

Deep Unfolding of Atomic Norm Minimization for DoA Estimation

Ali Raza*, Nhan Thanh Nguyen*, and Markku Juntti*

*Center for Wireless Communications (CWC), University of Oulu, Oulu 90014, Finland

Email: {ali.raza,nhan.nguyen,markku.juntti}@oulu.fi

Abstract—Direction-of-arrival (DoA) estimation is a critical problem in array signal processing with applications in radar sensing and mobile positioning. Among various approaches, atomic norm minimization (ANM) is an effective technique that promotes sparsity in the angular domain, making it well-suited for DoA estimation. However, the conventional ANM problem is typically solved using semi-definite programming (SDP), which is computationally intensive and limits its practical applicability. To address this challenge, we propose a novel DoA estimation approach based on deep unfolding of the ANM problem. Specifically, we introduce an approximate function for the atomic norm and solve the resulting problem using a deep unfolded projected gradient descent (PGD) method. In this framework, the step size of the PGD and the parameters of the approximated ANM function are optimized via data training to minimize the objective function efficiently. The proposed method unifies the PGD optimization with deep learning, delivering accurate DoA estimation with low complexity. Simulation results demonstrate that our method achieves competitive performance with state-of-the-art techniques while maintaining computational efficiency, making it a promising solution for practical DoA estimation.

Index Terms—direction of arrival (DOA), projected gradient descent (PGD), atomic norm, deep unfolding.

I. INTRODUCTION

Direction-of-arrival (DoA) estimation is a fundamental problem in array signal processing, enabling accurate source localization in radar, sonar, and wireless communication systems. The conventional subspace-based methods, such as multiple signal classification (MUSIC) [1] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [2], achieve high resolution but face several limitations. More specifically, they are sensitive to the signal-to-noise ratio (SNR), require a large number of snapshots, and depend on prior knowledge of the number of sources. These limitations degrade their performance at low SNRs or in under-determined scenarios where the number of sources exceeds that of sensors.

Sparse signal processing has emerged as a robust alternative, exploiting spatial sparsity. Compressive sensing involves discretizing the angular or frequency space into a grid, which is essential for the recovery of sparse signals. This discretization allows the representation of continuous signals in a manageable format [3]. However, a grid mismatch occurs when the true DoAs deviate from the discretized grid, leading to estimation errors. The atomic norm minimization (ANM) method overcomes this limitation via gridless sparse recovery for super-resolution accuracy [4]. The ANM problem is typically solved via semi-definite programming (SDP). The

ANM was developed in [4] by leveraging the Vandermonde structure of signals, enabling off-grid recovery. Subsequent works extended the ANM to diverse applications, such as wideband underwater acoustic systems [5] and MIMO radar [6], [7], improving robustness in underdetermined cases. Despite its advantages, the ANM's reliance on SDP imposes high computational costs. To mitigate this, the alternating direction method of multipliers (ADMM) was employed to decompose the SDP into tractable subproblems [8]. While the ADMM accelerates the ANM, its limited scalability remains challenging for large antenna arrays.

Deep learning (DL) offers a complementary approach to traditional signal processing techniques, with deep neural networks (DNNs) directly mapping spatial covariance matrices [9]–[11] or raw sampled signals [12] to estimate the DoAs. These conventional black-box DL-based estimators excel in noisy environments but lack interpretability and demand extensive training data. Deep unfolding, a typical model-based DL technique, bridges this gap by embedding numerical algorithms into the DNNs, transforming iterative algorithms into trainable layers [13]. Applied to the ANM, deep unfolding enhances efficiency while preserving super-resolution. For instance, the ANM-ADMM-Net and decoupled ANM-ADMM-Net [14], [15] employ deep unfolding to optimize the ADMM iterations, enabling faster and more accurate 1D and 2D DoA estimation across diverse array geometries. Through data-driven training, these methods achieve faster convergence of the estimator without compromising estimation accuracy. Recent studies [14], [15] demonstrate that deep unfolding methods can address the ANM problem and achieve a similar accuracy as traditional optimization methods but with significantly lower computational complexity.

This paper presents a novel DoA estimation method that advances the state of the art by overcoming the computational limitations of the SDP-based ANM. However, unlike most of existing methods that directly address the ANM problem, we introduce an approximate formulation of the atomic norm that is more tractable, allowing the ANM problem to be solved without the SDP. We then propose a deep unfolded projected gradient descent (PGD) framework to solve the approximate problem. Unlike the conventional solutions, this unfolded PGD method leverages data training to optimize both the step sizes in the PGD updates and the parameters of the approximated atomic norm. As a result, it achieves high accuracy with significantly reduced complexity compared with the SDP-

based ANM. The proposed approach integrates explainable optimization principles with DL, enabling fast yet interpretable DoA estimation. The simulation results demonstrate that our method delivers state-of-the-art performance while remaining computationally efficient and fast execution.

II. SYSTEM MODEL

Consider a uniform linear array (ULA) comprising M antenna elements receiving signals from K far-field sources. The array receives signals originating from K distinct angles, denoted as $\theta_1, \theta_2, \dots, \theta_K$. The inter-element spacing of the ULA is fixed to $d = \lambda/2$, where λ represents the wavelength of the incident signal. Under this configuration, the received signal at the m -th antenna element ($m = 0, 1, \dots, M-1$) at time t can be expressed as

$$y_m(t) = \sum_{k=0}^{K-1} s_k(t) e^{j2\pi \frac{md}{\lambda} \sin \theta_k} + n_m(t), \quad (1)$$

where $s_k(t)$ and $n_m(t)$ denotes the signal transmitted by the k -th source and additive white Gaussian noise (AWGN) sample with distribution $\mathcal{CN}(0, \sigma^2)$, respectively. In (1), the factor $e^{j2\pi \frac{md}{\lambda} \sin \theta_k}$ represents the phase shift due to the DoA θ_k . By aggregating the signals from all M antenna elements, we can rewrite (1) in a compact form as

$$\begin{aligned} \mathbf{y}(t) &= \sum_{k=0}^{K-1} s_k(t) \mathbf{a}(\theta_k) + \mathbf{n}(t) \\ &= \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}(t) + \mathbf{n}(t), \end{aligned} \quad (2)$$

where $\mathbf{y}(t) = [y_0(t), y_1(t), \dots, y_{M-1}(t)]^\top \in \mathbb{C}^{M \times 1}$ is the received signal vector, $\mathbf{s}(t) = [s_0(t), s_1(t), \dots, s_{K-1}(t)]^\top \in \mathbb{C}^{K \times 1}$ denotes the source signal vector, and $\mathbf{n}(t) = [n_0(t), n_1(t), \dots, n_{M-1}(t)]^\top \in \mathbb{C}^{M \times 1}$ is the vector of AWGN samples. In (2), $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)] \in \mathbb{C}^{M \times K}$ is the steering matrix whose k -th column is given as

$$\mathbf{a}(\theta_k) = \left[1, e^{j2\pi \frac{d}{\lambda} \sin(\theta_k)}, \dots, e^{j2\pi \frac{(M-1)d}{\lambda} \sin(\theta_k)} \right]^\top. \quad (3)$$

Let \mathbf{x} be a sparse signal vector corresponding to K distinct DoA components, represented as

$$\mathbf{x} = \sum_{k=0}^{K-1} s_k(t) \mathbf{a}(\theta_k). \quad (4)$$

Each atom $\mathbf{a}(\theta_k)$ belongs to an atomic set \mathcal{A} [6], [16]:

$$\mathcal{A} = \left\{ \mathbf{a}(\theta) : \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right\}. \quad (5)$$

The noiseless data \mathbf{x} can be written as the linear combination of K atoms in \mathcal{A} . The atomic norm of \mathbf{x} is defined as

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf_{\theta_k} \left\{ K \in \mathbb{Z} : \mathbf{x} = \sum_{k=0}^{K-1} s_k \mathbf{a}(\theta_k) \right\}. \quad (6)$$

Given the noisy received signal \mathbf{y} , the denoising problem can be formulated as a sparse signal reconstruction task in the

following ANM problem:

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \beta \|\mathbf{x}\|_{\mathcal{A}}, \quad (7)$$

where $\beta > 0$ is a regularization parameter that controls the trade-off between reconstruction accuracy and sparsity. This formulation imposes a continuous sparse prior on \mathbf{x} , promoting solutions with only a few active DoAs. Conventionally, this ANM problem is solved using the SDP method [4], [17]. SDP solvers are highly complex and slow. In the next section, we propose a computationally efficient model-based ML approach to overcome the limitations of SDP-based solutions.

III. PROPOSED DEEP UNFOLDING METHOD FOR ANM

Our goal is to develop a computationally efficient solution to (7). However, this is hindered by the atomic norm's complex and intractable formulation, as seen in (6). Specifically, it cannot be solved directly using standard convex optimization tools or linearized with first-order Taylor expansions. To address this, we first approximate the ANM with a simpler and more manageable form, which can be efficiently solved with PGD. These will be elaborated next.

A. Approximation of the Atomic Norm and PGD Approach

We propose the quadratic approximation $\|\mathbf{x}\|_{\mathcal{A}} \approx \alpha \|\mathbf{x}\|_2^2$, where $\alpha \geq 0$ cast as a hyperparameter. As a result, the ANM is approximated by the following quadratic problem

$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \beta \alpha \|\mathbf{x}\|_2^2. \quad (8)$$

This relaxation transforms the original objective function into a smooth and differentiable quadratic form, which is much more tractable than the conventional atomic norm. Then, we propose to solve the approximated problem via deep unfolded PGD, a learning-based iterative algorithm that exploits the gradient of the objective function to iteratively refine an estimate of \mathbf{x} . We will discuss the detailed learning aspects in the next section. Here, we focus on developing the PGD framework to solve problem (8).

The gradient of the approximate objective function, i.e., $f(\mathbf{x})$, is derived as

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = (1 + 2\beta\alpha) \mathbf{x} - \mathbf{y}. \quad (9)$$

As a result, the PGD update rule of \mathbf{x} at the $(\ell+1)$ -th iteration is given as

$$\begin{aligned} \mathbf{x}^{\ell+1} &= \mathbf{x}^{\ell} - \eta^{\ell} \nabla_{\mathbf{x}} f(\mathbf{x})|_{\mathbf{x}=\mathbf{x}^{\ell}} \\ &= \mathbf{x}^{\ell} - \eta^{\ell} ((1 + 2\beta\alpha) \mathbf{x}^{\ell} - \mathbf{y}), \end{aligned} \quad (10)$$

where η^{ℓ} denotes the step size. This iterative process continues until convergence is achieved. Finally, the DoA is determined as the peaks of the spatial spectrum

$$\mathcal{S}(\theta) = |\mathbf{a}^H(\theta) \mathbf{x}|^2. \quad (11)$$

The efficiency and convergence of the PGD method depend heavily on the step sizes η^{ℓ} , $\ell = 1, 2, \dots$, and manually tuning these values is challenging. Conventional techniques,

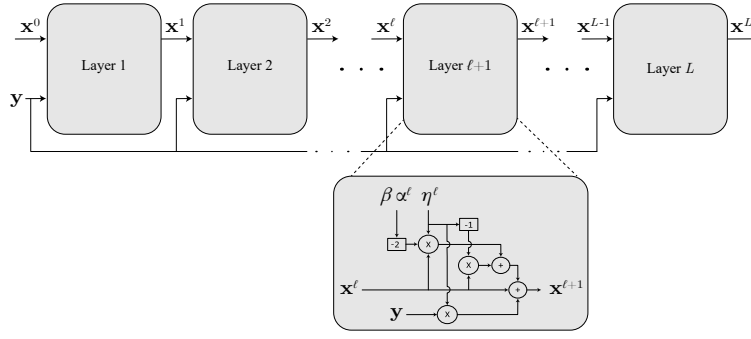


Fig. 1. The network architecture of the proposed deep unfolded PGD method for the ANM.

such as backtracking line search, can help adjust the step sizes to guarantee convergence. However, these methods introduce additional per-iteration complexity and may not necessarily lead to faster convergence and better solution [18]. In addition, from (10), we observe that the algorithm's efficiency also depends on the parameter α , which is used to approximate the atomic norm. In (8), for a given β , the parameter α controls the weight of the term $\|\mathbf{x}\|_2^2$. Hence, α must be properly selected to balance the contributions of the two terms in the overall objective function.

To address the challenges of setting $\{\eta^\ell\}$ and α , we propose modeling the PGD iterations as layers of a DNN, where $\{\eta^\ell\}$ and α are treated as trainable parameters. These parameters are optimized during training to enhance the efficiency of the unfolded PGD method. The detailed architecture and training procedure of this method are discussed in the next section.

B. Deep Unfolded PGD Architecture

1) *Network Architecture*: Let L denote the number of layers in the unfolded DNN model, which is predetermined via fine-tuning and fixed during inference. This contrasts with conventional optimization algorithms, which must iterate until convergence to achieve satisfactory performance. Moreover, the fixed network architecture results in a constant computational complexity, making the proposed approach well-suited for practical implementation. Based on the deep unfolding principle, each iteration of the conventional PGD algorithm is mapped to one layer in the unfolded DNN model. To efficiently set parameter α , we also model it as a layer-specific learning parameter. More specifically, similar to $\{\eta^\ell\}$, $\{\alpha^\ell\}$ are optimized via data training. In this way, the output of the $(\ell + 1)$ -th layer of the unfolded model can be expressed as:

$$\mathbf{x}^{\ell+1} = \mathbf{x}^\ell - \eta^\ell ((1 + 2\beta\alpha^\ell) \mathbf{x}^\ell - \mathbf{y}). \quad (12)$$

From the update procedure in (12), we note the following observations:

- First, unlike existing deep unfolded PGD models where only the step sizes are optimized [19], [20], our proposed method also optimizes the parameters $\{\alpha^\ell\}$ via data training. This enhances the efficiency of the quadratic approximation and improves the accuracy of the DoA estimation in the deep unfolded model.

- Second, as evident from the product $\beta\alpha^\ell$, the regularization term associated with the atomic norm is influenced by $\{\alpha^\ell\}$. Therefore, optimizing $\{\alpha^\ell\}$ through data-driven training enables dynamic regularization, which adaptively balances the two terms in the objective function in (8). This contrasts with conventional methods, where the regularization parameter β is fixed and evaluated as $\sqrt{\sigma^2 M \log M}$, where the σ^2 is the noise variance [21].

It is also observed in (12) that generating $\mathbf{x}^{\ell+1}$ requires \mathbf{x}^ℓ , \mathbf{y} , η^ℓ , and α^ℓ . Here, \mathbf{x}^ℓ and \mathbf{y} serve as the inputs to the ℓ -th layer, while η^ℓ and α^ℓ are layer-specific hyperparameters, which are trained and optimized during the training phase. The network architecture of the proposed unfolded DNN is illustrated in Fig. 1. It shows that the proposed unfolding network closely follows the optimization flow of the PGD approach, ensuring its explainability and stable performance, as will be demonstrated later through simulation results. The details of the model training are discussed next.

2) *Training Loss Function*: In the proposed unfolding PGD approach, $\{\eta^\ell\}$ and $\{\alpha^\ell\}$ are optimized via data training. To this end, we employ the following loss function:

$$\mathcal{L} = \sum_{\ell=1}^L \|\mathbf{y} - \mathbf{x}^\ell\|_2^2 + \beta \alpha^\ell \|\mathbf{x}^\ell\|_2^2. \quad (13)$$

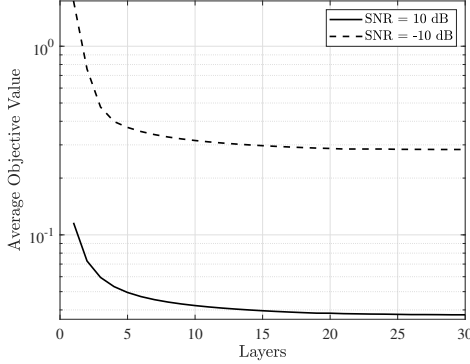
where we recall that \mathbf{x}^ℓ is the output of the ℓ -th layer. The loss function reflects the total objective values of problem (8) achieved by \mathbf{x}^ℓ , for $\ell = 1, \dots, L$. As a result, not only the final output but also those of the earlier layers contribute to the loss, enabling faster convergence of the solution. Furthermore, it is noted that an unsupervised training approach is employed, meaning that ground truths and labels are not required for training the proposed unfolded PGD model.

C. Overall Algorithm

Following the training phase, the trained unfolded network is applied as an online solver for the optimization problem (8). In each layer of the DNN, the estimate is updated based on (12) using the optimized $\{\eta^\ell, \alpha^\ell\}$. After L layers, it outputs \mathbf{x}^L as the final estimate of \mathbf{x} . Finally, the DoA is determined as the angle at which the peaks of the spatial spectrum $\mathcal{S}(\theta) = |\mathbf{a}^H(\theta)\mathbf{x}^L|^2$ are achieved. The operation

Algorithm 1 Deep Unfolded PGD for DoA Estimation

- 1: **Input:** Observation \mathbf{y} , trained parameters $\{\eta^\ell, \alpha^\ell\}_{\ell=1}^L$.
- 2: **Output:** DOAs.
- 3: **Initialize:** $\mathbf{x}^0 = \mathbf{0}$.
- 4: **for** $\ell = 0 \rightarrow L - 1$ **do**
- 5: Compute gradient of $f(\mathbf{x})$ based on (9) at $\mathbf{x} = \mathbf{x}^\ell$.
- 6: Update $\mathbf{x}^{\ell+1}$ based on (12).
- 7: **end for**
- 8: **Estimate DoAs:** Choose the peak values of the spatial spectrum $\mathcal{S}(\theta) = |\mathbf{a}^H(\theta)\mathbf{x}^L|^2$ and set as the DoAs.

**Fig. 2.** Convergence performance of the proposed method.

of the proposed unfolded PGD method is summarized in Algorithm 1.

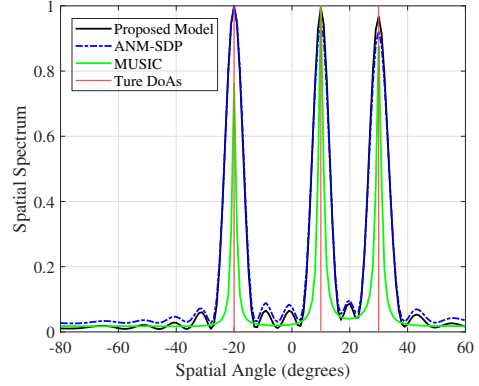
The proposed method has a complexity of only $\mathcal{LO}(M^2)$, which is mostly to compute the gradients and updating \mathbf{x} , as seen in steps 5 and 6 of Algorithm 1. Whereas, the complexity of the conventional ANM-SDP is $I_{\text{SDP}}\mathcal{O}(M^6)$, where I_{SDP} is the number of iterations until convergence [22]. It is clear that the proposed method is much more computational efficient than the conventional SDP counterpart.

IV. SIMULATION RESULTS

The proposed DoA estimation method based on the deep unfolding of the ANM is implemented in Python v3.12 using PyTorch v2.6. The unfolding DNN is trained on a synthetically generated dataset, with key simulation parameters detailed in Table I. We trained the unfolding DNN over 1000 epochs using the SGD optimizer with a learning rate of 0.6×10^{-4} and a batch size of 100. The testing dataset includes 100 samples.

Fig. 2 illustrates the convergence behavior of the proposed method by plotting the average objective value versus the number of unfolding layers with $\text{SNR} = \{10, -10\}$ dB. In both cases, the objective values decrease monotonically and converge after about 20 layers. With $\text{SNR} = 10$ dB, the network converges faster and reaches a much lower objective value than that with $\text{SNR} = -10$ dB.

In Fig. 3, we show the spatial spectrum obtained by the proposed deep unfolding, the ANM-SDP, and the MUSIC approaches. We consider the range $\theta \in [-80^\circ, 60^\circ]$ to evaluate the spectrum based on (11). It is observed that even with quadratic approximation of the original ANM problem, the

**Fig. 3.** Spatial spectrum comparison with different methods.**Table I.** Summary of parameter settings for the simulations.

Parameter	Value
Number of sources K	3
Number of antennas M	16
Number of unfolding layers L	30
Antenna separation d	0.5λ
Snapshots	100
SNR for Fig. 3 and 5	10 dB
True DoAs	$-20^\circ, 10^\circ$, and 30°

proposed method still accurately estimates the true DoAs with performance comparable to the original atomic norm. Moreover, the proposed method offers significantly lower computational complexity than ANM-SDP, as discussed in Section III-C.

In Fig. 4, we compare the performance of the proposed method with ANM-SDP and MUSIC under different SNRs. The estimation performance is measured using the root mean square error (RMSE), computed as

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{sim}}K} \|\boldsymbol{\theta}_{\text{est}} - \boldsymbol{\theta}_{\text{true}}\|_2^2}, \quad (14)$$

where $\boldsymbol{\theta}_{\text{est}}$ and $\boldsymbol{\theta}_{\text{true}}$ represent the estimated and true DoA vectors, respectively, and $N_{\text{sim}} = 1000$ is the number of Monte Carlo simulations. In Fig. 4, it is observed that the proposed method and ANM-SDP exhibit nearly identical performance, both achieving the lowest RMSE values compared to the MUSIC method. An improved performance gap compared to the ANM-SDP is observed between -5 to 10 dB SNRs, during which the proposed method approximately improved by 30.5% and 72.8% at 0 and 5 dB SNRs, respectively.

In Fig. 5, we compare the runtime of the proposed method with the conventional ANM-SDP. It is observed that the proposed method has a faster runtime of only 0.057 s, about 800 times faster than the ANM-SDP (8.06 s) at $M = 128$. Overall, as M increases, the runtime for ANM-SDP increases sharply, while that of the proposed method maintains a relatively low runtime. This agrees with the comparison on computational complexities of the two scheme in Section (III-C) and demonstrate the fast execution of the proposed unfolding method.

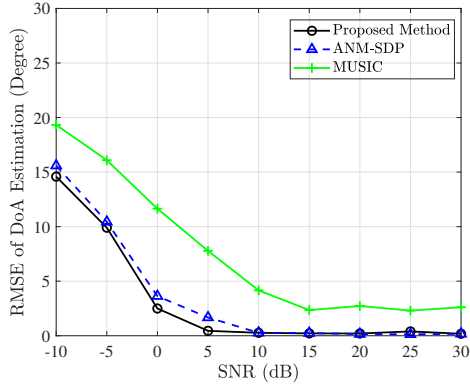


Fig. 4. DoA estimation performance comparison with different methods.

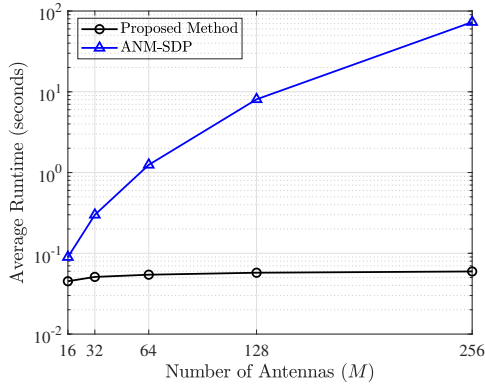


Fig. 5. Runtime complexity comparison with classical ANM.

V. CONCLUSION

In this paper, we introduced a novel approach to solve the ANM problem for DoA estimation. Our method incorporates a quadratic approximation of the original ANM problem, which is efficiently solved through the deep unfolded PGD framework. Extensive simulation results show that our method accurately estimates true DoAs and reduces estimation error. Furthermore, our approach is much less computationally complex than SDP-based ANM, making it more computationally efficient and scalable for practical use. In future work, we will investigate ℓ_1 -norm approximations of the atomic norm to enhance resolution and promote sparsity, thereby overcoming the limitations of the current quadratic approximation method. Additionally, we plan to extend our approach to DoA estimation in reconfigurable intelligent surface (RIS)-aided systems.

ACKNOWLEDGEMENT

This work was supported by the Research Council of Finland through the 6G Flagship program (grant no. 369116), projects DIRECTION (grant no. 354901), DYNAMICS (grant no. 367702), and by the EU's Horizon Europe Research and Innovation Programme through project EXACT-6G (grant no. 101120297).

REFERENCES

- [1] J.-D. Lin, W.-H. Fang, Y.-Y. Wang, and J.-T. Chen, "Fsf music for joint doa and frequency estimation and its performance analysis," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4529–4542, 2006.
- [2] F.-M. Han and X.-D. Zhang, "An esprit-like algorithm for coherent doa estimation," *IEEE Antennas and Wireless Propagation Letters*, vol. 4, pp. 443–446, 2005.
- [3] J. Fang, F. Wang, Y. Shen, H. Li, and R. S. Blum, "Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 18, pp. 4649–4662, 2016.
- [4] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.
- [5] J. Shi, Q. Zhang, W. Tan, L. Mao, L. Huang, and W. Shi, "Underdetermined doa estimation for wideband signals via focused atomic norm minimization," *Entropy*, vol. 22, no. 3, 2020. [Online]. Available: <https://www.mdpi.com/1099-4300/22/3/359>
- [6] W.-G. Tang, H. Jiang, and S.-X. Pang, "Grid-free dod and doa estimation for mimo radar via duality-based 2d atomic norm minimization," *IEEE Access*, vol. 7, pp. 60 827–60 836, 2019.
- [7] S. Gao, H. Ma, H. Liu, and Y. Yang, "Dod and doa estimation from incomplete data based on parafac and atomic norm minimization method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–14, 2023.
- [8] Z. Li and T. Wang, "Admm-based low-complexity off-grid space-time adaptive processing methods," *IEEE Access*, vol. 8, pp. 206 646–206 658, 2020.
- [9] H. Wang, X. Wang, X. Lan, and T. Su, "Cascaded deep neural network for off-grid direction-of-arrival estimation," *IEICE Transactions on Communications*, vol. E107-B, no. 10, pp. 633–644, 2024.
- [10] J. Cong, X. Wang, M. Huang, and L. Wan, "Robust doa estimation method for mimo radar via deep neural networks," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7498–7507, 2021.
- [11] H. Xiang, B. Chen, T. Yang, and D. Liu, "Improved de-multipath neural network models with self-paced feature-to-feature learning for doa estimation in multipath environment," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5068–5078, 2020.
- [12] P. Chen, Z. Chen, L. Liu, Y. Chen, and X. Wang, "Sdoa-net: An efficient deep learning-based doa estimation network for imperfect array," *IEEE Transactions on Instrumentation and Measurement*, 2024.
- [13] J. R. Hershey, J. L. Roux, and F. Weninger, "Deep unfolding: Model-based inspiration of novel deep architectures," *arXiv preprint arXiv:1409.2574*, 2014.
- [14] Z. Mao, S. Liu, Y. D. Zhang, L. Han, and Y. Huang, "Joint doa-range estimation using space-frequency virtual difference coarray," *IEEE Transactions on Signal Processing*, vol. 70, pp. 2576–2592, 2022.
- [15] H. Zhu, W. Feng, C. Feng, T. Ma, and B. Zou, "Deep unfolded gridless doa estimation networks based on atomic norm minimization," *Remote Sensing*, vol. 15, no. 1, p. 13, 2022.
- [16] J. Wang, Y. Wang, and L. Du, "A sparse array doa estimation approach via atomic norm minimization," in *2024 IEEE 7th International Conference on Computer and Communication Engineering Technology (CCET)*. IEEE, 2024, pp. 134–138.
- [17] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 64, no. 5, pp. 1257–1269, 2015.
- [18] M. Ma, T. Fang, N. Shlezinger, A. Swindlehurst, M. Juntti, and N. Nguyen, "Model-based machine learning for max-min fairness beamforming design in JCAS systems," IEEE, 2025.
- [19] O. Lavi and N. Shlezinger, "Learn to rapidly and robustly optimize hybrid precoding," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 5814–5830, 2023.
- [20] N. T. Nguyen, L. V. Nguyen, N. Shlezinger, Y. C. Eldar, A. L. Swindlehurst, and M. Juntti, "Joint communications and sensing hybrid beamforming design via deep unfolding," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 5, pp. 901–916, 2024.
- [21] B. N. Bhaskar and B. Recht, "Atomic norm denoising with applications to line spectral estimation," in *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2011, pp. 261–268.
- [22] W. Zhao, X. Li, Y. Gao, and C. Hao, "Fast gridless direction-of-arrival estimation for wideband linear frequency modulated signals based on fractional fourier transform," *Signal, Image and Video Processing*, vol. 19, no. 1, pp. 1–14, 2025.