drawn for each of the $K$ subcarriers and the bandwidth is $\Delta f_d = f_{\max} - f_{\min}$.

*3) Broadband Scenario 3:* A combination of the two previous scenarios and consists of OFDM signals modulated on different carrier frequencies

$$s_d(t) = \exp\left(2\pi j f_{c,d} t\right)\, s_{d,\text{OFDM}}, \tag{20}$$

where $f_{c,d}$ is drawn randomly from $\mathcal{U}(f_{\min}, f_{\max} - \Delta f_d)$ to account for the signal bandwidth $\Delta f_d$.

*Results:* Figs. 16, 17, and 18 present the RMSPE obtained when localizing $D = 2$ broadband signals from Broadband Scenario 1, 2, and 3 respectively. The number of snapshots goes as high as the sampling frequency $f_s = 2000$ [Hz] and is given logarithmically. This high number is suitable for the MB broadband MUSIC algorithm to achieve reliable transformation from the time domain to the frequency domain. DA-MUSIC is again only trained for the case $T = 200$, as indicated by a circle around the marker, yet manages to perform similarly well with a higher number of snapshots. As expected, the classic narrowband MUSIC algorithm completely fails to operate with these broadband signals, while DA-MUSIC consistently achieves the most accurate estimates, outperforming the MB broadband MUSIC algorithm, except for Broadband Scenario 3 with a very large number of snapshots $T > 10^3$, where DA-MUSIC trained with much shorter sequences is slightly outperformed by the MB estimator. These results demonstrate the suitability of DA-MUSIC for coping with broadband scenarios with a limited number of observations.

Figs. 19, 20, and 21 analyze the performances with differently sized frequency ranges $[f_{\min}, f_{\max}]$. It is noted that the architecture of DA-MUSIC is almost invariant towards such scalings and manages to handle signals during inference with much larger bandwidths or modulated with higher carrier frequencies than the signals of the training data. Specifically, DA-MUSIC is only trained for signals with carrier frequencies and bandwidths within 0 to 1000 [Hz] as indicated by the circle around the marker. The results in Figs. 19–21 show that DA-MUSIC, whose complexity is fixed, learns to achieve the most accurate estimates, outperforming the MB broadband MUSIC; the latter requires an increase of the number of IFBs to overcome a larger frequency range and has a constant 10 [Hz] per bin in the depicted results leading to a severe increase in computational complexity.

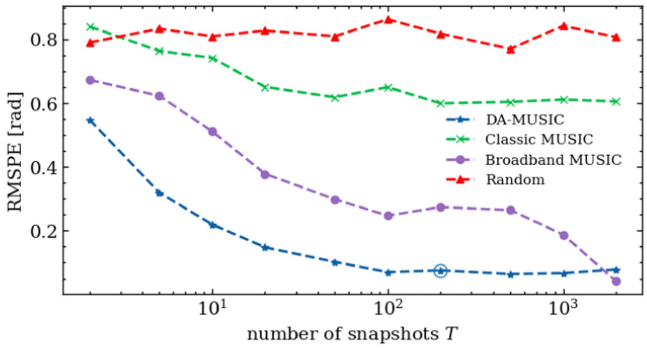


Fig. 16. RMSPE for various $T$ with known number of $D = 2$ signals from Broadband Scenario 1.

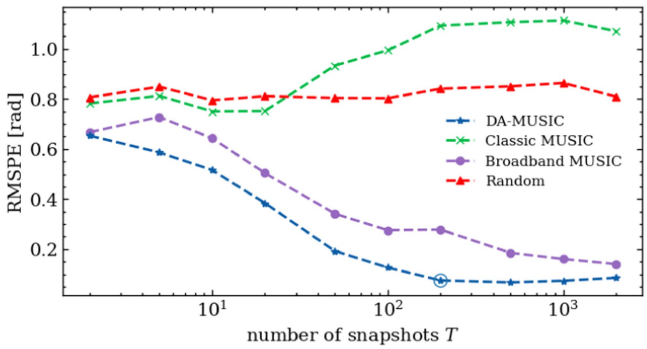


Fig. 17. RMSPE for various $T$ with known number of $D = 2$ signals from Broadband Scenario 2.

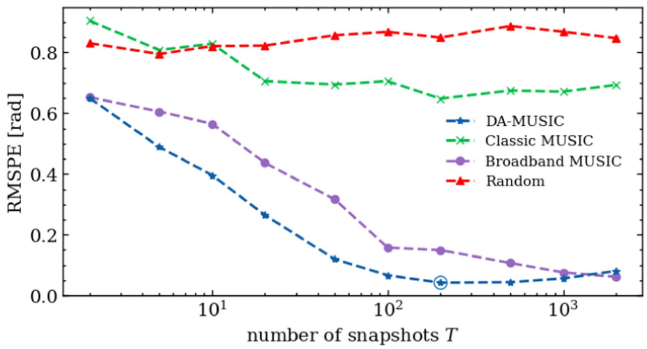


Fig. 18. RMSPE for various $T$ with known number of $D = 2$ signals from Broadband Scenario 3.

### C. Non-Synthetic Data: Azimuth Estimation in Seismic Arrays

In this section, we demonstrate the feasibility and the performance of DA-MUSIC in processing non-synthetic seismic data. The seismic data were recorded by the German Experimental Seismic System (GERES) array located in the Bavarian Forest, Germany. GERES is part of the Comprehensive nuclear-Test Ban Treaty Organization (CTBTO) international monitoring system, and is a well-maintained and calibrated station. Data from GERES is continuously streamed to the International Data Centre (IDC) of the CTBTO, where it is analyzed. The array is composed of 25 vertical seismometers with a minimal distance

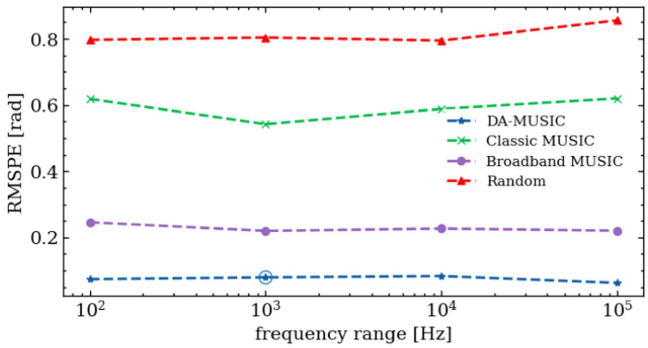


Fig. 19. RMSPE across different frequency ranges for the carrier frequencies of $D = 2$ signals from Broadband Scenario 1.

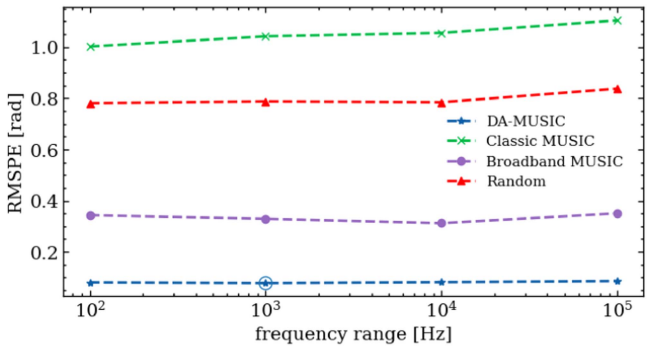


Fig. 20. RMSPE across different frequency ranges for the bandwidth of $D = 2$ signals from Broadband Scenario 2.

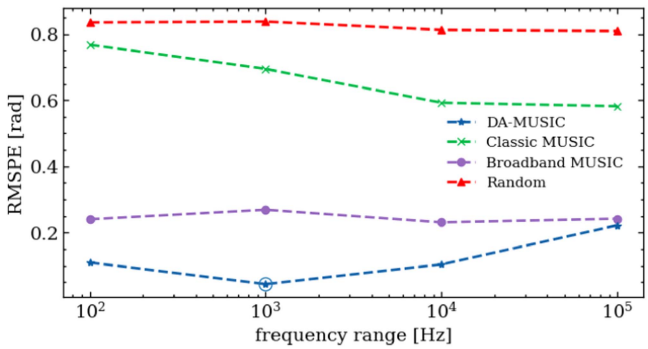


Fig. 21. RMSPE across different frequency ranges for the carrier frequencies of $D = 2$ signals from Broadband Scenario 3.

between the sites of 124 [m] and an aperture of approximately 2.13 [km]. The seismic signal at each sensor is sampled at 40 [Hz]. Details about the GERES arrays and the exact array configuration can be found in [62].

We use data from October to December 2021, in which GERES detected arrivals for 2904 events of which 2816 were used during this analysis, 2534 for training, and 282 for testing. We employ sliding windows of length 100 seconds with a shift of 25 seconds around the signal arrival time as designated by the IDC analysis. The following parameters for each event are obtained from the IDC's Reviewed Event Bulletin: the azimuth DoA angle $\theta$, the slowness value $u$, and the sensor positions

TABLE V
RMSPE IN [RAD] OF DIFFERENT DOA ESTIMATION ALGORITHMS FOR SEISMIC DATA

| | DA-MUSIC | Broadband MUSIC | Classic MUSIC | Beam-former | Random |
|---|---|---|---|---|---|
| RMSPE | **0.6269** | 1.0075 | 1.2475 | 0.9383 | 1.7097 |

$\mathbf{r}_1, \ldots, \mathbf{r}_M$. Using these parameters, the steering vectors take the following form:

$$\mathbf{a}(f, u, \psi, \alpha) = \left[e^{-j2\pi f u \mathbf{r}_1 \mathbf{k}(\psi,\alpha)}, \ \ldots, \ e^{-j2\pi f u \mathbf{r}_M \mathbf{k}(\psi,\alpha)}\right],$$

for some frequency $f$ and with the wave vector for certain elevation $\alpha$ and azimuth $\psi$ of interest,

$$\mathbf{k}(\psi, \alpha) = [\sin\alpha\cos\psi, \sin\alpha\sin\psi, \cos\alpha].$$

We compare DA-MUSIC with following DoA estimators for the setting settings $\alpha = -\pi/4$ and $f = 1$ [Hz]:

- Broadband MUSIC, corresponding to an incoherent broadband extension of MUSIC [46], and instead of using the constant $f = 1$ [Hz] it utilizes 10 IFB with frequencies in [0, 20] [Hz].
- The classic MUSIC algorithm, implemented in its narrowband format as described in Section II-D with additionally filtering the measurements with an experimentally calibrated low-pass filter, allowing only frequencies within [0, 10] [Hz] to pass.
- A conventional beamformer [8].
- Choosing a DoA angle at random.

The results, reported in Table V, show that DA-MUSIC manages to outperform the MB estimators by not only focusing the frequency component but also by concurrently focusing the interdependent elevation angle to receive a stable azimuth estimation. On the other hand, the MB algorithms require further knowledge of the elevation and frequency at hand to operate reliably. While the errors in Table V may appear to be relatively large, it is noted that the average error achieved via expert analysis reported by the IDC Reviewed Event Bulletin is 0.4243 [rad]. This indicates the ability of DA-MUSIC to achieve comparable results and to notably outperform the MB estimators while operating with simplified and approximated model parameters.

## V. CONCLUSION

We presented a hybrid MB/DD implementation of the MUSIC algorithm for DoA estimation. The proposed DA-MUSIC was shown to mitigate some of the limitations and drawbacks of the classic method. DA-MUSIC is operable with an unknown number of sources and with broadband signals while being adaptable to various scenarios and robust towards severe mismatches in the array geometry. The proposed hybrid MB/DD approach provides a viable alternative in both low and high snapshot domains and shows a remarkable resolution capability compared to both MB and DD benchmarks in various settings.

## ACKNOWLEDGMENT

We thank Dr. Yochai Ben Horin for constructive and valuable joint discussions on the seismic data, and for providing the data.

## REFERENCES

[1] J. P. Merkofer, G. Revach, N. Shlezinger, and R. J. G. van Sloun, "Deep augmented music algorithm for data-driven DoA estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2022, pp. 3598–3602.

[2] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, "The rise of radar for autonomous vehicles: Signal processing solutions and future research directions," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 20–31, Sep. 2019.

[3] D. Rahamim, J. Tabrikian, and R. Shavit, "Source localization using vector sensor array in a multipath environment," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3096–3103, Nov. 2004.

[4] J. Foutz, A. Spanias, and M. K. Banavar, "Narrowband direction of arrival estimation for antenna arrays," *Synth. Lectures Antennas*, vol. 3, no. 1, pp. 1–76, 2008.

[5] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Apr. 1996.

[6] Z. Ahmad and I. Ali, "Three decades of development in DoA estimation technology," *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 12, pp. 6297–6312, 2014.

[7] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[8] M. S. Bartlett, "Smoothing periodograms from time-series with continuous spectra," *Nature*, vol. 161, no. 4096, pp. 686–687, May 1948.

[9] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[10] T. E. Tuncer, T. K. Yasar, and B. Friedlander, "Narrowband and wideband DoA estimation for uniform and nonuniform linear arrays," in *Classical and Modern Direction-of-Arrival Estimation*. Boston, MA, USA: Academic Press, 2009, ch.4, pp. 125–160.

[11] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 823–831, Aug. 1985.

[12] G. Su and M. Morf, "The signal subspace approach for multiple wide-band emitter location," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 6, pp. 1502–1522, Dec. 1983.

[13] P.-A. Grumiaux, S. Kitić, L. Girin, and A. Guérin, "A survey of sound source localization with deep learning methods," *J. Acoustical Soc. Amer.*, vol. 152, no. 1, pp. 107–151, 2022.

[14] S. Chakrabarty and E. A. P. Habets, "Multi-speaker DoA estimation using deep convolutional networks trained with noise signals," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 1, pp. 8–21, Mar. 2019.

[15] 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.

[16] L. li Wu, Z. M. Liu, and Z. T. Huang, "Deep convolution network for direction of arrival estimation with sparse prior," *IEEE Signal Process. Lett.*, vol. 26, no. 11, pp. 1688–1692, Nov. 2019.

[17] H. Hammer, S. E. Chazan, J. Goldberger, and S. Gannot, "Dynamically localizing multiple speakers based on the time-frequency domain," *EURASIP J. Audio, Speech, Music Process.*, vol. 2021, no. 1, pp. 1–10, 2021.

[18] G. K. Papageorgiou, M. Sellathurai, and Y. C. Eldar, "Deep networks for direction-of-arrival estimation in low SNR," *IEEE Trans. Signal Process.*, vol. 69, pp. 3714–3729, 2021.

[19] A. Barthelme and W. Utschick, "A machine learning approach to DoA estimation and model order selection for antenna arrays with subarray sampling," *IEEE Trans. Signal Process.*, vol. 69, pp. 3075–3087, 2021.

[20] 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.

[21] M. L. L. de Oliveira and M. Bekooij, "Deep-MLE: Fusion between a neural network and MLE for a single snapshot DOA estimation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2022, pp. 3673–3677.

[22] M. L. L. de Oliveira and M. Bekooij, "ResNet applied for a single-snapshot DOA estimation," in *Proc. IEEE Radar Conf.*, 2022, pp. 1–6.

[23] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *Proc. IEEE*, vol. 111, no. 5, pp. 465–499, May 2023.

[24] N. Shlezinger, Y. C. Eldar, and S. P. Boyd, "Model-based deep learning: On the intersection of deep learning and optimization," *IEEE Access*, vol. 10, pp. 115384–115398, 2022.

[25] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *Proc. IEEE*, vol. 111, no. 5, pp. 465–499, May 2023, doi: 10.1109/JPROC.2023.3247480.

[26] A. M. Elbir, "DeepMUSIC: Multiple signal classification via deep learning," *IEEE Sensors Lett.*, vol. 4, no. 4, pp. 1–4, Apr. 2020.

[27] C. Liu, W. Feng, H. Li, and H. Zhu, "Single snapshot DOA estimation based on spatial smoothing MUSIC and CNN," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput*, 2021, pp. 1–5.

[28] A. Barthelme and W. Utschick, "DoA estimation using neural network-based covariance matrix reconstruction," *IEEE Signal Process. Lett.*, vol. 28, pp. 783–787, 2021.

[29] D. T. Hoang and K. Lee, "Deep learning-aided coherent direction-of-arrival estimation with the FTMR algorithm," *IEEE Trans. Signal Process.*, vol. 70, pp. 1118–1130, 2022.

[30] W. Zhang, Y. Han, M. Jin, and X. Qiao, "Multiple-toeplitz matrices reconstruction algorithm for DoA estimation of coherent signals," *IEEE Access*, vol. 7, pp. 49504–49512, 2019.

[31] S. Chakrabarty and E. A. Habets, "Broadband DoA estimation using convolutional neural networks trained with noise signals," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, 2017, pp. 136–140.

[32] W. Zhu and M. Zhang, "A deep learning architecture for broadband DoA estimation," in *Proc. IEEE 19th Int. Conf. Commun. Technol.*, 2019, pp. 244–247.

[33] L. li Wu and Z. T. Huang, "Coherent SVR learning for wideband direction-of-arrival estimation," *IEEE Signal Process. Lett.*, vol. 26, no. 4, pp. 642–646, Apr. 2019.

[34] B. Friedlander, "Wireless direction-finding fundamentals," in *Classical and Modern Direction-of-Arrival Estimation*, T. E. Tuncer and B. Friedlander, Eds. Boston, MA, USA: Academic Press, 2009, ch. 1, pp. 1–51.

[35] Z. Chen, G. Gokeda, and Y. Yu, *Introduction to Direction-of-Arrival Estimation*. Norwood, MA, USA: Artech House, 2010.

[36] S. Ge, K. Li, and S. N. M. Rum, "Deep learning approach in DoA estimation: A systematic literature review," *Mob. Inf. Syst.*, vol. 2021, pp. 6392875:1–6392875: 14, 2021 .

[37] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[38] B. Friedlander, "The root-music algorithm for direction finding with interpolated arrays," *Signal Process.*, vol. 30, no. 1, pp. 15–29, 1993.

[39] T.-Jun Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 806–811, Aug. 1985.

[40] Q. Chen and R. Liu, "On the explanation of spatial smoothing in music algorithm for coherent sources," in *Proc. IEEE Int. Conf. Inf. Sci. Technol.*, 2011, pp. 699–702.

[41] A. Paulraj, R. Roy, and T. Kailath, "Estimation of signal parameters via rotational invariance techniques- ESPRIT," in *Proc. 19th Asilomar Conf. Circuits, Syst. Comput.*, 1985, pp. 83–89.

[42] V. Vasylyshyn, "Direction of arrival estimation using ESPRIT with sparse arrays," in *Proc. IEEE Eur. Radar Conf.*, 2009, pp. 246–249.

[43] M. Wax, T.-J. Shan, and T. Kailath, "Spatio-temporal spectral analysis by eigenstructure methods," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 4, pp. 817–827, Aug. 1984.

[44] S. Chandran and M. K. Ibrahim, "DoA estimation of wide-band signals based on time-frequency analysis," *IEEE J. Ocean. Eng.*, vol. 24, no. 1, pp. 116–121, Jan. 1999.

[45] S. Argentieri and P. Danes, "Broadband variations of the music high-resolution method for sound source localization in robotics," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2007, pp. 2009–2014.

[46] Y.-S. Yoon, L. M. Kaplan, and J. H. McClellan, "DoA estimation of wideband signals," in *Proc. Adv. Direction-of-Arrival Estimation*, Artech, 2006, pp. 49–68.

[47] J. Krolik and D. Swingler, "Focused wide-band array processing by spatial resampling," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 2, pp. 356–360, Feb. 1990.

[48] T.-S. Lee, "Efficient wideband source localization using beamforming invariance technique," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1376–1387, Jun. 1994.

[49] B. Friedlander and A. Weiss, "Direction finding for wideband signals using an interpolated array," in *Proc. 25th Asilomar Conf. Signals, Syst., Comput.*, 1991, pp. 583–587.

[50] E. D. di Claudio and R. Parisi, "WAVES: Weighted average of signal subspaces for robust wideband direction finding," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2179–2191, Oct. 2001.
[51] Y.-S. Yoon, L. M. Kaplan, and J. H. McClellan, "TOPS: New DOA estimator for wideband signals," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 1977–1989, Jun. 2006.
[52] F. Ma and X. Zhang, "Wideband DOA estimation based on focusing signal subspace," *Signal, Image Video Process.*, vol. 13, PP. 675–682, 2019.
[53] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall International Editions, 1993, pp. 27–37.
[54] H. Krim and J. Proakis, "Smoothed eigenspace-based parameter estimation," *Automatica*, vol. 30, no. 1, pp. 27–38, 1994.
[55] O. Solomon et al., "Deep unfolded robust PCA with application to clutter suppression in ultrasound," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1051–1063, Apr. 2020.
[56] T. Routtenberg and J. Tabrikian, "Bayesian parameter estimation using periodic cost functions," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1229–1240, Mar. 2012.
[57] T. Routtenberg and J. Tabrikian, "Non-Bayesian periodic Cramér-Rao bound," *IEEE Trans. Signal Process.*, vol. 61, no. 4, pp. 1019–1032, Apr. 2013.
[58] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[59] Z. Zhang, Z. Shi, and Y. Gu, "Ziv-Zakai bound for DOAs estimation," *IEEE Trans. Signal Process.*, vol. 71, pp. 136–149, 2023.
[60] T. Routtenberg and J. Tabrikian, "A general class of outage error probability lower bounds in Bayesian parameter estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2152–2166, May 2012.
[61] S. B. Weinstein, "The history of orthogonal frequency-division multiplexing [history of communications]," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 26–35, Nov. 2009.
[62] H.-P. Harjes, "Design and siting of a new regional array in central Europe," *Bull. Seismological Soc. America*, vol. 80, pp. 1801–1817, 1990.

**Julian P. Merkofer** (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from the Federal Institute of Technology in Zurich (ETHZ), Zurich, Switzerland, in 2019, and 2021, respectively. He is currently working toward the Ph.D. degree with the Biomedical Diagnostics Laboratory, Eindhoven University of Technology, Eindhoven, The Netherlands. His research interests include machine learning for signal processing, model-based deep learning, and magnetic resonance spectroscopy.

**Guy Revach** (Graduate Student Member, IEEE) received the B.Sc. (*cum laude*) and M.Sc. degrees from the Andrew and Erna Viterbi Department of Electrical and Computer Engineering, Technion – Israel Institute of Technology, Haifa, Israel, in 2008 and 2017, respectively. Since 2019, he has been working toward the Ph.D. degree with the Institute for Signal and Information Processing (ISI), ETH Zürich, Zürich, Switzerland, supervised by Dr. Hans-Andrea Loeliger. He did his master's thesis under the supervision of Prof. Nahum Shimkin on planning for cooperative multiagents. He is currently a Researcher with a proven industry track record as an innovator and system engineer. His research interests include the intersection of machine learning with signal processing, and more specifically combining sound theoretical principles from classical signal processing with state-of-the-art machine learning algorithms for tracking and detection problems. Before coming to ETH Zürich, he was with the Israeli wireless communication industry for more than ten years, first as a real-time embedded Software Engineer and then a Software Manager. He was the main innovator behind state-of-the-art, software-defined radio (SDR) for wireless communication, game-changing and groundbreaking in terms of size, weight, and power. As a System Engineer, he defined major aspects of SDR requirements and architecture for hardware, software, network, cyber defense, signal processing, data analysis, and control algorithms.

**Nir Shlezinger** (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical and computer engineering from Ben-Gurion University, Beersheba, Israel, in 2011, 2013, and 2017, respectively. He is currently an Assistant Professor wtih the School of Electrical and Computer Engineering, Ben-Gurion University. From 2017 to 2019, he was a Postdoctoral Researcher with the Technion, and from 2019 to 2020, he was a Postdoctoral Researcher with the Weizmann Institute of Science, where he was awarded the FGS prize for outstanding research achievements. His research interests include communications, information theory, signal processing, and machine learning.

**Tirza Routtenberg** (Senior Member, IEEE) received the B.Sc. degree from the Technion Israel Institute of Technology, Haifa, Israel, in 2005, and the M.Sc. and Ph.D. degrees in electrical and computer engineering from Ben-Gurion University, Beersheba, Israel, in 2007 and 2012, respectively. She is currently an Associate Professor with the School of Electrical and Computer Engineering, Ben-Gurion University, Beersheba, Israel. She was a Postdoctoral Fellow with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA, from 2012 to 2014. Since 2014, she has been a faculty Member with the School of Electrical and Computer Engineering, and Ben-Gurion University of the Negev, Beer-Sheva. During 2022–2023, she was a William R. Kenan, Jr. Visiting Professor for Distinguished Teaching with Princeton University, Princeton, NJ, USA. Her research interests include signal processing in smart grids, statistical signal processing, estimation and detection theory, and signal processing on graphs. She is an Associate Editor for IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKSand IEEE SIGNAL PROCESSING LETTERS, and a Member of the IEEE Signal Processing Theory and Methods Technical Committee. She was the recipient of the Best Student Paper Award at the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2011, IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) 2013 (co-author), ICASSP 2017 (co-author), and IEEE Workshop on Statistical Signal Processing (SSP) 2018 (co-author). She was awarded the Negev scholarship in 2008, Lev-Zion scholarship in 2010, Marc Rich Foundation Prize in 2011, and Toronto Prize for Excellence in Research in 2021.

**Ruud J. G. van Sloun** (Member, IEEE) received the M.Sc. and Ph.D. degree (*cum laude*) in electrical engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2014, and 2018, respectively. He is currently an Associate Professor with the Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands. From 2019 to 2020, he was a Visiting Professor with the Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel, and from 2020 to 2023, he was a kickstart AI Fellow with Philips Research. His current research interests include closed-loop image formation, deep learning for signal processing and imaging, active signal acquisition, model-based deep learning, compressed sensing, ultrasound imaging, and probabilistic signal and image reconstruction. He was the recipient of an ERC starting grant, NWO VIDI grant, NWO Rubicon grant, and Google Faculty Research Award.