# Optimization of High-resolution and Ambiguity-free Sparse Planar Array Geometry for Automotive MIMO Radar

Mingsai Huan, *Student Member, IEEE*, Junli Liang, *Senior Member, IEEE*, Yugang Ma, *Senior Member, IEEE*,
Wei Liu, *Senior Member, IEEE*, Yifan Wu, and Yonghong Zeng, *Fellow, IEEE*

***Abstract*—The next-generation 4D imaging automotive radar is characterized by high angular resolution, unambiguous detection, low latency, low cost, and small size. This study provides an enhanced analysis of the angular ambiguity function (AAF) for planar MIMO arrays, and pioneers a method for a more accurate evaluation of angular resolution using the main lobe width (MLW). Then the 2D expanded beam pattern (EBP) is introduced to assess the field-of-view (FOV), region of interest (ROI), sidelobe level (SLL), and normalized resolution intuitively and precisely. After constructing the sophisticated 2D element spacing and aperture constraints for planar MIMO arrays, the optimization of array geometry is creatively formulated as a novel Domino sparse optimization problem aiming to minimize the MLW while sufficiently suppressing the SLL, which is inspired by the sequential fall of dominoes. This non-convex and non-smooth constrained problem is efficiently solved by a hybrid optimization framework, which integrates the alternating direction multiplier method (ADMM), aggregate function, modified real genetic algorithm (MGA), and non-uniform fast Fourier transform (NUFFT). Numerical simulations demonstrate that angular resolution varies with array geometry, even under the same aperture size. The proposed arrays outperform others with equal aperture size, exhibiting narrower MLW and lower Cramér-Rao bound (CRB), thereby enhancing angular resolution with fewer antennas and without preprocessing in standard single-snapshot 2D DOA estimation methods.**

***Index Terms*—Colocated MIMO radar, Sparse planar array, Angular ambiguity function, Angular resolution, Non-convex optimization.**

## I. INTRODUCTION

OVER the past decades, automotive safety has led to the adoption of radar sensors for detecting various entities [1], and multiple-input multiple-output (MIMO) radar has been a focus due to its high angular resolution with limited antennas [2], used in current-generation advanced driver-assistance systems (ADAS) and explored for next-generation high-resolution imaging in autonomous driving. The next-generation MIMO radar aims for finer imaging capability for better target recognition, and therefore demands high angular resolution, no ambiguity, low latency, low hardware cost, and small size, as defined by the most competitive sensor products [3].

The angular resolution capacity of an antenna array is essential for automotive radar [4]. Various super-resolution direction of arrival (DOA) estimation methods have been developed, such as multiple signal classification (MUSIC) [5], compressed sensing (CS) [6] and the iterative adaptive approach (IAA) [7], but the resolution is still limited by the number of antennas in uniform arrays. Non-uniform arrays, which offer cost reduction and larger aperture, have thus gained significant attention [8]. These arrays are divided into two categories: regular sparse arrays (RSAs), with elements on a regular grid, and irregular sparse arrays (ISAs), where elements are arbitrarily distributed.

RSAs, also known as thinned arrays, select elements from a uniform array [9]. The difference coarray concept is essential for RSA as it appears naturally in the cross correlation between two elements in the covariance matrix [4]. This allows the thinned array to be re-formed as a virtual coarray with larger aperture size [10], thus achieving better DOA estimation performance. However, most of RSAs cannot deal with coherent sources [11], and multiple snapshots are required for accurate covariance matrix estimation [3]. This is challenging for 4D imaging radar in dynamic automotive scenarios, where only a few or even a single snapshot might be available after range-Doppler two-dimensional fast Fourier transform (2D-FFT) [3]. A recent method used structured matrix completion to fill the missing elements of covariance matrix with RSA using only a single snapshot [12], but this is computationally intensive.

Unlike RSAs, ISAs have arbitrarily distributed elements on the antenna aperture, increasing the degrees of freedom (DoFs) for the antenna geometry. A key problem with sparse arrays is angular ambiguity, which arises when the radar system cannot distinguish between true targets and ghost targets generated by similar steering vectors. Various approaches have been proposed to design an ambiguity-free sparse array. [13] provides a simple ambiguity measure using MUSIC, while [14] finds the optimal circular array by measuring the average

---

This work was supported by the China Scholarship Council (CSC) under Grant Liu Jin Xuan [2022] No. 87, the Key Research and Development Program of Shaanxi under Grant 2021SF-166, Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University, Aeronautical Science Foundation of China under Grant 201920053001, and National Research Foundation, Singapore and Infocomm Media Development Authority under its Future Communications Research & Development Programme (Grant award number: FCP-NUS-RG-2022-018, FCP-ASTAR-TG-2022-003). (Corresponding author: Junli Liang.)

Mingsai Huan, Junli Liang and Yifan Wu are with School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China. (e-mail: huanmingsai@163.com; liangjunli@nwpu.edu.cn; wuyf@mail.nwpu.edu.cn). Yugang Ma and Yonghong Zeng are with A*STAR, Singapore, Institute for Infocomm Research (I2R), Singapore 138632 (e-mail: mayg@i2r.a-star.edu.sg; yhzeng@i2r.a-star.edu.sg). Wei Liu is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, E1 4NS, UK (e-mail: w.liu@qmul.ac.uk).

SLL and RMSE of the signal-subspace MUSIC spectrum. Another method imposes restrictions, such as the maximal inter-sensor spacing (typically half the wavelength), on the array geometry to ensure first-order ambiguity-free arrays [15], [16]. [17] reduces the Euclidean distance between the steering vectors. Selecting the objective function for geometry optimization is an important issue in ISA design. Advanced designs of ISA typically focus on two main objectives for 2D DOA estimation: enhancing angular accuracy using the Cramér-Rao bound (CRB) [13] as a criterion, and avoiding ambiguities by employing the ambiguity function [18] as a guiding principle. This paper refers to the latter as the angular ambiguity function (AAF) to distinguish it from the radar range-Doppler ambiguity function.

Cramér-Rao bounds (CRBs) for sparse arrays have been extensively studied [19]. An isotropic array is proposed in [20], where the CRB on the DOA of a single source is uniform for all angles. The CRB, used to evaluate DOA estimation accuracy, is only a tight bound for maximum likelihood estimates of DOA at high signal-to-noise ratios (SNRs), but not at low SNRs [21]. A two-target CRB is used in [22] to reduce the sidelobe level (SLL) at the cost of a wider beamwidth. Other bounds for random parameter estimation, such as the Bayesian CRB (BCRB) [21], the Weiss–Weinstein bound (WWB) [23], the Ziv–Zakai bound (ZZB) [24], [25], and the Bobrovsky-Zakai Bound (BZB) [26], are used to evaluate ambiguity-free sparse arrays. These bounds are tighter and more realistic than CRB, especially at low SNRs, but have complex mathematical expressions. The angular ambiguity function (AAF), which measures the similarity of any two steering vectors of an array, is a more suitable metric for ambiguity assessment due to its interpretability and computational simplicity. Consequently, both the SLL of AAF and the CRB are considered in optimizing ISA geometry for high-accuracy and ambiguity-free DOA estimation [27]–[30].

Angular resolution is typically evaluated using the Rayleigh criterion, which determines resolution based on aperture size [31]. Although the Rayleigh criterion offers a convenient approach, it becomes less accurate for non-uniform arrays where the beamwidth varies with array geometry [32]. The mainlobe width (MLW) of AAF provides a more precise method for evaluating resolution among arrays with equal aperture sizes [33], [34]. The relationship between angular accuracy and resolution has been explored, with some derivations from CRB to beamwidth in [20], [27]. The half-power beamwidth can be approximated by the CRB performance using a second-order Taylor series approximation, but this leads to estimation errors [32]. [35] suggests that the angular resolution limit (ARL) is inversely proportional to the variance of array element positions, which is also drawn from the two-source CRB approximation. High-resolution ISA geometry is designed based on array manifold properties in [36], but this causes a high SLL. Using the manifold, CRB, and other Bayesian bounds as performance metrics can be problematic in practice, as array calibration errors tend to dominate DOA estimation errors [37], [38]. The real performance of arrays, considering undesirable factors, can be more effectively analyzed through AAF measurements of a prototype [39], and further improved by a calibration [40].

Conventional pattern synthesis methods usually minimize the SLL and control the MLW indirectly by selecting a sidelobe region, and heuristic algorithms are often utilized to optimize the ISA geometry for a low SLL given their global convergence, like genetic algorithm (GA) [27], [41], simulated annealing (SA) [42] and particle swarm optimization (PSO) [43]. GA is exploited to maximize the area of the ambiguity-free region in [39], which defines the 2-D-angular field-of-view (FOV) of a MIMO radar, but the authors simply assume that the angular resolution is determined by the aperture size without considering the MLW. While low SLL arrays have been extensively studied, a narrow MLW is also crucial as it provides high angular resolution and accuracy for many DOA estimation methods [37]. To achieve a higher resolution, MLW is also incorporated into the objective function in recent studies [25], [44], and new search techniques have been explored to solve the multi-objective optimization problem with element position constraints [25].

However, multi-objective optimization struggle with MLW minimization as reaching the MLW-SLL pareto front boundary with few iterations is challenging [25], [33]. No solution has been reported to directly reduce the MLW of the angular ambiguity function (AAF) by planar array geometry optimization. Recently, sparse optimization algorithms achieving minimum beamwidth have been proposed for focused beam pattern synthesis [45], [46], which could allow a narrow-MLW planar array design. Nevertheless, the ISA geometry optimization problem differs from beam pattern synthesis, as optimizing the element positions is a strongly nonlinear and non-convex problem with variables in exponential terms.

This paper focuses on the optimization of planar MIMO ISA geometry, aiming to improve angular resolution within constraints of no angular ambiguity, maximum aperture size, and antenna spacing limits. The key contributions are:

1) Based on a detailed analysis of the AAF for planar arrays in various coordinate systems, we pioneer a method for estimating the angular resolution of arbitrary planar arrays using MLW evaluation, more accurate than the well-known Rayleigh criterion. This method builds on the established understanding that MLW is influenced by the array geometry, even when the aperture size remains constant [32], [33], [47]. Due to the high dimensionality of AAF, the 2D expanded beam pattern (EBP) is introduced to assess the FOV, ROI, SLL, and 2D normalized resolution more intuitively and precisely.

2) To tackle the practical challenges of limited aperture size, antenna overlap, and mutual coupling in planar MIMO arrays, advanced 2D MIMO element position constraints are constructed, which are notably more sophisticated than the linear array case reported in our previous work [33]. Aggregate functions are skillfully employed to deal with multiple non-smooth functions in the constraints.

3) It is a difficult task to directly minimize the MLW of EBP. In this study, the planar MIMO array geometry optimization problem is modeled as an innovative 2D Domino sparse optimization problem for the first time, aiming to minimize the MLW while sufficiently suppressing the SLL. This model, inspired by the sequential fall of dominoes, prioritizes the

sparsification of outer variables to ensure a sequential collapse towards the main lobe peak. This non-convex problem with many nonlinear constraints is iteratively solved by the ADMM, and a hybrid algorithm with an acceleration using NUFFT and an initialization using MGA is proposed to improve efficiency.

The rest of this paper is organized as follows. In Section II, the problem is formulated after analysis of the angular ambiguity function and antenna position constraints. The MIMO array geometry optimization problem and associated algorithms are developed in Section III. In Section IV, numerical examples are provided to evaluate the performance of the optimized arrays. Finally, conclusions are drawn in Section V.

*Notation*: Vectors and matrices are denoted by boldface lowercase and uppercase letters, respectively. $\|\cdot\|$ denotes the Frobenius norm, while $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$ and $(\cdot)^{-1}$ are the transpose, Hermitian transpose and matrix inverse operations, respectively. $\mathbf{0}_{m\times n}$ and $\boldsymbol{I}_n$ represent the $m\times n$ zero matrix and $n\times n$ identity matrix, respectively. $\Re\{\}$ and $\Im\{\}$ represent the real and imaginary parts with $j=\sqrt{-1}$. $|\cdot|$ and $\angle\{\cdot\}$ are the magnitude and phase of a complex-valued scalar, respectively. $\otimes$ is the Kronecker product, and $\succeq$ a type of generalized inequalities, with $\boldsymbol{x}\succeq 0$ indicating that all its elements are greater than 0.

## II. SYSTEM MODEL, ANALYSIS, AND PROBLEM FORMULATION

In this section, the concepts of angular ambiguity function, forward-looking angular resolution, expanded beam pattern, element position constraints and aggregate function are introduced, followed by the problem formulation.

### *A. Angular Ambiguity Function*

Consider a planar array for a collocated MIMO radar system with $M$ transmitting antennas and $N$ receiving antennas. In the 3D coordinate system, the planar array is assumed to be placed on the xoz plane if it is used for forward-looking radar. Under the narrow-band far-field condition, the incident signal can be considered as a plane wave. As shown in Fig. 1, $\varphi$ and $\vartheta$ represent the forward-looking azimuth (FLA) and forward-looking elevation (FLE), respectively [39]. $\varphi\in[-\pi/2,\pi/2]$ and $\vartheta\in[-\pi/2,\pi/2]$ are the corresponding angular ranges in the half space in front of the array. Specifically, $\varphi>0$ means that the target is on the right front of the array (positive x-axis direction), and $\vartheta>0$ indicates that it is above the array (positive z-axis direction).

Let

$$\boldsymbol{x}_t=\begin{bmatrix}x_{t,1}\\ \vdots\\ x_{t,M}\end{bmatrix},\boldsymbol{z}_t=\begin{bmatrix}z_{t,1}\\ \vdots\\ z_{t,M}\end{bmatrix},\boldsymbol{x}_r=\begin{bmatrix}x_{r,1}\\ \vdots\\ x_{r,N}\end{bmatrix},\boldsymbol{z}_r=\begin{bmatrix}z_{r,1}\\ \vdots\\ z_{r,N}\end{bmatrix} \tag{1}$$

describe the x- and z-coordinates of the $M$ transmitters and $N$ receivers respectively in MIMO radar. Then, the transmitting and receiving steering vectors associated with FLA $\varphi$ and FLE $\vartheta$ can be written as

$$\boldsymbol{a}_t(\varphi,\vartheta)=\bar{g}(\varphi,\vartheta)\exp(-\mathrm{j}\frac{2\pi}{\lambda}(\boldsymbol{x}_t\cos\vartheta\sin\varphi+\boldsymbol{z}_t\sin\vartheta)),$$

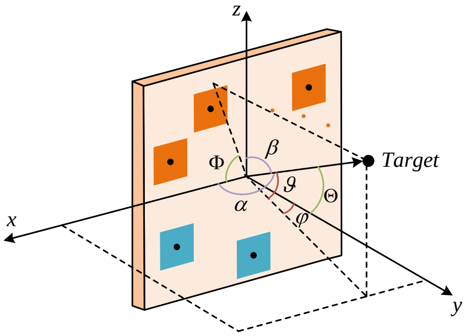


Fig. 1. The system model of a forward-looking MIMO radar and the representation of a target's 2D DOA in different coordinate systems. The forward-looking azimuth and elevation angles are denoted by $\varphi$ and $\vartheta$, respectively; $\alpha$ and $\beta$ are called electric angles [4]; $\Phi$ and $\Theta$ are azimuth and elevation in a spherical coordinate system.

$$\boldsymbol{a}_r(\varphi,\vartheta)=\bar{g}(\varphi,\vartheta)\exp(-\mathrm{j}\frac{2\pi}{\lambda}(\boldsymbol{x}_r\cos\vartheta\sin\varphi+\boldsymbol{z}_r\sin\vartheta)). \tag{2}$$

where $\bar{g}(\varphi,\vartheta)$ is an average element radiation pattern, which includes the information on the mutual coupling environment, and it can be exploited to approximate each element pattern in the array with high accuracy [33], [48]. With multiple virtual channels created in MIMO radar by transmitting orthogonal waveforms [2], the virtual array steering vector can be expressed as the Kronecker product of the transmitting and receiving steering vectors, i.e.,

$$\boldsymbol{a}(\varphi,\vartheta)=\boldsymbol{a}_t(\varphi,\vartheta)\otimes\boldsymbol{a}_r(\varphi,\vartheta). \tag{3}$$

Here, the angular ambiguity function (AAF) [18] is used to describe the array's ability to discriminate between the signals of targets. Specifically, it measures the cosine similarity of the steering vectors corresponding to the two directions $\boldsymbol{a}(\varphi_i,\vartheta_i)$ and $\boldsymbol{a}(\varphi_j,\vartheta_j)$:

$$\begin{aligned}AF\left(\varphi_i,\vartheta_i,\varphi_j,\vartheta_j\right)&=\frac{\boldsymbol{a}\left(\varphi_j,\vartheta_j\right)^{\mathsf{H}}\boldsymbol{a}\left(\varphi_i,\vartheta_i\right)}{\left\|\boldsymbol{a}\left(\varphi_i,\vartheta_i\right)\right\|\left\|\boldsymbol{a}\left(\varphi_j,\vartheta_j\right)\right\|}\\ &=\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m}+x_{r,n})(\cos\vartheta_i\sin\varphi_i\\ &\quad-\cos\vartheta_j\sin\varphi_j)+(z_{t,m}+z_{r,n})(\sin\vartheta_i-\sin\vartheta_j)]\}\end{aligned} \tag{4}$$

The derived AAF eliminates the effect of element radiation patterns. To simplify expressions, define the following spatial frequency variables

$$\begin{aligned}u_i&=\cos\alpha_i=\cos\vartheta_i\sin\varphi_i=\sin\Theta_i\cos\Phi_i,\\ u_j&=\cos\alpha_j=\cos\vartheta_j\sin\varphi_j=\sin\Theta_j\cos\Phi_j,\\ v_i&=\cos\beta_i=\sin\vartheta_i=\sin\Theta_i\sin\Phi_i,\\ v_j&=\cos\beta_j=\sin\vartheta_j=\sin\Theta_j\sin\Phi_j,\end{aligned} \tag{5}$$

where $\alpha\in[0,\pi]$ and $\beta\in[0,\pi]$ are called electric angles, $\Phi\in[0,2\pi]$ and $\Theta\in[0,\pi/2]$ are azimuth and elevation in a spherical coordinate system [4]. Thus, we obtain two other representations for the AAF:

$$AF\left(\alpha_i,\beta_i,\alpha_j,\beta_j\right)$$

$$= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m}+x_{r,n})(\cos\alpha_i-\cos\alpha_j) + (z_{t,m}+z_{r,n})(\cos\beta_i-\cos\beta_j)]\}, \quad (6)$$

$$AF\left(u_i,v_i,u_j,v_j\right) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m}+x_{r,n})(u_i-u_j) + (z_{t,m}+z_{r,n})(v_i-v_j)]\}. \quad (7)$$

The AAFs defined in (4), (6), and (7) depend on four variables. If the subscript $j$ is used to indicate the real DOAs and the subscript $i$ the estimated ones [39], then the AAF patterns with two variables for estimated DOAs can be obtained by fixing the real DOAs, namely $AF_{\varphi_j,\vartheta_j}(\varphi_i,\vartheta_i)$, $AF_{\alpha_j,\beta_j}(\alpha_i,\beta_i)$, and $AF_{u_j,v_j}(u_i,v_i)$.

It is worth noting that the maximum absolute value of (4), (6) and (7) when $i \neq j$ can be regarded as the mutual coherence of sensing matrix in compressed sensing [6], [49], defined by

$$coh \triangleq \max_{i\neq j}\frac{\left|\boldsymbol{a}(\varphi_j,\vartheta_j)^{\mathsf{H}}\boldsymbol{a}(\varphi_i,\vartheta_i)\right|}{\|\boldsymbol{a}(\varphi_i,\vartheta_i)\|\,\|\boldsymbol{a}(\varphi_j,\vartheta_j)\|} \quad (8)$$

$coh$ takes values from 0 to 1, and a low coherence value is desired to ensure optimal sparse recovery properties for accurate DOA estimation.

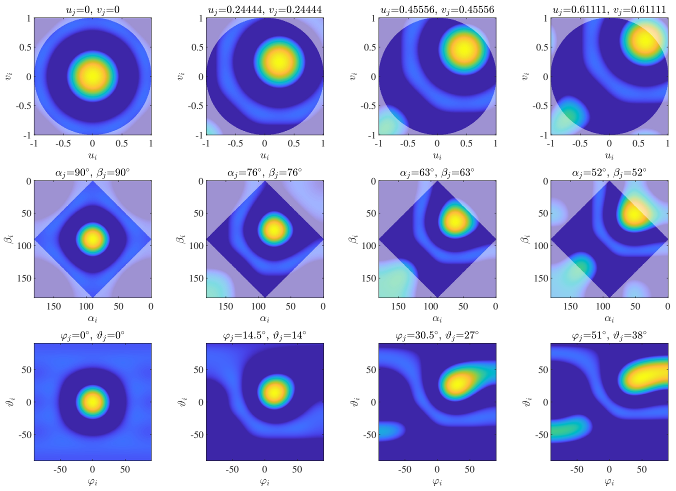


Fig. 2. The color maps of AAF amplitude for 1T8R UCA: from top to bottom are $|AF_{u_j,v_j}(u_i,v_i)|$ ,$|AF_{\alpha_j,\beta_j}(\alpha_i,\beta_i)|$, and $|AF_{\varphi_j,\vartheta_j}(\varphi_i,\vartheta_i)|$, where color distribution represents values ranging from $-10$ dB to 0 dB.

To investigate the properties of AAF (as well as $coh$) expressed in different coordinate systems, the color maps of AAF amplitude for 1 transmitter and 8 receivers (1T8R) uniform circular array (UCA) with half wavelength spacing are given in Fig. 2. Four real DOAs are assumed, and each column of the figures corresponds to one of them, which has different representations in different coordinate systems. The main lobe region is defined as the region bounded by the first null contour around the real DOA, which is indicated by the yellow areas on the images, while the remaining areas are side lobe region. The opaque areas in Fig. 2 represent the physically feasible DOA domain, which can be derived from (5). Some properties can be observed as described in the following.

1) The image of $|AF_{u_j,v_j}(u_i,v_i)|$ is symmetrical about the point $(u_j,v_j)$, that is,

$$|AF_{u_j,v_j}(u_i,v_i)| = |AF_{u_j,v_j}(2u_j-u_i,2v_j-v_i)|. \quad (9)$$

As the real DOA changes, the pattern simply translates with it, and new side lobe region appears, while the main lobe width does not change.

2) The images of $|AF_{\alpha_j,\beta_j}(\alpha_i,\beta_i)|$ and $|AF_{\varphi_j,\vartheta_j}(\varphi_i,\vartheta_i)|$ are no longer symmetrical when the real DOA deviates from the boresight direction. As the deviation angle increases, the MLWs in $\alpha_i$- and $\beta_i$- direction increase, which implies reduction in angular resolution. For the forward-looking coordinate system $(\varphi_i,\vartheta_i)$, the FLA resolution is degraded more than the FLE resolution.

### B. Forward-looking Angular Resolution Estimation

The well-known Rayleigh angular resolution for bore-sight direction used in [39] is given by

$$\Delta\varphi_{\text{rayl}} = 1.22\frac{\lambda}{d_x}, \qquad \Delta\vartheta_{\text{rayl}} = 1.22\frac{\lambda}{d_z}, \quad (10)$$

where $d_x$ and $d_z$ represent the $x$-aperture size and $z$-aperture size (of virtual array for MIMO radar). (10) is derived from half of the first null beamwidth (HFNBW) of Fraunhofer diffraction pattern [50], which is calculated by continuous Fourier transform (FT) of the aperture function. However, radar antenna arrays typically involve discrete spatial sampling, and the AAF formulas presented in (4), (6), and (7) are essentially discrete analogs of the Fraunhofer diffraction integral, necessitating the use of the discrete Fourier transform (DFT) or, for non-uniform arrays, the non-uniform discrete Fourier transform (NUDFT). For uniform arrays, DFT can approximate FT well, thus making (10) a reasonable estimator of resolution. However, for non-uniform arrays, the significant differences between NUDFT and FT results lead to inaccuracies when (10) is applied. This discrepancy is critical as it highlights the limitations of traditional optical resolution criteria when applied to radar systems, especially in the case of non-uniform arrays.

Considering the inaccuracies involved in directly applying (10) to non-uniform arrays, the HFNBW of AAF is used to evaluate the resolution of arbitrary sparse planar arrays in this study, which determines the minimum interval between two targets required for estimating their powers without distortion. Therefore, the 2D MLWs of $|AF_{u_j,v_j}(u_i,v_i)|$, denoted by $\Delta u$ and $\Delta v$, are indicated as the HFNBWs of $|AF_{u_j,v_j}(u_i,v_i)|$ along $u_i$-direction and $v_i$-direction, and they describe the distance from the peak of main lobe to the corresponding first null point. As shown in Fig. 2, since $\Delta u$ and $\Delta v$ are only related to the array geometry and do not vary with the real DOA, they can be used as normalized resolution (NR).

When the real DOA is bore-sight direction $\varphi_j = \vartheta_j = 0$, the forward-looking angular resolution in $\varphi_i$-direction and $\vartheta_i$-direction is denoted by $\Delta\varphi$ and $\Delta\vartheta$, respectively, which are indicated as the HFNBWs of $|AF_{0,0}(\varphi_i, \vartheta_i)|$. Then, once $\Delta u$ and $\Delta v$ are obtained by numerical calculations, the bore-sight angular resolution is estimated by

$$\begin{aligned} \Delta\varphi_{\text{prop}} &= \arcsin(\Delta u), \\ \Delta\vartheta_{\text{prop}} &= \arcsin(\Delta v), \end{aligned} \tag{11}$$

which is derived in Appendix A.

Unlike the Rayleigh resolution only applicable to uniform arrays, (11) gives a more general angular resolution estimation approach, which is especially useful for non-uniform arrays. In the simulation section, we will compare the angular resolution estimated by the Rayleigh formula and the proposed one for non-uniform arrays, and show that the angular resolution vary with the array geometry even under the same aperture size.

### C. 2D Expanded Beam Pattern

The AAF is a function of four variables $u_j, v_j, u_i, v_i$, and thus it is hard to evaluate AAF patterns of all possible steering directions. Our previous work [33] introduced the concept of EBP for linear arrays so that the dimensionality of AAF can be reduced. Here, a 2D EBP is introduced for planar arrays in this subsection, which includes AAF patterns associated with all $u_j, v_j, u_i, v_i$ in a region of interest (ROI) [25], [28]. Define

$$\begin{aligned} u &= u_i - u_j = \sin\Theta_i \cos\Phi_i - \sin\Theta_j \cos\Phi_j \\ v &= v_i - v_j = \sin\Theta_i \sin\Phi_i - \sin\Theta_j \sin\Phi_j \end{aligned} \tag{12}$$

as the differences between estimated DOAs and real DOAs, and then the four-variable function $AF(u_i, v_i, u_j, v_j)$ in (7) can be simplified to the EBP function

$$\begin{aligned} AF(\boldsymbol{d}, u, v) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m} + x_{r,n})u \\ + (z_{t,m} + z_{r,n})v]\} \end{aligned} \tag{13}$$

In this paper, we focus on array geometry optimization, and hence the position of array elements $\boldsymbol{d} = [\boldsymbol{x}_t^\mathsf{T}, \boldsymbol{z}_t^\mathsf{T}, \boldsymbol{x}_r^\mathsf{T}, \boldsymbol{z}_r^\mathsf{T}]^\mathsf{T}$ becomes a variable in (13).

To determine the ROI of EBP, assume that $\Theta_i \in [0, \hat{\Theta}_{\max}]$ specifies a user-defined field-of-view (FOV) for DOA estimation. As the radiation characteristics of the orthogonal waveform MIMO radar is equivalent to that of a single array element, the range of real DOAs depends on the element radiation pattern. Then, $\Theta_j \in [0, \Theta_{\max}]$ can be determined by the beamwidth of radiation pattern, such as $-10$ dB beamwidth [47]. According to (5), the range of $u_i, v_i, u_j, v_j$ is given by

$$\begin{aligned} u_i^2 + v_i^2 &= \sin^2\Theta_i \le \sin^2\hat{\Theta}_{\max}, \\ u_j^2 + v_j^2 &= \sin^2\Theta_j \le \sin^2\Theta_{\max}. \end{aligned} \tag{14}$$

Using the Cauchy-Schwarz inequality $(u_i u_j + v_i v_j)^2 \le (u_i^2 + v_i^2)(u_j^2 + v_j^2)$, the range of $u, v$ is derived by

$$\begin{aligned} u^2 + v^2 &= (u_i - u_j)^2 + (v_i - v_j)^2 \\ &= (u_i^2 + v_i^2) + (u_j^2 + v_j^2) - 2(u_i u_j + v_i v_j) \\ &\le \sin^2\hat{\Theta}_{\max} + \sin^2\Theta_{\max} + 2\sin\hat{\Theta}_{\max}\sin\Theta_{\max}, \end{aligned} \tag{15}$$

and thus the ROI of EBP is

$$\Omega_{\text{ROI}} = \{(u, v) | u^2 + v^2 \le radius^2\}, \tag{16}$$

where $radius = \sin\hat{\Theta}_{\max} + \sin\Theta_{\max}$.

For example, Fig. 4 in the following section illustrates the EBP with an ROI of $radius = 1.73$, i.e., $\hat{\Theta}_{\max} = 60^\circ, \Theta_{\max} = 60^\circ$, and thus all possible sidelobes of the whole set of AAF patterns can be evaluated. The SLL within the ROI reflects the gross error probability [27], and determines whether angular ambiguity will appear in DOA estimation. To design an ambiguity-free array, the SLL should be maintained below a certain threshold.

In terms of the relationship between the AAF (7) and coherence (8), a lower SLL of EBP is equivalent to a smaller coherence when indexes $i$ and $j$ are quite different, while a narrower MLW corresponds to a smaller coherence when $i$ and $j$ are similar. If DOA estimation is performed by compressed sensing, the SLL threshold should be chosen according to the relationship between the coherence and restricted isometry constant [49]: $\delta_{2S} \le (2S-1)coh$, where $S$ is the sparsity of spatial signals.

### D. Element Position Constraints and Aggregate Function

In order to control the size of antennas, avoid overlaps and mitigate mutual coupling effects, the 2D element spacing and aperture size constraints are constructed for a planar MIMO array in this subsection. Different from the 1D position constraints studied in our previous work [33], the 2D position constraints become non-convex and more steps are needed to deal with the challenging non-smooth constraints.

Consider a multistatic MIMO radar with separation between the transmitter area and receiver area like [39], the 2D position constraints include $x$-aperture size for transmitters, $x$-aperture size for receivers, $z$-aperture size for transmitters, $z$-aperture size for receivers, i.e.,

$$\begin{aligned} \max_{m=1,\ldots,M}\{x_{t,m}\} - \min_{m=1,\ldots,M}\{x_{t,m}\} &\le X_t, \\ \max_{n=1,\ldots,N}\{x_{r,n}\} - \min_{n=1,\ldots,N}\{x_{r,n}\} &\le X_r, \\ \max_{m=1,\ldots,M}\{z_{t,m}\} - \min_{m=1,\ldots,M}\{z_{t,m}\} &\le Z_t, \\ \max_{n=1,\ldots,N}\{z_{r,n}\} - \min_{n=1,\ldots,N}\{z_{r,n}\} &\le Z_r, \end{aligned} \tag{17}$$

$x$-or-$z$-spacing for transmitters and $x$-or-$z$-spacing for receivers, i.e.,

$$\begin{aligned} &|x_{t,m_1} - x_{t,m_2}| \ge \Delta x_t \quad \text{or} \quad |z_{t,m_1} - z_{t,m_2}| \ge \Delta z_t, \\ &\quad (m_1 < m_2,\ m_1 = 1, \ldots, M-1,\ m_2 = 2, \ldots, M), \\ &|x_{r,n_1} - x_{r,n_2}| \ge \Delta x_r \quad \text{or} \quad |z_{r,n_1} - z_{r,n_2}| \ge \Delta z_r, \\ &\quad (n_1 < n_2,\ n_1 = 1, \ldots, N-1,\ n_2 = 2, \ldots, N), \end{aligned} \tag{18}$$

An example of the 2D MIMO element spacing and aperture size constraints is given in Fig. 3, where $X_t$ and $Z_t$ denote the maximum available aperture for transmitters, while $X_r$ and $Z_r$ are for receivers. Considering antenna size and reduction of mutual coupling effects, $\Delta x_t$ and $\Delta z_t$ give the minimum

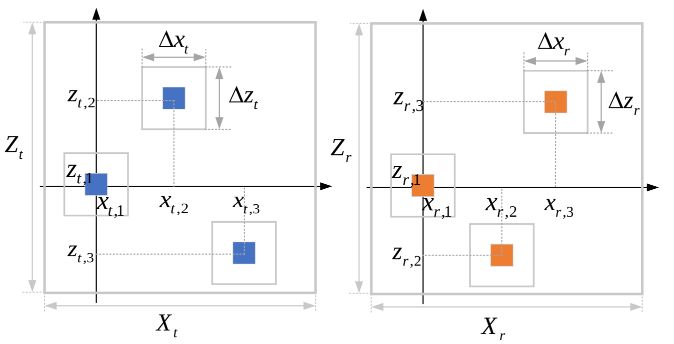


Fig. 3. An example of the 2D element spacing and aperture size constraints.

$x$-spacing and $z$-spacing between transmitters, while $\Delta x_r$ and $\Delta z_r$ are for the receivers. (18) can be rewritten as

$$\begin{aligned}&\max\left\{|x_{t,m_1}-x_{t,m_2}|-\Delta x_t,\ |z_{t,m_1}-z_{t,m_2}|-\Delta z_t\right\}\geq 0,\\&\qquad(m_1<m_2,\ m_1=1,\ldots,M-1,\ m_2=2,\ldots,M),\\&\max\left\{|x_{r,n_1}-x_{r,n_2}|-\Delta x_r,\ |z_{r,n_1}-z_{r,n_2}|-\Delta z_r\right\}\geq 0,\\&\qquad(n_1<n_2,\ n_1=1,\ldots,N-1,\ n_2=2,\ldots,N).\end{aligned}\tag{19}$$

Because the max-value and min-value functions in (17) and (19) are non-smooth, traditional gradient-based techniques for constrained optimization may encounter difficulties in solving them. To smooth multiple max-value functions and min-value functions in the constraints, the aggregate function [51] is introduced as follows

$$\begin{aligned}f_p(\boldsymbol{x})&=\frac{1}{p}\ln\left(\sum_{i=1}^{I}\exp\left\{pg_i(\boldsymbol{x})\right\}\right),\\f_{-p}(\boldsymbol{x})&=\frac{1}{-p}\ln\left(\sum_{i=1}^{I}\exp\left\{-pg_i(\boldsymbol{x})\right\}\right),\end{aligned}\tag{20}$$

where $p>0$ is called the smoothing parameter. According to [51], the aggregate function is often used to approximate max-value or min-value functions, and the following inequality holds:

$$\begin{aligned}&\max_i\left\{g_i(\boldsymbol{x})\right\}\leq f_p(\boldsymbol{x})\leq\max_i\left\{g_i(\boldsymbol{x})\right\}+\frac{1}{p}\ln I,\\&\min_i\left\{g_i(\boldsymbol{x})\right\}-\frac{1}{p}\ln I\leq f_{-p}(\boldsymbol{x})\leq\min_i\left\{g_i(\boldsymbol{x})\right\}.\end{aligned}\tag{21}$$

The maximum approximation error is $\frac{1}{p}\ln I$ if and only if $g_1(\boldsymbol{x})=g_2(\boldsymbol{x})=\cdots=g_I(\boldsymbol{x})$. Otherwise, the approximation error is much smaller than this. Under the condition of not causing computation overflow, the smoothing parameter should be as large as possible for a smaller approximation error. For example, consider $M=16$ and $x_{t,m}\in[0,20\lambda]$ in (17), a $p=35$ is feasible and the maximum approximation error of the max-value or min-value function is $0.079\lambda$, while the average approximation error is only about $5\times10^{-4}\lambda$, which is acceptable. For the functions in (19), the error is smaller because there are only two elements ($I=2$) in the max-value functions.

Although the aggregate function (20) can be used to smooth these functions, it will increase the computational complexity especially when the number of array elements is large. A better approach is to sort the elements by $x$-coordinate, and then linear inequality constraints could be imposed to replace some of the non-smooth constraints in (17). Since the EBP depends only on relative position of the array elements rather than the absolute position [28], the first transmitter and the first receiver can be fixed at the origin as the reference position of the first MIMO virtual element. Therefore, (17) and (19) can be rewritten as

$$\begin{aligned}&x_{t,1}=z_{t,1}=x_{r,1}=z_{r,1}=0,\\&x_{t,m}-x_{t,m-1}\geq 0,\quad m=2,\ldots,M\\&x_{r,n}-x_{r,n-1}\geq 0,\quad n=2,\ldots,N\\&x_{t,M}\leq X_t,\\&x_{r,N}\leq X_r,\\&\max_{m=1,\ldots,M}\{z_{t,m}\}-\min_{m=1,\ldots,M}\{z_{t,m}\}\leq Z_t,\\&\max_{n=1,\ldots,N}\{z_{r,n}\}-\min_{n=1,\ldots,N}\{z_{r,n}\}\leq Z_r,\\&\max\left\{x_{t,m_2}-x_{t,m_1}-\Delta x_t,\ |z_{t,m_2}-z_{t,m_1}|-\Delta z_t\right\}\geq 0,\\&\qquad(m_1<m_2,\ m_1=1,\ldots,M-1,\ m_2=2,\ldots,M),\\&\max\left\{x_{r,n_2}-x_{r,n_1}-\Delta x_r,\ |z_{r,n_2}-z_{r,n_1}|-\Delta z_r\right\}\geq 0,\\&\qquad(n_1<n_2,\ n_1=1,\ldots,N-1,\ n_2=2,\ldots,N).\end{aligned}\tag{22}$$

In a matrix-vector form and with aggregate function, we rewrite (22) as

$$\begin{aligned}&\boldsymbol{C}\boldsymbol{d}\preceq\boldsymbol{c},\\&TA(\boldsymbol{d})\leq 0,\\&RA(\boldsymbol{d})\leq 0,\\&TS_{m_1m_2}(\boldsymbol{d})\leq 0,\ (m_1<m_2,\ m_1=1,\ldots,M-1,\ m_2=2,\ldots,M),\\&RS_{n_1n_2}(\boldsymbol{d})\leq 0,\ (n_1<n_2,\ n_1=1,\ldots,N-1,\ n_2=2,\ldots,N),\end{aligned}\tag{23}$$

where all the coordinate variables are concatenated into a vector $\boldsymbol{d}=[\boldsymbol{x}_t^{\mathsf{T}},\boldsymbol{z}_t^{\mathsf{T}},\boldsymbol{x}_r^{\mathsf{T}},\boldsymbol{z}_r^{\mathsf{T}}]^{\mathsf{T}}$. $\boldsymbol{C},\boldsymbol{c}$ and the aggregate functions $TA,RA,TS_{m_1m_2},RS_{n_1n_2}$ are given in Appendix B.

### E. Problem Formulation

The normalized resolution is expected to be as high as possible without ambiguity in the ROI. Thus, our goal is to optimize the antenna positions under element position constraints so that the MLW of $|AF(\boldsymbol{d},u,v)|$ can be as narrow as possible subject to an SLL constraint. The problem can be formulated as

$$\begin{aligned}\min_{\boldsymbol{d}}\quad&w_u\cdot\Delta u+w_v\cdot\Delta v\\\text{s.t.}\quad&(23),\\&SLL\leq\epsilon,\end{aligned}\tag{24}$$

where $w_u$ and $w_v$ are weight coefficients corresponding to $\Delta u$ and $\Delta v$. Since there are no exact analytical expressions for $\Delta u$ and $\Delta v$, the task of this study is to develop an equivalent mathematical model for (24).

## III. GEOMETRY OPTIMIZATION

In this section, the optimization problem of planar MIMO array geometry for the minimum 2D MLW is modeled as a sparse optimization problem with 2D Domino $l_0$-norm, which can be seen as a very significant improvement of the 1D MLW optimization in our previous work [33] (Algorithm 2). Then, the ADMM with relaxation of the $l_0$-norm and multiple complicated constraints is presented. The synthesis procedure of the proposed MGA-ADMM hybrid algorithm is summarized at the end.

### *A. Slack Mainlobe Constraints and Strict Sidelobe Constraints*

According to [45], [46], to automatically determine the MLW in EBP synthesis, we set a slack mainlobe region as

$$\Omega_{\text{ROIslack}} = \{(u,v)|\frac{u^2}{a^2}+\frac{v^2}{b^2}\le 1\}, \tag{25}$$

where $a \in (0, radius)$ and $b \in (0, radius)$ specify the major and minor axes of the elliptical slack main lobe region, and they can be larger than the initialized HFNBW. $a/b$ will affect the shape of main lobe after optimization; in other words, it will affect the ratio of the normalized resolution in the $u$-direction to that in the $v$-direction: $\Delta u/\Delta v$. Accordingly, the remaining region is defined as a strict sidelobe region

$$\Omega_{\text{ROIstrict}} = \Omega_{\text{ROI}} - \Omega_{\text{ROIslack}}. \tag{26}$$

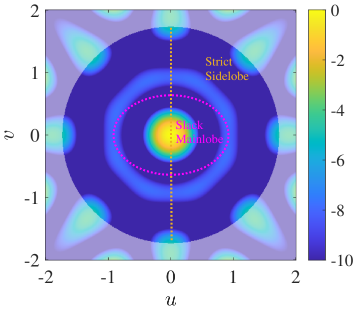


Fig. 4. The color map of EBP for 1T8R UCA, where the ROI is defined in (16), and the slack mainlobe region and strict sidelobe region are defined in (27).

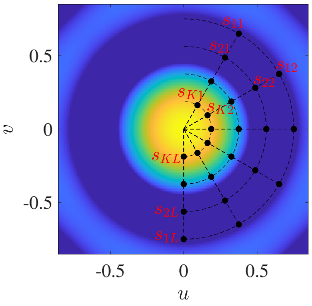


Fig. 5. Sampling values of $s(u,v)$ in a circular slack mainlobe region.

As shown in Fig. 4, symmetry of the EBP allows us to only consider half of the ROI for reduced computational effort, such as

$$\begin{aligned}&\Omega = \Omega_{\text{ROI}} \cap \{(u,v)|u\ge 0\},\\&\Omega_{\text{slack}} = \Omega_{\text{ROIslack}} \cap \{(u,v)|u\ge 0\}\\&\Omega_{\text{strict}} = \Omega_{\text{ROIstrict}} \cap \{(u,v)|u\ge 0\}.\end{aligned} \tag{27}$$

To synthesize an EBP with a small MLW and SLL, different constraint strategies are adopted for the slack mainlobe region and strict sidelobe region. In $\Omega_{\text{strict}}$, strict constraints are imposed to limit the power of EBP to be lower than the given SLL $\epsilon$. In $\Omega_{\text{slack}}$, a two-variable function $s(u,v)\ge 0$ is introduced to allow the power to be higher than the SLL. The constraints for relaxed EBP synthesis can be described as

$$\begin{aligned}\text{s.t.}\ \ &|AF(\boldsymbol{d},u,v)|^2 \le \epsilon + s(u,v), && (u,v)\in\Omega_{\text{slack}},\\&|AF(\boldsymbol{d},u,v)|^2 \le \epsilon, && (u,v)\in\Omega_{\text{strict}},\end{aligned} \tag{28}$$

where the power $|AF(\boldsymbol{d},u,v)|^2$ smooths the absolute value function since gradient-based techniques for constrained optimization may encounter difficulties in tackling the non-smooth absolute value functions. To facilitate the processing of ADMM, $\Omega_{\text{slack}}$ and $\Omega_{\text{strict}}$ are discretized to $K\times L$ and $E$ sampling points, namely $(u_{kl}, v_{kl})$ and $(u_e, v_e)$, respectively, where $kl$ is the index of sampling points in slack region, $e$ the index for strict region. More specifically, $u_{k1},\ldots,u_{kL}$ lie on the same elliptical contour described by $u^2/a^2+v^2/b^2=\sigma_k^2$, where $\sigma_k = 1-\frac{k-1}{K}$ for $k=1,\ldots,K$ denotes a scaling factor of the $k$-th contour. For example, we assume a circular slack mainlobe region (when $a=b$) as shown in Fig. 5. The sampling values of $s(u,v)$ at these $K\times L$ points can be expressed as a matrix

$$\boldsymbol{S} = \begin{bmatrix}\boldsymbol{s}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{s}_K^{\mathrm{T}}\end{bmatrix} = \begin{bmatrix}s_{11} & \ldots & s_{1L}\\ \vdots & \ddots & \vdots\\ s_{K1} & \ldots & s_{KL}\end{bmatrix} = \begin{bmatrix}s(u_{11},v_{11}) & \ldots & s(u_{1L},v_{1L})\\ \vdots & \ddots & \vdots\\ s(u_{K1},v_{K1}) & \ldots & s(u_{KL},v_{KL})\end{bmatrix} \tag{29}$$

Then, we can write (28) in a discrete form and simplify the expression

$$\begin{aligned}\text{s.t.}\ \ &|AF_{kl}(\boldsymbol{d})|^2 \le \epsilon + s_{kl}, && k=1,\ldots,K,\ l=1,\ldots,L,\\&|AF_e(\boldsymbol{d})|^2 \le \epsilon, && e=1,\ldots,E,\end{aligned} \tag{30}$$

where $AF_{kl}(\boldsymbol{d})$ and $AF_e(\boldsymbol{d})$ are abbreviations of $AF(\boldsymbol{d},u_{kl},v_{kl})$ and $AF(\boldsymbol{d},u_e,v_e)$, respectively. To have an effective relaxation, every element of $\boldsymbol{S}$ must be non-negative so that the constraints (30) are feasible, i.e.,

$$\text{s.t.}\ \ s_{kl}\ge 0,\quad k=1,\ldots,K,\ l=1,\ldots,L \tag{31}$$

It should be noted that the discretization may lead to some errors in SLL computation, depending on the selection of $K$, $L$, and $E$, and there is a trade-off between speed and accuracy of SLL computation. As the number of sampling points increases, the SLL error decreases, but the computational cost increases dramatically. In the test, we first use the exhaustive method to obtain accurate SLLs for many different array geometries, and then compare the root mean square errors (RMSE) of the SLLs computed with different numbers of $K$, $L$, and $E$. To balance the computational cost and accuracy, we believe that an SLL RMSE of 0.2 dB is acceptable, thus the parameters are set to $K=L=20$ and $E=159600$ for 4T8R and 16T16R configurations.

### *B. Domino Sparse Optimization Problem*

According to (30), if a side lobe higher than $\epsilon$ occurs at $(u_{kl},v_{kl})$ in the chosen $\Omega_{\text{slack}}$ region, $s_{kl}$ will be greater than zero, while the constraints with $s_{kl}=0$ hold when the corresponding side lobes are lower than $\epsilon$. Since fewer non-zero entries in $\boldsymbol{S}$ imply a narrower mainlobe [45], [46], one can minimize the MLW by sparsifying $\boldsymbol{S}$. However, to conform to the monotonicity of the main lobe, the approach given in [45] requires many constraints on $\boldsymbol{S}$. In this subsection, a Domino sparse method is proposed to avoid imposing monotonicity constraints on $\boldsymbol{S}$.

In this method, the sparsification of $\boldsymbol{S}$ is liken to the falling of a set of dominoes arranged on concentric ellipses. If we were to simultaneously knock down all the dominoes on the outermost ellipse, the inner ellipses would follow suit, falling one after the other from the outside in. Firstly, to ensure that the dominoes on the $k$-th ellipse fall simultaneously, it is equivalent to knocking down the current tallest domino on that ellipse. Therefore, the maximum value of $s_{k1}, s_{k2}, \dots, s_{kL}$ is calculated, which can be approximated by the aggregate function

$$f_p(\boldsymbol{s}_k) = \frac{1}{p} \ln \left( \sum_{l=1}^{L} \exp\{p s_{kl}\} \right) \approx \max_{l=1,\dots,L} \{s_{kl}\}. \tag{32}$$

Subsequently, to facilitate the sequential collapse of the dominoes from the outer to the inner ellipses, a 2D Domino $l_0$-norm is defined as

$$g(\boldsymbol{S}) = \left\| \begin{bmatrix} f_p(\boldsymbol{s}_1) \\ \frac{1}{2}[f_p(\boldsymbol{s}_1) + f_p(\boldsymbol{s}_2)] \\ \vdots \\ \frac{1}{K}[f_p(\boldsymbol{s}_1) + f_p(\boldsymbol{s}_2) + \cdots + f_p(\boldsymbol{s}_K)] \end{bmatrix} \right\|_0, \tag{33}$$

which indicates the sequence in which each ellipse of dominoes falls. (33) is inspired by a design of sparse FIR filters with joint optimization of sparsity and filter order [52]. Different from the random sparse method, the Domino $l_0$-norm gives the variables corresponding to the outer ring of the slack mainlobe region a higher optimization priority, and thus achieving monotonicity. Combining this new objective function and aforementioned constraints, the original problem (24) can be established as a Domino $l_0$-norm optimization problem, given by

$$\begin{aligned} \min_{\boldsymbol{d},\boldsymbol{S}} \quad & g(\boldsymbol{S}) \\ \text{s.t.} \quad & (23), \\ & |AF_{kl}(\boldsymbol{d})|^2 \le \epsilon + s_{kl}, \quad k = 1,\dots,K,\ l = 1,\dots,L, && (34a) \\ & |AF_e(\boldsymbol{d})|^2 \le \epsilon, \qquad e = 1,\dots,E, && (34b) \\ & s_{kl} \ge 0, \qquad k = 1,\dots,K,\ l = 1,\dots,L, && (34) \end{aligned}$$

(34) is a non-convex sparse optimization problem with a large number of highly nonlinear constraints. To relax the $l_0$-norm, the log-sum-exp penalty function is employed [45], which is a special case of the aggregate function with $p = 1$ mentioned in (20), and then the objective function becomes convex:

$$\begin{aligned} \hat{g}(\boldsymbol{S}) = & \ln \Big( \exp\{f_p(\boldsymbol{s}_1)\} + \exp\{\frac{1}{2}[f_p(\boldsymbol{s}_1) + f_p(\boldsymbol{s}_2)]\} + \cdots \\ & + \exp\{\frac{1}{K}[f_p(\boldsymbol{s}_1) + f_p(\boldsymbol{s}_2) + \cdots + f_p(\boldsymbol{s}_K)]\} \Big) \end{aligned} \tag{35}$$

Define new functions $h_{kl}(\boldsymbol{d}, s_{kl}), h_e(\boldsymbol{d})$ and (34a)-(34b) can be rewritten as

$$\begin{aligned} \text{s.t.} \quad & h_{kl}(\boldsymbol{d}, s_{kl}) \triangleq |AF_{kl}(\boldsymbol{d})|^2 - \epsilon - s_{kl} \le 0, \\ & h_e(\boldsymbol{d}) \triangleq |AF_e(\boldsymbol{d})|^2 - \epsilon \le 0. \end{aligned} \tag{36}$$

For brevity, $k = 1,\dots,K,\ l = 1,\dots,L$, and $e = 1,\dots,E$ are omitted below. To relax the complicated inequality constraints, auxiliary variables $w_{kl}, w_e$ are introduced, and (34a)-(34b) are converted to equality constraints:

$$\begin{aligned} \min_{\boldsymbol{d},\boldsymbol{S},\{w_{kl}\},\{w_e\}} \quad & \hat{g}(\boldsymbol{S}) \\ \text{s.t.} \quad & h_{kl}(\boldsymbol{d}, s_{kl}) + w_{kl}^2 = 0, && (37a) \\ & h_e(\boldsymbol{d}) + w_e^2 = 0, && (37b) \\ & (23), \quad s_{kl} \ge 0. && (37) \end{aligned}$$

where $w_{kl}^2 \ge 0,\ w_e^2 \ge 0$. The condition for the equalities to hold is the same as that of (36).

### C. ADMM

ADMM [53] has been widely used for solving various optimization problems due to its distinct decomposition-coordination procedure and superior convergence property. Therefore, ADMM framework is applied to solve (37). First, the augmented Lagrangian function is constructed by

$$\begin{aligned} & \mathcal{L}(\boldsymbol{d}, \boldsymbol{S}, \{w_{kl}\}, \{w_e\}, \{\lambda_{kl}\}, \{\mu_e\}) \\ = & \hat{g}(\boldsymbol{S}) \\ & + \sum_{k=1}^{K} \sum_{l=1}^{L} \Big\{ -\lambda_{kl}[h_{kl}(\boldsymbol{d}, s_{kl}) + w_{kl}^2] + \frac{\rho_1}{2}[h_{kl}(\boldsymbol{d}, s_{kl}) + w_{kl}^2]^2 \Big\} \\ & + \sum_{e=1}^{E} \Big\{ -\mu_e[h_e(\boldsymbol{d}) + w_e^2] + \frac{\rho_2}{2}[h_e(\boldsymbol{d}) + w_e^2]^2 \Big\} \end{aligned} \tag{38}$$

where $\lambda_{kl}$ and $\mu_e$ are dual variables, and $\rho_1$ and $\rho_2$ are user-defined step sizes for the equality constraints (37a) and (37b). In order to eliminate $w_{kl}$ and $w_e$, by $\frac{\partial \mathcal{L}}{\partial w_{kl}} = 0$ and $\frac{\partial \mathcal{L}}{\partial w_e} = 0$, one can deduce that

$$\begin{aligned} w_{kl}^2 &= \begin{cases} \frac{1}{\rho_1}[\lambda_{kl} - \rho_1 h_{kl}(\boldsymbol{d}, s_{kl})], & \text{if } \lambda_{kl} - \rho_1 h_{kl}(\boldsymbol{d}, s_{kl}) > 0 \\ 0, & \text{otherwise} \end{cases} \\ w_e^2 &= \begin{cases} \frac{1}{\rho_2}[\mu_e - \rho_2 h_e(\boldsymbol{d})], & \text{if } \mu_e - \rho_2 h_e(\boldsymbol{d}) > 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{39}$$

Substituting them into (38), we have

$$\begin{aligned} & \mathcal{L}(\boldsymbol{d}, \boldsymbol{S}, \{\lambda_{kl}\}, \{\mu_e\}) \\ = & \hat{g}(\boldsymbol{S}) + \frac{1}{2\rho_1} \sum_{k=1}^{K} \sum_{l=1}^{L} \left[ (\min\{0, \lambda_{kl} - \rho_1 h_{kl}(\boldsymbol{d}, s_{kl})\})^2 - \lambda_{kl}^2 \right] \\ & + \frac{1}{2\rho_2} \sum_{e=1}^{E} \left[ (\min\{0, \mu_e - \rho_2 h_e(\boldsymbol{d})\})^2 - \mu_e^2 \right] \end{aligned} \tag{40}$$

Based on ADMM, $\boldsymbol{S}$, $\boldsymbol{d}$, $\{\lambda_{kl}\}$ and $\{\mu_e\}$ are determined by the following iterative steps, where $t$ represents the number of iterations.

**Step 1:** $\boldsymbol{S}^{(t+1)}$ is determined with given $\boldsymbol{d}^{(t)}$, $\{\lambda_{kl}^{(t)}\}$ and $\{\mu_e^{(t)}\}$.

$$\begin{aligned} \min_{\boldsymbol{S}} \quad & \mathcal{L}(\boldsymbol{d}^{(t)}, \boldsymbol{S}, \{\lambda_{kl}^{(t)}\}, \{\mu_e^{(t)}\}) \\ \text{s.t.} \quad & s_{kl} \ge 0. \end{aligned} \tag{41}$$

Ignoring the constant terms in the Lagrangian function, this subproblem becomes a convex one with piecewise convex objective function and linear constraints shown as

$$\begin{aligned}\min_{\{s_{kl}\}} \quad & \hat{g}(\boldsymbol{S}) + \frac{1}{2\rho_1}\sum_{k=1}^{K}\sum_{l=1}^{L}(\min\{0, \lambda_{kl}^{(t)} - \rho_1 h_{kl}(\boldsymbol{d}^{(t)}, s_{kl})\})^2 \\ \text{s.t.} \quad & s_{kl} \geq 0, \end{aligned} \tag{42}$$

which can be effectively solved by the primal-dual interior-point method [54] and CVX toolbox [55], [56]. The gradient and Hessian matrix of the aggregate function are given in Appendix C.

**Step 2:** $\boldsymbol{d}^{(t+1)}$ is determined with given $\boldsymbol{S}^{(t+1)}$, $\{\lambda_{kl}^{(t)}\}$ and $\{\mu_e^{(t)}\}$.

$$\begin{aligned}\min_{\boldsymbol{d}} \quad & \mathcal{L}(\boldsymbol{d}, \boldsymbol{S}^{(t+1)}, \{\lambda_{kl}^{(t)}\}, \{\mu_e^{(t)}\}) \\ \text{s.t.} \quad & (23) \end{aligned} \tag{43}$$

Ignore the constant terms and (43) can be rewritten as

$$\begin{aligned}\min_{\boldsymbol{d}} \quad & \sum_{k=1}^{K}\sum_{l=1}^{L}\frac{1}{2\rho_1}(\min\{0, \lambda_{kl} - \rho_1 h_{kl}(\boldsymbol{d}, s_{kl}^{(t+1)})\})^2 \\ & + \sum_{e=1}^{E}\frac{1}{2\rho_2}(\min\{0, \mu_e - \rho_2 h_e(\boldsymbol{d})\})^2 \\ \text{s.t.} \quad & (23) \end{aligned} \tag{44}$$

Sequential quadratic programming (SQP) [57] can be used to solve (44). The gradients of the objective function and constraints are derived in Appendix C.

Note that the objective function and its gradient are computationally intensive, because 2D EBP synthesis requires a sufficiently dense discrete sampling grid in the strict sidelobe region, which results in a very large $E$. Since the EBP (see (13)) and its gradient (see (61)) are expressions of NUDFT, they can be efficiently computed based on non-uniform fast Fourier transform (NUFFT) [58]. Therefore, NUFFT is exploited to speed up the computation, and its complexity is given in Appendix D.

**Step 3:** The Lagrangian multipliers $\{\lambda_{kl}^{(t+1)}\}$ and $\{\mu_e^{(t+1)}\}$ are updated with given $\boldsymbol{d}^{(t+1)}$ and $\boldsymbol{S}^{(t+1)}$. By substituting (39) into the multiplier update rule

$$\begin{aligned}\lambda_{kl}^{(t+1)} &= \lambda_{kl}^{(t)} - \rho_1(h_{kl}(\boldsymbol{d}^{(t+1)}, s_{kl}^{(t+1)}) + w_{kl}^2), \\ \mu_e^{(t+1)} &= \mu_e^{(t)} - \rho_2(h_e(\boldsymbol{d}^{(t+1)}) + w_e^2), \end{aligned} \tag{45}$$

we obtain

$$\begin{aligned}\lambda_{kl}^{(t+1)} &= \min\{0, \lambda_{kl}^{(t)} - \rho_1 h_{kl}(\boldsymbol{d}^{(t+1)}, s_{kl}^{(t+1)})\}, \\ \mu_e^{(t+1)} &= \min\{0, \mu_e^{(t)} - \rho_2 h_e(\boldsymbol{d}^{(t+1)})\}. \end{aligned} \tag{46}$$

**Step 4:** According to an easily implementable stopping criterion given by [59], if $t > T$ ($T$ is the maximum number of iterations) or

$$\max\{\alpha, \beta\} \leq \delta, \tag{47}$$

where $\delta$ is the stop tolerance and

$$\begin{aligned}\alpha &= \left|\hat{g}(\boldsymbol{S}^{(t+1)}) - \hat{g}(\boldsymbol{S}^{(t)})\right|, \\ \beta &= \sqrt{\sum_{k=1}^{K}\sum_{l=1}^{L}(h_{kl}^{(t+1)} + w_{kl}^2)^2 + \sum_{e=1}^{E}(h_e^{(t+1)} + w_e^2)^2} \\ &= \sqrt{\sum_{k=1}^{K}\sum_{l=1}^{L}(\max\{h_{kl}^{(t+1)}, \frac{\lambda_{kl}^{(t+1)}}{\rho_1}\})^2 + \sum_{e=1}^{E}(\max\{h_e^{(t+1)}, \frac{\mu_e^{(t+1)}}{\rho_2}\})^2}, \end{aligned}$$

the algorithm will stop; otherwise, set $t := t + 1$, and go to **Step 1**.

### *D. MGA Initialization*

The proposed ADMM algorithm aims to auto-determine and reduce the MLW under the strict sidelobe and element position constraints. But the strict sidelobe constraints in (34b) do not appear to be easily satisfied by a commonly used random initialization, which makes the ADMM maybe run outside the feasible region from the start, because the constraints are highly non-convex. Also, the local optimal solution of a non-convex problem depends heavily on the initial value. Therefore, we employ an efficient modified genetic algorithm (MGA) [41] specifically to provide initial coarse solutions for ADMM that meet the strict SLL constraints. The main advantage of the MGA is that it explores a reduced solution space through an indirect representation of individuals, and it can prevent infeasible solutions during the optimization process using two innovative genetic operators. To enhance efficiency and pursue a globally optimal solution, we introduce a hybrid algorithm, MGA-ADMM, that integrates the global optimization capabilities of heuristic algorithms with the precision of numerical algorithms. Based on the aforementioned process, the overall algorithm is described in **Algorithm 1**.

**Algorithm 1** MGA-ADMM Hybrid Algorithm

```
Input: Number of transmitting antennas $M$ and receiving antennas $N$, geometry constraints $X_t$, $X_r$, $Z_t$, $Z_r$, $\Delta x_t$, $\Delta x_r$, $\Delta z_t$ and $\Delta z_r$, ROI $radius$, slack mainlobe region size $a$ and $b$, desired SLL $\epsilon$, ADMM step size $\rho_1$ and $\rho_2$, initialized Lagrangian multipliers $\{\lambda_{kl}^{(0)}\}$ and $\{\mu_e^{(0)}\}$, maximum number of iterations $T$, stop tolerance $\delta$, population size, crossover rate, mutation rate, maximum number of generations for MGA.
Step 0: According to the MGA, the initialized antenna positions $\boldsymbol{d}^{(0)}$ are synthesized to achieve a minimum SLL, which satisfies the strict sidelobe constraints in (34b).
while $\max\{\alpha, \beta\} > \delta$ and $t \leq T$ do
    Step 1: $\boldsymbol{S}^{(t+1)}$ is determined by the primal-dual interior-point method and CVX toolbox.
    Step 2: $\boldsymbol{d}^{(t+1)}$ is determined by SQP, in which NUFFT is used to accelerate the computation of EBP and its gradient.
    Step 3: Update the Lagrangian multipliers $\{\lambda_{kl}^{(t+1)}\}$, $\{\mu_e^{(t+1)}\}$.
    $t = t + 1$.
end while
Output: optimal antenna positions $\boldsymbol{d}$ and optimized objective function $\hat{g}(\boldsymbol{S})$.
```

## IV. Performance Evaluation

### *A. Parameter Settings*

In the context of the ADMM, the initial values of the Lagrange multipliers $\{\lambda_{kl}^{(0)}\}$ and $\{\mu_e^{(0)}\}$ have a minimal impact on the convergence of the algorithm. Hence, setting their initial values to zero is a safe and commonly used strategy [53]. On the other hand, the step sizes $\rho_1$ and $\rho_2$ significantly influence the algorithm. Our experience suggests that a step size of around $10^3$ offers a good balance between convergence speed and algorithm stability.

As an outstanding work, [39] proposed a method for array performance assessment based on AAF and presented a geometry of 2D MIMO antenna arrays optimized via genetic algorithm (GA). Under the constraints of a certain SLL, aperture size and element spacing, it tries to maximize the ambiguity-free region and simply assumes that the angular resolution is proportional to the aperture size. To have a fair comparison considering the difference in objective functions, the parameters used here are consistent with the 4T8R result given in [39]: the maximum aperture sizes are $X_t = 2.55\lambda$, $Z_t = 5.10\lambda$, $X_r = 6.12\lambda$ and $Z_r = 6.63\lambda$, and single-element dimension amounts to $\Delta x_t = \Delta x_r = 2.88$ mm and $\Delta z_t = \Delta z_r = 3$ mm for 76.5 GHz. Since [39] does not use EBP to define the unambiguous region, the optimized SLL and unambiguous ROI are calculated using numerical methods; specifically, the SLL is $-6.3$ dB and $radius = 1.87$.

Moreover, in addition to the 4T8R configuration, we also provide a 16T16R result with $\Delta x_t = \Delta x_r = \Delta z_t = \Delta z_r = 0.5\lambda$ and an SLL of $-13.3$ dB, and the other parameters remain the same. For comparison, a uniform 18T24R array with equal aperture size of the virtual array is also involved in the simulation as a benchmark.

## B. Optimization Results

The aforementioned MGA-ADMM algorithm for the 4T8R configuration is run 500 times and the result that minimizes the objective function in (37) is then selected. The convergence curve of ADMM objective function (35) is shown in Fig. 6. We also give 500 results of MGA-ADMM, purely GA and random search to demonstrate the advantages of the proposed algorithms. The SLLs versus average MLWs (AMLW, $\frac{\Delta u+\Delta v}{2}$) of the results are plotted in Fig. 7, which indicate that the results with lower SLL and smaller AMLW can be obtained by our MGA-ADMM. The selected 4T8R result and the one given in [39] are marked by black symbols.

The element positions of the optimized 4T8R array and 16T16R array[1] are presented in Fig. 8 and Fig. 11, respectively. For comparison, the 4T8R result in [39] and the uniform 18T24R are given in Figs. 9 and 10, respectively. Their EBPs are shown in Fig. 12. Specifically, Table I lists the SLL, MLW and CRB for 2D DOA estimation [60] of the four arrays, as well as AMLW and average CRB (ACRB). The CRB is given by [28]:

$$CRB_u = \frac{\alpha}{\text{var}\{\boldsymbol{x}\} - \frac{\text{cov}\{\boldsymbol{x},\boldsymbol{z}\}^2}{\text{var}\{\boldsymbol{z}\}}}, \quad CRB_v = \frac{\alpha}{\text{var}\{\boldsymbol{z}\} - \frac{\text{cov}\{\boldsymbol{x},\boldsymbol{z}\}^2}{\text{var}\{\boldsymbol{x}\}}},$$
$$\alpha = \frac{\lambda^2}{8\pi^2 SMN}\frac{1}{\text{SNR}}, \qquad \text{ACRB} = \frac{1}{2}(CRB_u + CRB_v), \tag{48}$$

where $S$, var{} and cov{} denote the number of snapshots, the variance and covariance of the virtual element positions.

It can be seen that the proposed 4T8R array achieves a smaller $\Delta u$ and CRB than the GA result in [39], and the SLL remains the same. Although the uniform array is used as a benchmark in this experiment, it is surprising that the optimized $\Delta u$ and $CRB_u$ of the proposed 4T8R are even lower than those of the uniform 18T24R with the same aperture size due to the larger var$\{\boldsymbol{x}\}$ of the virtual element positions shown in Fig. 8. However, the AMLW and ACRB of the proposed array are still larger than those of the uniform array because the number of 4T8R elements is very small, resulting in a limited degree of freedom for optimization. It can be found that the MLW of non-uniform arrays can vary with array geometry, while the MLW of uniform arrays only depends on the aperture size along the corresponding space dimension. This gives us the opportunity to adjust the azimuth and elevation resolution to some extent by array geometry optimization, even at the same aperture size. The ratio $\Delta u/\Delta v$ can be controlled by adjusting $a$ and $b$ mentioned in (25).

Note that the SLL of 4T8R configurations is about 7 dB higher than the uniform 18T24R which is a significant gap. As the element sparsity rate of the virtual array of 4T8R configuration is only 7%, the SLL inevitably rises compared to the dense uniform array. When we increase the element sparsity rate to 59%, i.e., the proposed 16T16R, the SLL can be a little lower to the uniform 18T24R, and both $\Delta u$ and $\Delta v$ are smaller than that of the uniform array.

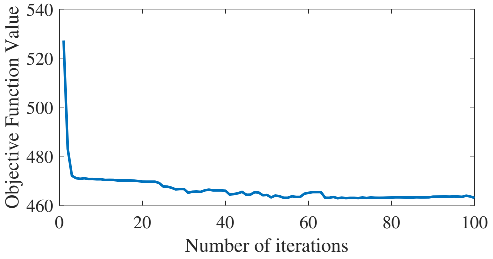


Fig. 6. The convergence curve of ADMM objective function (35) for 4T8R configuration.

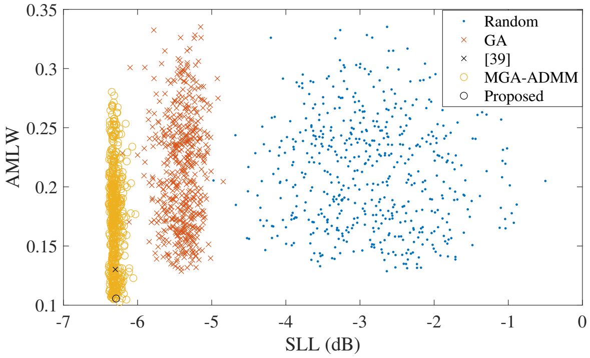


Fig. 7. 500 results of MGA-ADMM, GA and random search for 4T8R configuration.

---

[1] Element position data files for the arrays are available on the “Media” page of this paper at IEEE Xplore.

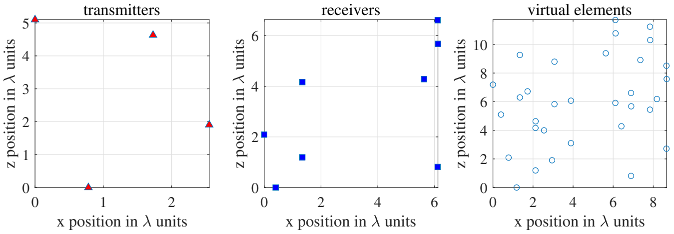


Fig. 8. The proposed 4T8R and virtual element positions optimized via MGA-ADMM.

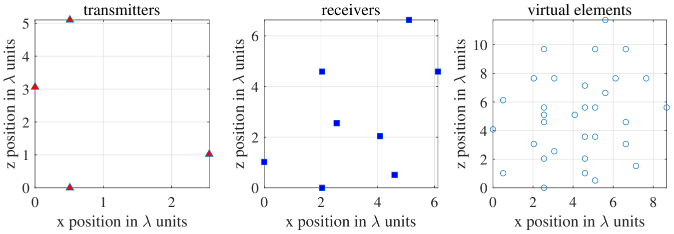


Fig. 9. The 4T8R and virtual element positions optimized via GA [39].

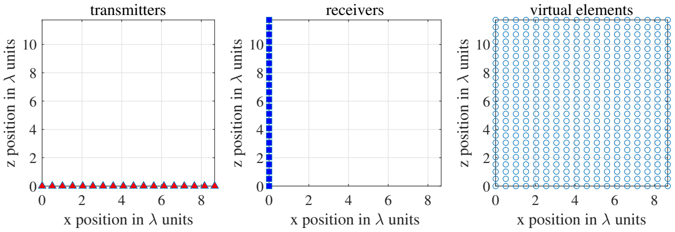


Fig. 10. The uniform 18T24R and virtual element positions.

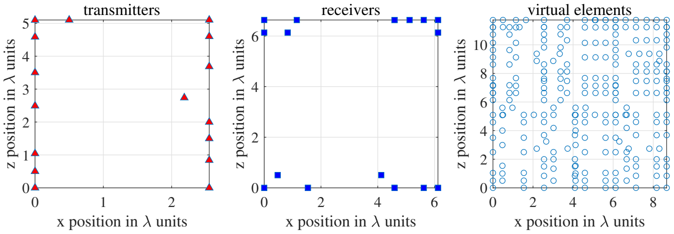


Fig. 11. The proposed 16T16R and virtual element positions optimized via MGA-ADMM.

### C. Angular Resolution Estimation

To compare the accuracy of the angular resolution estimated by the proposal (11) and the Rayleigh criterion (10), spatial spectral power estimation at two known targets is implemented on the four arrays using NUFFT under an ideal scenario without noise, where the powers of the two targets are $P_1$ and $P_2$, respectively.

In optical theory [50], the angular resolution between two point sources is equal to the angle at which the main maximum of the power distribution from the first point source coincides with the first minimum of the power distribution from the second point source. Since this minimum is zero, the angular resolution can be quantified by the RMSE between the estimated power and the actual power at the locations of the two targets. This power RMSE also eliminates the grid error effect inherent in DOA estimation algorithms.

To evaluate the bore-sight $\Delta\varphi$, let $\vartheta_1 = \vartheta_2 = 0$, and the FLAs are set to $\varphi_1 = -\Delta\varphi/2$ and $\varphi_2 = \Delta\varphi/2$, where $\Delta\varphi$ is a variable. As a benchmark for comparison, the true angular resolution $\Delta\varphi_{\text{true}}$ is evaluated by the first null point of the RMSE of estimated powers corresponding to the two targets, given by

$$\begin{aligned}\Delta\varphi_{\text{true}} =& \arg\min_{\Delta\varphi} \sqrt{\frac{|\hat{P}_1(\Delta\varphi) - P_1|^2 + |\hat{P}_2(\Delta\varphi) - P_2|^2}{2}}, \\ & \text{s.t.} \quad 0 < \Delta\varphi \leq \Delta\varphi_{\max} \end{aligned} \tag{49}$$

where $\hat{P}_1(\Delta\varphi)$ and $\hat{P}_2(\Delta\varphi)$ are the estimated powers of the spatial spectrum at the two targets, and they are related to the separation angle $\Delta\varphi$ between the targets. $\Delta\varphi_{\max}$ confines the first null point. Similarly, in order to evaluate the $\Delta\vartheta_{\text{true}}$, let $\varphi_1 = \varphi_2 = 0$, and the FLEs are set to $\vartheta_1 = -\Delta\vartheta/2$ and $\vartheta_2 = \Delta\vartheta/2$. Then $\Delta\vartheta_{\text{true}}$ is evaluated by

$$\begin{aligned}\Delta\vartheta_{\text{true}} =& \arg\min_{\Delta\vartheta} \sqrt{\frac{|\hat{P}_1(\Delta\vartheta) - P_1|^2 + |\hat{P}_2(\Delta\vartheta) - P_2|^2}{2}}, \\ & \text{s.t.} \quad 0 < \Delta\vartheta \leq \Delta\vartheta_{\max} \end{aligned} \tag{50}$$

The true angular resolutions and the ones estimated by the proposed formulas (11) and the Rayleigh criterion (10) are presented in Tab. I, which indicates that the proposed angular resolution is much closer to the true value than the Rayleigh resolution. Due to a smaller $\Delta u$, the $\Delta\varphi_{\text{true}}$ of the proposed arrays is also smaller than those of [39] and the uniform 18T24R array.

### D. DOA Estimation Validation

In this subsection, a numerical study on the angular RMSE of two-target DOA estimation is conducted for four arrays under various conditions. Two commonly used DOA estimation methods are employed in the simulations: CS $l_1$-norm minimization [6] with a single snapshot and MUSIC with 10 snapshots. The power levels of the two uncorrelated signal sources used in the DOA estimation are identical, and 500 Monte Carlo trials are performed.

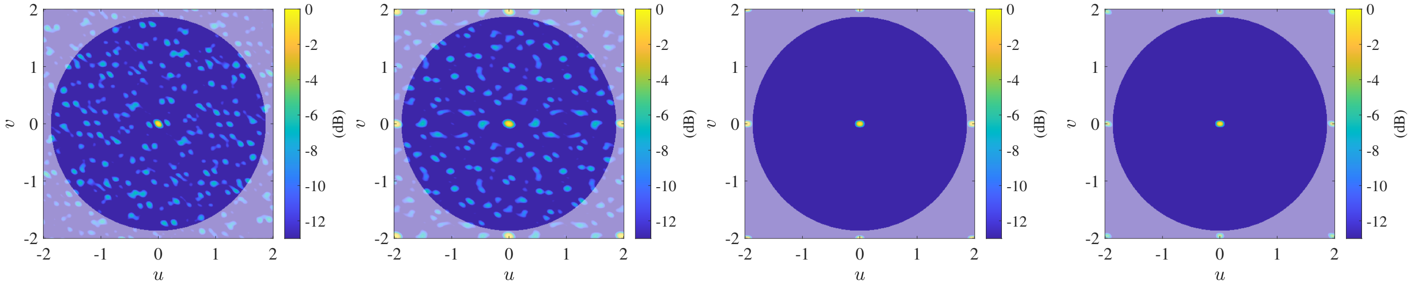
The angular RMSE is evaluated as a function of the SNR, and the noise generated in the simulation is white Gaussian. Considering the two targets with $\varphi$ separation, they are positioned at $(\varphi_1 = -\Delta\varphi_{\text{rayl}}/2, \vartheta_1 = 0)$ and $(\varphi_2 = \Delta\varphi_{\text{rayl}}/2, \vartheta_2 = 0)$, respectively corresponding to the Rayleigh angular resolution of the four arrays in Table I. For the two targets with $\vartheta$ separation, they are positioned at $(\varphi_1 = 0, \vartheta_1 = -\Delta\vartheta_{\text{rayl}}/2)$ and $(\varphi_2 = 0, \vartheta_2 = \Delta\vartheta_{\text{rayl}}/2)$. Fig. 13 illustrates the RMSE versus SNR of the four arrays using CS $l_1$-norm minimization for DOA estimation, and Fig. 14 presents the results using MUSIC.

Additionally, the RMSE as a function of angle separation is assessed in Fig. 15 using CS $l_1$-norm minimization. For the two targets with $\varphi$ separation, $\varphi_2 - \varphi_1$ varies from 2° to 8°. Similarly, for the two targets with $\vartheta$ separation, $\vartheta_2 - \vartheta_1$ also spans from 2° to 8°. The SNR is fixed at 10 dB.

When CS with a single snapshot is employed, except for ultra low SNR conditions ($-10$ dB) in Fig. 13 and very small

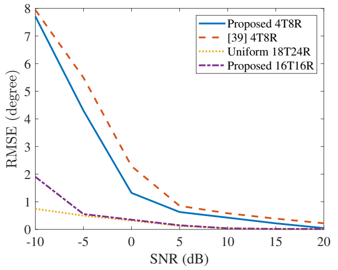
TABLE I
PERFORMANCE COMPARISON

| Array | SLL | $\Delta u$ | $\Delta v$ | AMLW | $CRB_u$ | $CRB_v$ | ACRB | $\Delta\varphi_{\text{true}}$ | $\Delta\varphi_{\text{prop}}$ | $\Delta\varphi_{\text{rayl}}$ | $\Delta\vartheta_{\text{true}}$ | $\Delta\vartheta_{\text{prop}}$ | $\Delta\vartheta_{\text{rayl}}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Proposed 4T8R | $-6.3$ dB | 0.0957 | 0.1154 | 0.1056 | $0.1383\alpha$ | $0.1164\alpha$ | $0.1274\alpha$ | 5.5235° | 5.4944° | 8.0624° | 7.0388° | 6.6243° | 5.9702° |
| [39] 4T8R | $-6.3$ dB | 0.1453 | 0.1152 | 0.1302 | $0.2254\alpha$ | $0.1175\alpha$ | $0.1715\alpha$ | 8.7499° | 8.3532° | 8.0624° | 7.4708° | 6.6152° | 5.9592° |
| Uniform 18T24R | $-13.2$ dB | 0.1089 | 0.0817 | 0.0953 | $0.1425\alpha$ | $0.0801\alpha$ | $0.1113\alpha$ | 6.2445° | 6.2541° | 8.0624° | 4.6823° | 4.6857° | 5.9592° |
| Proposed 16T16R | $-13.3$ dB | 0.1052 | 0.0796 | 0.0924 | $0.1322\alpha$ | $0.0752\alpha$ | $0.1037\alpha$ | 6.1034° | 6.0409° | 8.0624° | 4.6723° | 4.5681° | 5.9566° |



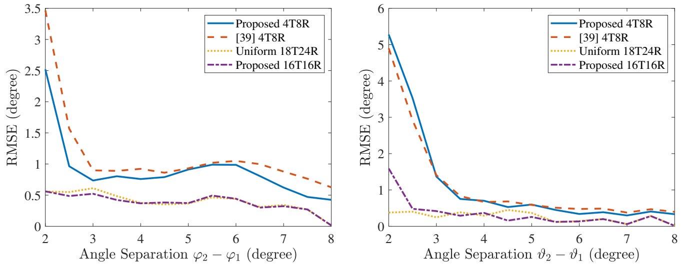
(a) Proposed 4T8R (b) [39] 4T8R (c) Uniform 18T24R (d) Proposed 16T16R

Fig. 12. The expanded beam pattern $|AF(\boldsymbol{d},u,v)|$ color map.



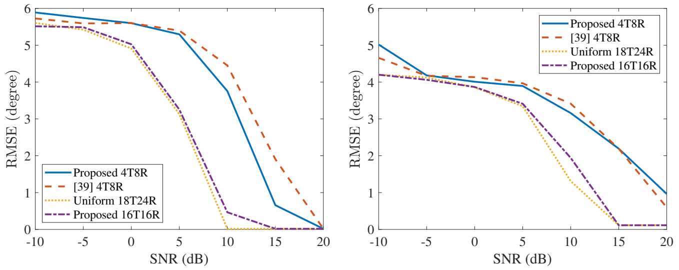
(a) $\varphi$ separation (b) $\vartheta$ separation

Fig. 13. RMSE of DOA estimation versus SNR using CS $l_1$-norm minimization. The angle separation of the two targets is the corresponding Rayleigh angular resolution in Tab. I, and the number of snapshots is 1.



(a) $\varphi$ separation (b) $\vartheta$ separation

Fig. 15. RMSE of DOA estimation versus angle separation using CS $l_1$-norm minimization. The SNR is 10 dB and the number of snapshots is 1.



(a) $\varphi$ separation (b) $\vartheta$ separation

Fig. 14. RMSE of DOA estimation versus SNR using MUSIC. The angle separation of the two targets is the corresponding Rayleigh angular resolution in Tab. I, and the number of snapshots is 10.

$\vartheta$ separation (2°) in Fig. 15(b), the RMSE of the proposed 16T16R is very close to, and occasionally lower than, that of the uniform 18T24R. This suggests that the proposed design might be suitable for 4D imaging automotive radar applications with fewer antennas and without any preprocessing before single-snapshot DOA estimation. The consistent performance of the uniform array at low SNR can be attributed to the fact that its first sidelobe level is $-13.2$ dB, with all other sidelobes being significantly lower than this level.

Concerning the RMSE for $\varphi$ separation as shown in Figs. 13(a), 14(a) and 15(a), the proposed 4T8R generally performs better than the 4T8R in [39] due to a considerably smaller $\Delta u$. They exhibit comparable performance for $\vartheta$ separation as demonstrated in Figs. 13(b), 14(b) and 15(b). Evidence suggests that the mutual coherence given in (8) plays a significant role in achieving uniform recovery guarantees for CS algorithms [6]. Indeed, the MLW minimization with SLL suppression developed in this work is equivalent to reducing coherence length with uniform recovery guarantees, which facilitates high-resolution recovery via $l_1$-norm minimization.

When MUSIC is applied, as shown in Fig. 14, although the proposed 16T16R exhibits marginally lower MLW, CRB and SLL, it does not surpass the performance of the uniform 18T24R at an SNR of 10 dB. This outcome is possibly due to the reduced number of elements in the sparse array, which diminishes the accuracy of estimating the noise subspace. Additionally, at lower SNRs, the RMSE of the proposed 4T8R is slightly higher than that of the 4T8R in [39], despite the former's smaller $\Delta u$. The optimization in this study, based on the analysis of the type I ambiguity function [18], offers foundational performance guarantees for single-snapshot DOA estimation methods such as NUFFT and CS. However, for subspace-based super-resolution methods like MUSIC, the resolution analysis method introduced herein may not provide sufficient accuracy. In this context, type II ambiguity function [61] emerges as a promising avenue for future investigation.

## V. CONCLUSIONS

In this study, a thorough investigation into various aspects of sparse planar MIMO arrays has been carried out, encompassing angular ambiguity functions, angular resolution estimation, expanded beam pattern, and antenna position constraints. The novelty in this paper lies in a more accurate analysis of the angular resolution which is determined by the main lobe width, not just the aperture size. Thus an innovative model was established aimed at reducing the MLW with SLL suppression, in order to improve the resolution under limited aperture size without ambiguities. Inspired by the sequential fall of dominoes, the so-called Domino sparse optimization model auto-determines the 2D MLW, which is a complicated non-convex problem with many highly nonlinear constraints, and it was efficiently solved by the proposed MGA-ADMM algorithm.

In the simulations, two results for 4T8R and 16T16R configurations were provided, and compared with the existing 4T8R result in [39] and the uniform 18T24R as a benchmark. The comparisons were made across the board, including SLL, MLW, CRB and angular resolution, and the RMSE of 2D DOA estimation were evaluated using CS $l_1$-norm minimization and MUSIC under different SNRs and angle separations. In most cases, the performance of the proposed 16T16R array was close to that of the uniform 18T24R array with the same aperture size. In terms of $\varphi$ resolution, the proposed 4T8R generally performs better than the result given in [39].

The proposed method for sparse planar array design aims to not only improve performance for single-snapshot 2D DOA estimation, but also reduce antenna costs, control the size of devices and avoid extra preprocessing, paving the way toward the realization of a low-cost 4D imaging automotive radar.

## APPENDIX A
## ESTIMATION OF $\Delta\varphi$ AND $\Delta\vartheta$

**Proof:** Assume the real DOA is the bore-sight direction, i.e., $u_j = v_j = \varphi_j = \vartheta_j = 0$, and define the auxiliary variables $u_1 < 0 < u_2$. If

$$
\begin{aligned}
&|AF_{0,0}(u_1,0)| = |AF_{0,0}(u_2,0)| = 0,\\
&\begin{cases} \frac{\mathrm{d}}{\mathrm{d}u_i}|AF_{0,0}(u_i,0)| \geq 0, & \text{for } u_i \in [u_1, 0],\\ \frac{\mathrm{d}}{\mathrm{d}u_i}|AF_{0,0}(u_i,0)| \leq 0, & \text{for } u_i \in [0, u_2], \end{cases}
\end{aligned}
\tag{51}
$$

then the HFNBW $\Delta u$ can be expressed by

$$
\begin{aligned}
\Delta u &= (u_2 - u_1)/2\\
&= (\cos(0)\sin(\varphi_2) - \cos(0)\sin(\varphi_1))/2\\
&= \cos(\frac{\varphi_1+\varphi_2}{2})\sin(\frac{\varphi_2-\varphi_1}{2})\\
&= \cos(0)\sin(\Delta\varphi)
\end{aligned}
\tag{52}
$$

where $\varphi_1 + \varphi_2 = 0$ corresponds to $u_1 + u_2 = 0$, and $\Delta\varphi = (\varphi_2 - \varphi_1)/2$. Similarly, for $\vartheta$ separation, we also have

$$
\begin{aligned}
\Delta v &= (v_2 - v_1)/2\\
&= (\sin\vartheta_2 - \sin\vartheta_1)/2\\
&= \cos(\frac{\vartheta_1+\vartheta_2}{2})\sin(\frac{\vartheta_2-\vartheta_1}{2})\\
&= \cos(0)\sin(\Delta\vartheta),
\end{aligned}
\tag{53}
$$

where $\vartheta_1 + \vartheta_2 = 0$ corresponds to $v_1 + v_2 = 0$, and $\Delta\vartheta = (\vartheta_2 - \vartheta_1)/2$. Thus Eq. (11) holds.

## APPENDIX B
## NOTATIONS IN (23)

$$
\boldsymbol{C}_t = \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1\\ & & & 1 \end{bmatrix}_{M\times M}
\quad
\boldsymbol{C}_r = \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1\\ & & & 1 \end{bmatrix}_{N\times N}
$$

$$
\begin{aligned}
&\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_t & \mathbf{0}_{M\times M} & \mathbf{0}_{M\times N} & \mathbf{0}_{M\times N}\\ \mathbf{0}_{N\times M} & \mathbf{0}_{N\times M} & \boldsymbol{C}_r & \mathbf{0}_{N\times N} \end{bmatrix}, \boldsymbol{c} = \begin{bmatrix} \mathbf{0}_{(M-1)\times 1}\\ X_t\\ \mathbf{0}_{(N-1)\times 1}\\ X_r \end{bmatrix}\\
&TA(\boldsymbol{d}) = \frac{1}{p}\ln\left(\sum_{m=1}^{M}\exp\{pz_{t,m}\}\right) + \frac{1}{p}\ln\left(\sum_{m=1}^{M}\exp\{-pz_{t,m}\}\right) - Z_t\\
&RA(\boldsymbol{d}) = \frac{1}{p}\ln\left(\sum_{n=1}^{N}\exp\{pz_{r,n}\}\right) + \frac{1}{p}\ln\left(\sum_{n=1}^{N}\exp\{-pz_{r,n}\}\right) - Z_r\\
&TS_{m_1m_2}(\boldsymbol{d}) = \frac{1}{p}\ln\Big(\exp\{p(x_{t,m_2} - x_{t,m_1} - \Delta x_t)\} + \exp\{p(z_{t,m_2}\\
&\qquad - z_{t,m_1} - \Delta z_t)\} + \exp\{p(z_{t,m_1} - z_{t,m_2} - \Delta z_t)\}\Big)\\
&RS_{n_1n_2}(\boldsymbol{d}) = \frac{1}{p}\ln\Big(\exp\{p(x_{r,n_2} - x_{r,n_1} - \Delta x_r)\} + \exp\{p(z_{r,n_2}\\
&\qquad - z_{r,n_1} - \Delta z_r)\} + \exp\{p(z_{r,n_1} - z_{r,n_2} - \Delta z_r)\}\Big)
\end{aligned}
\tag{54}
$$

## APPENDIX C
## ON COMPUTATION OF THE GRADIENTS IN (44)

Let $F(\boldsymbol{d})$ denote the objective function in (44)

$$
\begin{aligned}
F(\boldsymbol{d}) =& \sum_{k=1}^{K}\sum_{l=1}^{L}\frac{1}{2\rho_1}(\min\{0, \lambda_{kl} - \rho_1 h_{kl}(\boldsymbol{d}, s_{kl}^{(t+1)})\})^2\\
&+ \sum_{e=1}^{E}\frac{1}{2\rho_2}(\min\{0, \mu_e - \rho_2 h_e(\boldsymbol{d})\})^2.
\end{aligned}
\tag{55}
$$

The gradient is derived by

$$
\begin{aligned}
\nabla_{\boldsymbol{d}}F =& \sum_{\tilde{k}}\sum_{\tilde{l}}(\rho_1 h_{\tilde{k}\tilde{l}}(\boldsymbol{d}, s_{\tilde{k}\tilde{l}}^{(t+1)}) - \lambda_{\tilde{k}\tilde{l}})\nabla_{\boldsymbol{d}}h_{\tilde{k}\tilde{l}}\\
&+ \sum_{\tilde{e}}(\rho_2 h_{\tilde{e}}(\boldsymbol{d}) - \mu_{\tilde{e}})\nabla_{\boldsymbol{d}}h_{\tilde{e}}
\end{aligned}
\tag{56}
$$

where

$$\begin{aligned}(\tilde{k},\tilde{l}) &\in \left\{(k,l)\middle|\lambda_{kl} - \rho_1 h_{kl}(\boldsymbol{d}, s_{kl}^{(t+1)}) < 0\right\},\\ \tilde{e} &\in \left\{e\middle|\mu_e - \rho_2 h_e(\boldsymbol{d}) < 0\right\}. \end{aligned} \tag{57}$$

The gradient of $h_{\tilde{k}\tilde{l}}$ and $h_{\tilde{e}}$ is given by

$$\nabla_{\boldsymbol{d}} h_{\tilde{k}\tilde{l}} = \nabla_{\boldsymbol{d}} |AF_{\tilde{k}\tilde{l}}|^2, \quad \nabla_{\boldsymbol{d}} h_{\tilde{e}} = \nabla_{\boldsymbol{d}} |AF_{\tilde{e}}|^2, \tag{58}$$

where

$$\nabla_{\boldsymbol{d}} |AF_{\tilde{k}\tilde{l}}|^2 = \nabla_{\boldsymbol{d}} AF_{\tilde{k}\tilde{l}} \times AF^*_{\tilde{k}\tilde{l}}(\boldsymbol{d}) + AF_{\tilde{k}\tilde{l}}(\boldsymbol{d}) \times (\nabla_{\boldsymbol{d}} AF_{\tilde{k}\tilde{l}})^*, \tag{59}$$

$$\begin{aligned}\nabla_{\boldsymbol{d}} AF_{\tilde{k}\tilde{l}} = \Big[&0, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial x_{t,2}}, \dots, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial x_{t,M}}, 0, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial z_{t,2}}, \dots, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial z_{t,M}},\\ &0, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial x_{r,2}}, \dots, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial x_{r,N}}, 0, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial z_{r,2}}, \dots, \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial z_{r,N}}\Big]^{\mathsf{T}}\end{aligned} \tag{60}$$

$$\begin{aligned}\frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial x_{t,m}} &= \frac{1}{MN}\sum_{n=1}^{N} -\mathrm{j}\frac{2\pi}{\lambda} u_{\tilde{k}} \exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m}+x_{r,n})u_{\tilde{k}}\\ &\qquad + (z_{t,m}+z_{r,n})v_{\tilde{l}}]\}\\ \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial x_{r,n}} &= \frac{1}{MN}\sum_{m=1}^{M} -\mathrm{j}\frac{2\pi}{\lambda} u_{\tilde{k}} \exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m}+x_{r,n})u_{\tilde{k}}\\ &\qquad + (z_{t,m}+z_{r,n})v_{\tilde{l}}]\}\\ \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial z_{t,m}} &= \frac{1}{MN}\sum_{n=1}^{N} -\mathrm{j}\frac{2\pi}{\lambda} v_{\tilde{l}} \exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m}+x_{r,n})u_{\tilde{k}}\\ &\qquad + (z_{t,m}+z_{r,n})v_{\tilde{l}}]\}\\ \frac{\partial AF_{\tilde{k}\tilde{l}}}{\partial z_{r,n}} &= \frac{1}{MN}\sum_{m=1}^{M} -\mathrm{j}\frac{2\pi}{\lambda} v_{\tilde{l}} \exp\{-\mathrm{j}\frac{2\pi}{\lambda}[(x_{t,m}+x_{r,n})u_{\tilde{k}}\\ &\qquad + (z_{t,m}+z_{r,n})v_{\tilde{l}}]\}\end{aligned} \tag{61}$$

Similarly, we can obtain $\nabla_{\boldsymbol{d}} AF_{\tilde{e}}$.

According to [51], the gradient and Hessian matrix of the aggregate function are derived by

$$\begin{aligned}\nabla_{\boldsymbol{x}} f_p(\boldsymbol{x}) =& \sum_{i=1}^{I} \mu_i(\boldsymbol{x}, p)\nabla_{\boldsymbol{x}} g_i(\boldsymbol{x}),\\ \nabla^2_{\boldsymbol{x}} f_p(\boldsymbol{x}) =& \sum_{i=1}^{I} \mu_i(\boldsymbol{x}, p)\nabla^2_{\boldsymbol{x}} g_i(\boldsymbol{x})\\ &+ p\sum_{i=1}^{I} \mu_i(\boldsymbol{x}, p)\nabla_{\boldsymbol{x}} g_i(\boldsymbol{x})\nabla_{\boldsymbol{x}} g_i(\boldsymbol{x})^{\mathsf{T}}\\ &- p\sum_{i=1}^{I} \mu_i(\boldsymbol{x}, p)\nabla_{\boldsymbol{x}} g_i(\boldsymbol{x})\sum_{i=1}^{I} \mu_i(\boldsymbol{x}, p)\nabla_{\boldsymbol{x}} g_i(\boldsymbol{x})^{\mathsf{T}},\end{aligned} \tag{62}$$

where

$$\mu_i(\boldsymbol{x}, p) = \frac{\exp\{p g_i(\boldsymbol{x})\}}{\sum_{i=1}^{I}\exp\{p g_i(\boldsymbol{x})\}} \in (0, 1], \quad \sum_{i=1}^{I} \mu_i(\boldsymbol{x}, p) = 1. \tag{63}$$

Then the gradients and Hessians of the objective function in (42) and constraints in (54) can be derived in a straightforward way.

## APPENDIX D
## CONVERGENCE AND COMPLEXITY ANALYSIS

The model (37) is a non-convex problem with convex objective function and many nonlinear equality constraints. Unfortunately, the convergence guarantee of the proposed ADMM with nonlinear equality constraints (neADMM) has not been explored because convergence conditions in existing literature cannot be applied to neADMM [62]; however, our results show that it outperforms existing state-of-the-art methods. It has been shown in literature that neADMM converges under some mild conditions such as the dual variable convergence [63]:

$$\lim_{t\to\infty} \lambda_{kl}^{(t+1)} - \lambda_{kl}^{(t)} = 0, \quad \lim_{t\to\infty} \mu_e^{(t+1)} - \mu_e^{(t)} = 0, \tag{64}$$

and then there exists a limit point $\{\boldsymbol{d}^\star, \boldsymbol{S}^\star, \{w_{kl}^\star\}, \{w_e^\star\}, \{\lambda_{kl}^\star\}, \{\mu_e^\star\}\}$, which is an optimal solution to (37). It provides some assurance on the reliability of the proposed algorithm.

We first consider the complexity of the subproblem in (42). The main complexity of the log-sum-exp function $\hat{g}(\boldsymbol{S})$ comes from two parts: computing each $f_p(\boldsymbol{s}_k)$, which is done $K$ times, each requiring $\mathcal{O}(L)$, and computing the summations, which is $\mathcal{O}(K^2)$. Thus, the overall complexity of the objective function in (42) is $\mathcal{O}(2KL + K^2)$. Since the subproblem (42) is convex, and $K$ and $L$ are usually small, The primal-dual interior-point method typically converges in less than 0.1 seconds on a modern mid-range PC processor. Most of the computing time is spent in the calculation of the objective function and its gradient in the SQP algorithm for the non-convex subproblem (44). For simplicity, we analyze the time complexity of the innermost function in detail. For each frequency point $(u, v)$, the function $AF(u, v)$ in (13) requires about $MN$ complex additions, which are equivalent to $2MN$ additions of real numbers. Its gradient $\nabla_{\boldsymbol{d}} AF(u, v)$ (see (60)-(61)) requires about $2MN$ complex multiplications and $2MN$ complex additions, which are equivalent to $4MN$ multiplications, $4MN$ multiplication-addition operations, and $4MN$ additions. Therefore, $\mathcal{O}(14MN(KL + E)T_{sqp}T)$ operations are required for the direct computation of EBP, where $KL + E$ is the total number of sampling points in $\Omega_{\text{slack}}$ and $\Omega_{\text{strict}}$, $T_{sqp}$ is the number of iterations of SQP, and $T$ is the number of iterations of ADMM. Aided by NUFFT [58], the time complexity of the algorithm is reduced to about $\mathcal{O}(14(KL + E)\log_2(KL + E)T_{sqp}T)$. Generally speaking, the more antennas there are, the larger the number of frequency sampling points is. For example, $KL + E = 16000$ is enough for the 16T16R configuration. The DFT requires $MN(KL + E) \approx 4 \times 10^6$ operations in a run, while the NUFFT requires only $(KL + E)\log_2(KL + E) \approx 2 \times 10^5$ operations. The SQP algorithm usually requires 5-50 iterations to converge, and ADMM requires at least 50 iterations.

It is worth mentioning that sparse array geometry optimization is not a task that requires to be performed in real-time:

a reasonable computation time is acceptable if the achieved results are excellent.

## REFERENCES

[1] C. Waldschmidt, J. Hasch, and W. Menzel, “Automotive radar — from first efforts to future systems,” *IEEE J. Microw.*, vol. 1, no. 1, pp. 135–148, 2021.

[2] J. Li and P. Stoica, *MIMO Radar Signal Processing*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Oct. 2008.

[3] S. Sun, A. P. Petropulu, and H. V. Poor, “MIMO radar for advanced driver-assistance systems and autonomous driving: Advantages and challenges,” *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 98–117, 2020.

[4] I. Aboumahmoud, A. Muqaibel, M. Alhassoun, and S. Alawsh, “A review of sparse sensor arrays for two-dimensional direction-of-arrival estimation,” *IEEE Access*, vol. 9, pp. 92 999–93 017, 2021.

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

[6] M. Rossi, A. M. Haimovich, and Y. C. Eldar, “Spatial compressive sensing for MIMO radar,” *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 419–430, Jan. 2014.

[7] W. Roberts, P. Stoica, J. Li, T. Yardibi, and F. A. Sadjadi, “Iterative adaptive approaches to MIMO radar imaging,” *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 1, pp. 5–20, Feb. 2010.

[8] W. Liu, M. Haardt, M. S. Greco, C. F. Mecklenbräuker, and P. Willett, “Twenty-five years of sensor array and multichannel signal processing: A review of progress to date and potential research directions,” *IEEE Signal Process. Mag.*, vol. 40, no. 4, pp. 80–91, Jun. 2023.

[9] P. Rocca, G. Oliveri, R. J. Mailloux, and A. Massa, “Unconventional phased array architectures and design methodologies—a review,” *Proc. IEEE*, vol. 104, no. 3, pp. 544–560, Mar. 2016.

[10] M. Pesavento, M. Trinh-Hoang, and M. Viberg, “Three more decades in array signal processing research: An optimization and structure exploitation perspective,” *IEEE Signal Process. Mag.*, vol. 40, no. 4, pp. 92–106, Jun. 2023.

[11] X. Zhang, Z. Zheng, W.-Q. Wang, and H. C. So, “DOA estimation of coherent sources using coprime array via atomic norm minimization,” *IEEE Signal Process. Lett.*, vol. 29, pp. 1312–1316, 2022.

[12] S. Sun and Y. D. Zhang, “4D automotive radar sensing for autonomous vehicles: A sparsity-oriented approach,” *IEEE J. Sel. Top. Signal Process.*, vol. 15, no. 4, pp. 879–891, Jun. 2021.

[13] C. Ang, C. See, and A. Kot, “Optimization of array geometry for identifiable high resolution parameter estimation in sensor array signal processing,” in *Proc. ICICS 1997 Int. Conf. Inf. Commun. Signal Process. Theme Trends Inf. Syst. Eng. Wirel. Multimed. Commun. Cat*, vol. 3, Sep. 1997, pp. 1613–1617 vol.3.

[14] G. Byun, H. Choo, and H. Ling, “Optimum placement of DF antenna elements for accurate DOA estimation in a harsh platform environment,” *IEEE Trans. Antennas Propag.*, vol. 61, no. 9, pp. 4783–4791, Sep. 2013.

[15] H. Gazzah and K. Abed-Meraim, “Optimum ambiguity-free directional and omnidirectional planar antenna arrays for DOA estimation,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3942–3953, 2009.

[16] A. Ghani, F. Keyvani, and S. H. Sedighy, “Antenna array placement on limited bound for isotropic and optimal direction-of-arrival estimation,” *IET Signal Process.*, vol. 12, no. 3, pp. 277–283, 2018.

[17] M. Gavish and A. Weiss, “Array geometry for ambiguity resolution in direction finding,” *IEEE Trans. Antennas Propag.*, vol. 44, no. 6, pp. 889–895, Jun. 1996.

[18] M. Eric, A. Zejak, and M. Obradovic, “Ambiguity characterization of arbitrary antenna array: Type I ambiguity,” *IEEE Int. Symp. Spread Spectr. Tech. Appl.*, vol. 3, no. October, pp. 955–957, 1998.

[19] Y. Liang, W. Liu, Q. Shen, W. Cui, and S. Wu, “A review of closed-form cramér-rao bounds for DOA estimation in the presence of gaussian noise under a unified framework,” *IEEE Access*, vol. 8, pp. 175 101–175 124, 2020.

[20] U. Baysal and R. Moses, “On the geometry of isotropic arrays,” *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1469–1478, Jun. 2003.

[21] U. Oktel and R. Moses, “A bayesian approach to array geometry design,” *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1919–1923, 2005.

[22] E. Tohidi, M. Coutino, S. P. Chepuri, H. Behroozi, M. M. Nayebi, and G. Leus, “Sparse antenna and pulse placement for colocated MIMO radar,” *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 579–593, Feb. 2019.

[23] M. A. González-Huici, D. Mateos-Núñez, C. Greiff, and R. Simoni, “Constrained optimal design of automotive radar arrays using the weiss-weinstein bound,” in *2018 IEEE MTT- Int. Conf. Microw. Intell. Mobil. ICMIM*, Apr. 2018, pp. 1–4.

[24] D. Khan and K. L. Bell, “Explicit ziv–zakai bound for analysis of DOA estimation performance of sparse linear arrays,” *Signal Processing*, vol. 93, no. 12, pp. 3449–3458, Dec. 2013.

[25] A. Gupta, U. Madhow, A. Arbabian, and A. Sadri, “Design of large effective apertures for millimeter wave systems using a sparse array of subarrays,” *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6483–6497, Dec. 2019.

[26] J. Tabrikian, O. Isaacs, and I. Bilik, “Cognitive antenna selection for automotive radar using bobrovsky-zakai bound,” *IEEE J. Sel. Top. Signal Process.*, vol. 15, no. 4, pp. 892–903, Jun. 2021.

[27] T. Birinci and Y. Tanik, “Optimization of nonuniform array geometry for DOA estimation with the constraint on gross error probability,” *Signal Processing*, vol. 87, no. 10, pp. 2360–2369, Oct. 2007.

[28] O. Lange and B. Yang, “Antenna geometry optimization for 2D direction-of-arrival estimation for radar imaging,” in *2011 Int. ITG Workshop Smart Antennas*, Feb. 2011, pp. 1–8.

[29] S. Kiani and A. M. Pezeshk, “A comparative study of several array geometries for 2D DOA estimation,” *Procedia Computer Science*, vol. 58, pp. 18–25, Jan. 2015.

[30] F. Keskin and T. Filik, “An optimum volumetric array design approach for both azimuth and elevation isotropic DOA estimation,” *IEEE Access*, vol. 8, pp. 183 903–183 912, 2020.

[31] L. R. F.R.S, “XXXI. investigations in optics, with special reference to the spectroscope,” *Lond. Edinb. Dublin Philos. Mag. J. Sci.*, Oct. 1879.

[32] M. Plotkin, “Beamwidth of phased arrays,” *IEEE Trans. Antennas Propag.*, vol. 21, no. 5, pp. 695–697, 1973.

[33] M. Huan, J. Liang, Y. Wu, Y. Li, and W. Liu, “SASA: super-resolution and ambiguity-free sparse array geometry optimization with aperture size constraints for MIMO radar,” *IEEE Trans. Antennas Propag.*, vol. 71, no. 6, pp. 4941–4954, Jun. 2023.

[34] J. Mayer, B. Nuss, J. Kowalewski, and T. Zwick, “Angular resolution estimation for conformal radar antenna arrays,” in *2018 IEEE MTT- Int. Conf. Microw. Intell. Mobil. ICMIM*, Apr. 2018, pp. 1–4.

[35] N. D. Tran, R. Boyer, S. Marcos, and P. Larzabal, “On the angular resolution limit for array processing in the presence of modeling errors,” *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4701–4706, Oct. 2013.

[36] N. Dowlut and A. Manikas, “A polynomial rooting approach to super-resolution array design,” *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1559–1569, Jun. 2000.

[37] M. Viberg and C. Engdahl, “Element position considerations for robust direction finding using sparse arrays,” in *Conf. Rec. Thirty-Third Asilomar Conf. Signals Syst. Comput. Cat NoCH37020*, vol. 2, 1999, pp. 835–839 vol.2.

[38] X. H. Wu, A. A. Kishk, and A. W. Glisson, “Antenna effects on a monostatic MIMO radar for direction estimation, a cramèr-rao lower bound analysis,” *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 2388–2395, Jun. 2011.

[39] A. Di Serio, P. Hugler, F. Roos, and C. Waldschmidt, “2-D MIMO radar: A method for array performance assessment and design of a planar antenna array,” *IEEE Trans. Antennas Propag.*, vol. 68, no. 6, pp. 4604–4616, 2020.

[40] C. Vasanelli, F. Roos, A. Durr, J. Schlichenmaier, P. Hugler, B. Meinecke, M. Steiner, and C. Waldschmidt, “Calibration and direction-of-arrival estimation of millimeter-wave radars: A practical introduction,” *IEEE Antennas Propag. Mag.*, vol. 62, no. 6, pp. 34–45, 2020.

[41] K. Chen, X. Yun, Z. He, and C. Han, “Synthesis of sparse planar arrays using modified real genetic algorithm,” *IEEE Trans. Antennas Propag.*, vol. 55, no. 4, pp. 1067–1073, Apr. 2007.

[42] T. Pavlenko, C. Reustle, Y. Dobrev, M. Gottinger, L. Jassoume, and M. Vossiek, “Design and optimization of sparse planar antenna arrays for wireless 3-D local positioning systems,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 7288–7297, 2017.

[43] X. He, C. Alistarh, and S. K. Podilchak, “Optimal MIMO sparse array design based on simulated annealing particle swarm optimization,” in *2022 16th Eur. Conf. Antennas Propag. EuCAP*, Mar. 2022, pp. 1–5.

[44] D. Mateos-Nunez, M. A. Gonzalez-Huici, R. Simoni, F. B. Khalid, M. Eschbaumer, and A. Roger, “Sparse array design for automotive MIMO radar,” *EuRAD 2019 - 2019 16th Eur. Radar Conf.*, pp. 249–252, 2019.

[45] W. Peng, T. Gu, Y. Zhuang, Z. He, and C. Han, “Pattern synthesis with minimum mainlobe width via sparse optimization,” *Digital Signal Processing*, vol. 128, p. 103632, Aug. 2022.

[46] K. Tang and X. Zhang, "Auto-determining mainlobe width for beampattern synthesis via relaxation optimization," *IEEE Signal Process. Lett.*, vol. 29, pp. 314–318, 2022.

[47] C. A. Balanis, *Antenna Theory: Analysis and Design*, fourth edition ed. Hoboken, New Jersey: Wiley, 2016.

[48] R. L. Haupt, *Antenna Arrays: A Computational Approach*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Mar. 2010.

[49] H. Rauhut, "Compressive sensing and structured random matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery*, M. Fornasier, Ed. DE GRUYTER, Jul. 2010, pp. 1–92.

[50] M. Born and E. Wolf, *Principles of Optics: 60th Anniversary Edition*, 7th ed. Cambridge: Cambridge University Press, 2019.

[51] E. Polak, R. S. Womersley, and H. X. Yin, "An algorithm based on active sets and smoothing for discretized semi-infinite minimax problems," *J Optim Theory Appl*, vol. 138, no. 2, pp. 311–328, Aug. 2008.

[52] A. Jiang, H. K. Kwan, Y. Zhu, X. Liu, N. Xu, and Y. Tang, "Design of sparse FIR filters with joint optimization of sparsity and filter order," *IEEE Trans. Circuits Syst. Regul. Pap.*, vol. 62, no. 1, pp. 195–204, Jan. 2015.

[53] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *FNT in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[54] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. Cambridge University Press, Mar. 2004.

[55] M. Grant and S. Boyd, "CVX: matlab software for disciplined convex programming, version 2.1," Mar. 2014.

[56] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.

[57] J. Nocedal and S. J. Wright, *Numerical Optimization Second Edition*. World Scientific, 1999.

[58] A. Dutt and V. Rokhlin, "Fast fourier transforms for nonequispaced data," *SIAM J. Sci. Comput.*, vol. 14, no. 6, pp. 1368–1393, Nov. 1993.

[59] D. Han and X. Yuan, "A note on the alternating direction method of multipliers," *J Optim Theory Appl*, vol. 155, no. 1, pp. 227–238, Oct. 2012.

[60] H. L. Van Trees, "Parameter estimation I: maximum likelihood," in *Optimum Array Processing*. John Wiley & Sons, Ltd, 2002, ch. 8, pp. 917–1138.

[61] M. Eric, A. Zejak, and M. Obradovic, "Ambiguity characterization of arbitrary antenna array: Type II ambiguity," in *1998 IEEE 5th Int. Symp. Spread Spectr. Tech. Appl. - Proc. Spread Technol. Afr. Cat No98TH8333*, vol. 3, Sep. 1998, pp. 955–958 vol.3.

[62] J. Wang and L. Zhao, "Nonconvex generalization of alternating direction method of multipliers for nonlinear equality constrained problems," *Results in Control and Optimization*, vol. 2, p. 100009, Apr. 2021.

[63] W. Fan, J. Liang, and J. Li, "Constant modulus MIMO radar waveform design with minimum peak sidelobe transmit beampattern," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4207–4222, Aug. 2018.

**Mingsai Huan** (S'24) was born in Weifang, China, in 1996. He received the B.Eng. degree in 2018 and the Ph.D. degree in 2024, from Northwestern Polytechnical University, Xi'an, China. He served as a visiting research student at the Institute for Infocomm Research (I2R) of A*STAR in Singapore from 2023 to 2024. He has accepted a position as a tenure-track lecturer at Ocean University of China in Qingdao, to begin in 2024. His research focuses on sparse arrays and their applications in MIMO radar direction finding, localization, and imaging.

**Junli Liang** (SM'16) was born in China. He received the Ph.D. degree in signal and information processing from the Chinese Academy of Sciences, Beijing, China, in 2007. He is currently a Professor with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China. His research interests include radar signal processing, machine learning, 6G, and their applications.

**Yugang Ma** (Senior Member, IEEE) received the B.S. degree in radio engineering and the M.S. degree in computer engineering from Chongqing University, China, in 1990 and 1995 respectively. Subsequently in 1998, he received the Ph.D. degree in telecommunications engineering from the University of Electronic Science and Technology of China (UESTC). Dr. Ma has been working with the Institute for Infocomm Research (I2R) of A*Star since 2013. He has been the Group Leader for white space communications research and is currently a Senior Principal Scientist. From 2014 to 2016, he was an Adjunct Professor of UETSC. From 2000 to 2013, he worked with the Singapore Research Laboratory, Sony Electronics, where he was a Senior Researcher. From 1998 to 2000, he was a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research areas include system and signal processing for radar sensing and communications, deep learning, MIMO mmWave vital sign & automotive radars, localization, OFDM, UWB, TVWS, DSA, time synchronization, time sensitive networks, interference suppression and antenna technologies. Dr. Ma holds over 50 granted patents. He was a Distinguished Engineer of Sony Corporation receiving 8 patent awards. He was also the recipient of Singapore Manufacturing Federation Award in 2014. He was the Chief Editor of IMDA TSAC WG6 IoT LoRa taskforce (2016-2017) and NB-IoT taskforce (2018-2019), as well as the Editor of IMDA TSAC WG6 5G NR CBS CMT taskforce (2020-2021).

**Wei Liu** (S'01-M'04-SM'10) received his BSc and LLB degrees from Peking University, China, in 1996 and 1997, respectively, MPhil from University of Hong Kong in 2001, and PhD from University of Southampton, UK, in 2003. He then worked as a postdoc first at Southampton and later at Imperial College London. In September 2005, he joined the University of Sheffield, UK, first as a Lecturer and then a Senior Lecturer. Since September 2023, he has been a Reader at the School of Electronic Engineering and Computer Science, Queen Mary University of London, UK. He has published 430+ journal and conference papers, six book chapters, and two research monographs titled "Wideband Beamforming: Concepts and Techniques" (Wiley, March 2010) and "Low-Cost Smart Antennas" (Wiley, March 2019), respectively. His research interests cover a wide range of topics in signal processing, with a focus on sensor array signal processing and its various applications.

**Yifan Wu** was born in China. He is currently working toward the Ph.D. degree with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China. His research focuses on radar signal processing and its application.

**Yonghong Zeng** (Fellow, IEEE) received the B.S. degree from Peking University, M.S. degree and Ph.D. degree from National University of Defense Technology, China. He has been working in the Institute for Infocomm Research, A*STAR, since 2004, and currently is a senior principal scientist. He is working on integrated sensing and communication, B5G/6G communications, ultra-reliable low-latency communication, vehicular radar and communication, and real time localization system. He has co-authored six books and more than 90 refereed journal papers. He received the 2007 and 2009 IES (Institute of Engineers Singapore) prestigious engineering achievement award in Singapore, and ministry-level Scientific and Technological Development Awards four times in China. He received the award for advances in communication from IEEE Communication society in 2022, the Inaugural IEEE Communications Society Asia-Pacific Best Paper Award in 2012, and a few other best paper awards. He received the Certificate of Appreciation for outstanding contributions to the IEEE 802.22 standard in 2011.