Proceedings of CONF-SEML 2026 Symposium: Computational Analysis and Modeling in Complex Intelligent Systems
DOI: 10.54254/2755-2721/2026.GU33273

Model-Driven Deep Learning DOA Estimation Methods Based on Sparse Parameter Estimation: A Review

Yunhao Zhang

School of Information and Electronic Engineering, Zhejiang University, Hangzhou, China
3352033736@qq.com

Abstract. In array signal processing, Direction of Arrival (DOA) estimation is employed to determine the directions of incoming signals. Under such conditions, high signal-to-noise ratios (SNR), an adequate number of snapshots, and optimally configured arrays enable classic methods such as Multiple Signal Classification (MUSIC) and Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) to achieve high accuracy. Yet, their accuracy suffers under low SNR, sparse snapshots, coherent multipath, or imperfect arrays. To address these limitations, sparse reconstruction formulates DOA estimation as a sparse recovery problem, but still encounters challenges such as high computational cost, sensitivity to hyperparameters, and grid mismatch. Recent advances in model-driven deep learning have boosted data adaptability by transforming iterative optimization into trainable networks, applied to enhanced MUSIC, sparse iterative networks, sparse Bayesian learning, atomic norm methods, 1-bit quantization, and non-ideal arrays. Thus, this paper investigates model-driven deep learning DOA estimation methods based on sparse parameter estimation, outlining sparse modeling paradigms and integrating algorithmic frameworks from process flow, sparse priors, model-learning fusion, to continuous angle estimation. Then, trade-offs among accuracy, complexity, and generalization are examined under low SNR, coherent multipath, and non-ideal arrays. The results show that these methods significantly improve the accuracy and robustness of multi-source DOA estimation, reduce the computational cost, and enable efficient feedforward inference in complex conditions.

Keywords: DOA estimation, Sparse parameter estimation, Model-driven deep learning, Sparse Bayesian learning

1. Introduction

In array signal processing, the estimation of Direction of Arrival (DOA) remains a major challenge. Although traditional subspace methods such as Multiple Signal Classification (MUSIC) can achieve high-resolution estimates under ideal conditions, their performance critically depends on stringent requirements including high signal-to-noise ratios, sufficient shot counts, signal non-coherence, and ideal array configurations, often resulting in insufficient robustness in complex real-world scenarios. Sparse representation theory casts DOA estimation as a sparse reconstruction problem, enhancing performance in underdetermined and noisy cases, while posing challenges like computational cost, hyperparameter sensitivity, and grid mismatch. Thus, deep learning, particularly model-driven deep

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.GU33273

learning, addresses these issues by turning classical optimization algorithms into trainable networks, enhancing data adaptability while preserving physical model interpretability, and offering a novel way to resolve these trade-offs. This paper examines model-driven deep learning methods for DOA estimation based on sparse parameter estimation, highlighting frameworks, algorithm improvements, sparse priors, model-learning fusion, and continuous angle estimation. It also evaluates performance trade-offs, computational complexity, and generalization capabilities across various methods under challenging conditions including low SNR, limited shot counts, coherent multipath effects, and non-ideal array configurations. Through literature review and comparative analysis, this study seeks to provide a clear technical roadmap and practical insights for optimizing algorithm design and system deployment in complex environments.

2. Sparse DOA parameter estimation and model-driven learning foundation

2.1. Sparse DOA parameter estimation problem

The sparse representation-based DOA method discretizes the continuous angle domain into G grid points and constructs an overcomplete dictionary matrix $A(\theta)$. The array received data can then be represented as:

$$y = A(\theta)x + n \tag{1}$$

where x denotes the sparse coefficient vector. Thus, angle estimation can be cast as a sparse vector reconstruction problem x[1-3]. This model offers theoretical advantages under underdetermined conditions and limited snapshots, and can be readily combined with multi-task statistical priors such as group sparsity. However, sparse DOA methods face several critical bottlenecks. Grid mismatch leads to spectral leakage and estimation errors. The solution process relies on iterative optimization (e.g., objective functions with L1 regularization), resulting in high computational complexity and sensitivity to hyperparameters. Additionally, model mismatches such as array coupling and element errors reduce the accuracy of steering vectors $A(\theta)$ and compromise dictionary consistency [1,4]. These factors collectively limit traditional sparse DOA performance in practice. To overcome this, model-driven deep learning transforms iterative processes of classical optimization algorithms (e.g., ISTA and SBL update steps) into trainable networks [2,3]. For instance, gridless continuous-domain optimization based on atomic norms has been implemented as neural networks to improve online convergence while retaining continuous angular domain advantages. By integrating physical models with data-driven learning, this approach yields higher accuracy and efficiency [1,5].

2.2. Basic ideas of model-based learning

The core of model-driven deep learning in DOA estimation lies in transforming iterative algorithms into learnable network layers, enhancing performance and efficiency while preserving physical constraints. For instance, LISTA parameterizes the linear transformation and soft thresholding steps of ISTA to enable the network to approximate multi-iteration optimization results with fewer layers, while PACT-LISTA incorporates array position information and complex Toeplitz structures to improve network interpretability and generalization capabilities [2,3]. Learnable parameters mainly focus on critical algorithmic components, including convergence/sparseness step sizes, thresholds, regularization weights, and prior hyperparameters in sparse Bayesian learning such as variance

DOI: 10.54254/2755-2721/2026.GU33273

update rules [1-3,5]. Moreover, parameters correcting array non-idealities like mutual coupling or element errors, along with modules improving traditional methods (e.g., DA-MUSIC improvements over MUSIC), can also serve as trainable parameters [6-8].

In terms of training methods, model-driven DOA networks often integrate supervised learning, self-supervised or consistency constraints, and structural regularization. In particular, supervised learning utilizes real DOA measurements or angle spectra as labels, self-supervised or consistency constraints are optimized through data reconstruction errors, while structural regularization imposes Toeplitz or positive-definite constraints. For method selection, DA-MUSIC adopts a hybrid method combining traditional processes with learning enhancements to retain network interpretability and improve robustness, whereas end-to-end methods like DeepMUSIC lean more toward data-driven strategies that directly enhance estimation speed and noise resistance via neural networks, forming complementary research pathways [6,7,9]. For evaluation, model-driven DOA approaches should consider metrics such as accuracy, resolution, robustness, computational cost, generalization, and practical usability, rather than relying solely on RMSE. Gridless networks based on atomic norms achieve high precision in continuous angle domains while lowering online solution costs, whereas SBL-based expansions maintain Bayesian robustness with fixed inference complexity [1,4,5]. These results demonstrate that learning key parameters within an interpretable physical framework can simultaneously improve performance, efficiency, and robustness.

3. Model-driven deep learning DOA method based on sparse parameter estimation

3.1. Learning methods at the algorithmic flow level

Traditional MUSIC performance relies on accurate separation of signal and noise subspaces. Under challenging conditions like low signal-to-noise ratios, limited fast-scan data, and signal coherence, their angular spectra often show artifacts including false peaks, broadened main peaks, and spectral shifts [6,7]. To boost robustness, MUSIC-based enhancement methods have been developed. They integrate deep neural networks as "enhancement modules" while preserving the MUSIC physical framework. By altering or replacing key steps (e.g., subspace estimation and spectrum generation), these methods markedly improve algorithm robustness in complex environments.

For example, DA-MUSIC integrates deep learning modules into the MUSIC workflow, using data-driven learning to tclarify signal subspaces or optimize spectral structures and compensate for covariance matrix errors under non-ideal conditions [6,7]. Its "retrogradability" allows the system to revert to standard MUSIC if the enhancement fails, maintaining basic performance. DeepMUSIC directly reconstructs or approximates angular spectra via neural networks for faster inference [9]. However, these approaches mainly perform local optimizations of existing physical processes and still rely on predefined spatial sampling grids, making off-grid mismatch hard to resolve. Embedded learning modules often depend heavily on training data, like specific array setups or element counts, limiting generalization when system parameters or scenarios change [6,7,9]. As a result, these approaches are best suited for enhancing the robustness of traditional algorithms under well-defined system constraints.

3.2. Learning methods at the sparse prior level

Sparse prior layer learning reduces the computational cost and sensitivity to hyperparameters (e.g., thresholds) in traditional sparse recovery algorithms such as LASSO for DOA estimation [2,3]. It converts iterative sparse optimization algorithms (e.g., ISTA) into deep neural networks, allowing

DOI: 10.54254/2755-2721/2026.GU33273

machine learning of key parameters such as iteration steps and thresholds, along with structural priors. And this enables fixed-depth networks to replicate and even surpass the slow, multi-iteration processes of traditional methods, turning iterative optimization into fast, deterministic feedforward computations.

Among such methods, PACT-LISTA is a representative example [2,3]. It explicitly incorporates array physical priors (e.g., complex Toeplitz structure, element positions) into network weights to integrate model architecture with learning. This design improves interpretability and generalization while compensating for model mismatches from non-ideal factors such as array coupling. After training, the network automatically learns sparse patterns suited for low SNR or coherent multipath, and turns high-latency iterations into fast, deterministic inference for real-time DOA applications like radar or communication beam tracking [2,3,10]. However, the performance of these methods remains dependent on training data distribution. Poorly designed objectives for learnable parameters (e.g., loss functions) can prevent convergence to physically reasonable solutions. Besides, networks tailored to specific architectures (e.g., particular array configurations) may also generalize poorly to different array geometries.

3.3. Integration of analysis models and learning methods

Traditional DOA systems suffer from non-ideal factors such as array intercoupling, element errors, channel imbalance, and nonlinear distortions from hardware limits like 1-bit quantization, which can break ideal observation models and sharply degrade performance [2,3,8,11]. To address this, fusion approaches combine analytical modeling with learning using a hybrid white-box/black-box strategy. Physical components, such as array manifolds and signal propagation models, remain in analytical form to constrain the solution space, while neural networks compensate for nonlinearities and distortions like intercoupling matrices and quantization errors [6-8]. For example, structured expansion networks like PACT-LISTA embed array physical structures such as complex Toeplitz characteristics into the network to learn unknown distortion parameters, and model-driven methods target nonlinear hardware constraints like 1-bit quantization [2,3]. By training efficient inference networks to bypass complex nonlinear analysis, these methods enable DOA recovery from strongly nonlinear observations [11].

This approach reduces the degrees of freedom and data volume required by the learning module via physical model constraints, boosting the model's predictability and generalization capabilities under unseen conditions. It also offers a practical way to handle complex distortions such as strong nonlinear quantization that traditional analytical methods cannot model [6,7,11]. The main risk is that the learning module may introduce biases or instability. Parameters are limited to key low-dimensional variables, loss functions enforce physical consistency, and fallback paths ensure the system can revert to classical methods if the learning module fails [6,7]. As such, these methods must balance performance gains with reliability and interpretability.

3.4. Joint estimation of angular parameters with contouring

Grid mismatch remains an inherent limitation of sparse DOA methods. When discrete dictionaries misalign with continuous angles, they can result in energy leakage and estimation bias. Accordingly, model-driven deep learning has progressed along two directions. The first applies grid-based offset correction, adding continuous angle offsets to discrete grids to jointly estimate sparse coefficients and angles. Sparse Bayesian Learning (SBL) is well suited for this approach due to its hierarchical priors. SBL-based deep expansion networks then learn hyperparameter updates to estimate offsets

DOI: 10.54254/2755-2721/2026.GU33273

adaptively [1,5]. The second approach is grid-free, using atomic norm minimization to solve sparse reconstruction directly in the continuous angle domain, eliminating grid errors. These methods also convert iterative optimization into feedforward deep expansion networks, achieving near-optimal solutions with lower computational cost [4].

Both approaches achieve efficient inference through the "expansion+learning" mechanism. For example, offset-correction methods such as SBL expansion networks provide a straightforward design, high computational efficiency, and seamless incorporation of prior knowledge from other models, although their performance is constrained by grid resolution, rendering them approximate solutions [1,5]. In contrast, ungrid-based methods like atomic norm expansion networks provide higher theoretical accuracy and resolution, but need more complex architectures, careful training, and better numerical stability [1,4,5]. Thus, practical implementation requires choosing solutions based on estimation accuracy, computational efficiency, and system complexity.

4. Model-based DOA method performance in complex scenarios

4.1. Performance performance in complex scenarios

In low SNR and limited fast-sampling scenarios, traditional subspace methods experience degraded performance due to covariance matrix estimation errors. In order to boost robustness, DA-MUSIC incorporates learning modules during subspace estimation and spectrum construction [6,7]. At the same time, sparse methods are sensitive to regularization parameters, but expansion networks such as LISTA enable adaptive tuning through learnable iteration steps and thresholds, ensuring stable performance. Similarly, SBL-based expansion networks adopt Bayesian frameworks to dynamically adjust sparsity levels, thereby achieving superior results under sparse sampling and coherent source conditions [1-3,5]. Thus, structured expansion networks like PACT-LISTA encode array physical structures as network constraints, focusing on compensating unknown mismatch parameters while balancing performance and computational cost. Gridless atomic norm expansion networks remove discrete errors, reaching theoretical optimality despite higher implementation complexity [2-4]. Furthermore, methods designed for non-ideal arrays and 1-bit quantization hardware extend these techniques' practical applicability, enabling system designers choose or combine solutions based on precision, speed, and robustness [8,11].

4.2. Current challenges and future research directions

Existing model-driven deep learning DOA methods face several key challenges. For example, the lack of unified benchmark datasets for evaluation prevents objective comparison of generalization capabilities across different approaches. Moreover, generalization performance is affected by array geometry and operational frequency bands, which can be adapted through conditional networks, learning array-invariant features, or meta-learning techniques. In this context, structured learning methods like PACT-LISTA improve versatility by embedding physical priors into neural networks [1,2]. Meanwhile, high-precision methods like gridless atomic norm expansion networks achieve theoretical accuracy but face limitations in training stability, numerical robustness, and adaptability to complex scenarios [4]. Consequently, practical applications require end-to-end joint calibration and estimation systems for non-ideal arrays and hardware constraints such as 1-bit quantization, along with lightweight networks for edge devices [8,11]. Besides, engineering reliability demands thorough analysis of network convergence and stability, validation against theoretical performance

DOI: 10.54254/2755-2721/2026.GU33273

bounds such as the Krammer-Roberts bound, and the inclusion of confidence assessment and fault detection mechanisms.

5. Conclusion

This paper reviews model-driven deep learning DOA methods based on sparse parameter estimation. When traditional sparse DOA methods face grid mismatch, high iterative complexity, or model mismatch, model-driven deep learning converts algorithm iterations into trainable networks and replaces manual tuning with adaptive learning, improving efficiency and robustness while keeping physical interpretability. Four main technical pathways have emerged in this field. Process-level enhancement such as DA-MUSIC embeds learning modules into conventional algorithms, whereas sparse prior layer expansion like PACT-LISTA turns iterative solvers into networks with physical structure priors. Simultaneously, integration of analytical models with learning compensates for array imperfections and hardware distortions, and continuous-angle joint estimation addresses grid mismatch through off-grid correction and grid-free networks for high-precision estimation. Typical methods convert high-latency iterations into feedforward inference and use physics-guided learning for non-ideal environments. However, remaining challenges include incomplete evaluation, limited generalization, stability issues in high-precision methods, and the need for lightweight closed-loop designs. Future efforts should enhance cross-environment robustness, adapt to hardware constraints, and deepen theoretical interpretability.

References

[1] Gong, Z., et al. (2023). Deep unfolding sparse Bayesian learning network for off-grid DOA estimation. Remote Sensing, 15(22), 5320.
[2] Guo, Y., et al. (2022). Position-enabled complex Toeplitz LISTA for DOA estimation. Signal Processing, 194, 108422.
[3] Guo, Y., Jin, J., Wang, Q., Chen, H., & Liu, W. (2022). Position-enabled complex Toeplitz LISTA for DOA estimation with unknow mutual coupling. Signal Processing, 194, 108422.
[4] Zhu, H., et al. (2022). Deep unfolded gridless DOA estimation networks based on atomic norm minimization. Remote Sensing, 15(1), 13.
[5] Li, N., Zhang, X., Lv, F., Zong, B., & Feng, W. (2024). A Bayesian Deep Unfolded Network for the Off-Grid Direction-of-Arrival Estimation via a Minimum Hole Array. Electronics, 13(11), 2139.
[6] Merkofer, J. P., Revach, G., Shlezinger, N., Routtenberg, T., & Van Sloun, R. J. (2023). DA-MUSIC: Data-driven DoA estimation via deep augmented MUSIC algorithm. IEEE Transactions on Vehicular Technology, 73(2), 2771-2785.
[7] Merkofer, J. P., Revach, G., Shlezinger, N., & van Sloun, R. J. (2022, May). Deep augmented music algorithm for data-driven DoA estimation. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 3598-3602.
[8] Wu, L., Liu, Z., & Liao, J. (2022, July). DOA estimation using an unfolded deep network in the presence of array imperfections. In 2022 7th International conference on signal and image processing (ICSIP), 182-187.
[9] Elbir, A. M. (2020). DeepMUSIC: Multiple signal classification via deep learning. IEEE Sensors Letters, 4(4), 1-4.
[10] Niu, X., Su, X., He, L., & Chen, G. (2025). Direction of arrival (DOA) estimation using a deep unfolded learned iterative shrinkage thresholding algorithm (LISTA) network in a non-uniform metasurface. Remote Sensing, 17(7), 1253.
[11] Yeganegi, F., Eamaz, A., Esmaeilbeig, T., & Soltanalian, M. (2024, July). Deep learning-enabled one-bit DoA estimation. In 2024 IEEE 13rd Sensor Array and Multichannel Signal Processing Workshop (SAM), 1-5.