On beam design for sparse arrays of subarrays using multi-objective optimization and estimation-theoretic criteria

A. Gupta[(1)], U. Madhow[(1)], A. Arbabian[(2)] and Ali Sadri[(3)]

(1) ECE Department, University of California, Santa Barbara, CA, USA
(2) EE Department, Stanford University, Stanford, CA, USA, (3) Intel Corporation
Emails: {anantgupta, madhow}@ece.ucsb.edu, arbabian@stanford.edu, ali.s.sadri@intel.com

***Abstract*—Compact antenna arrays with a moderately large number of elements (e.g., 16 elements in a 4×4 planar configuration) can be realized at low cost at millimeter wave frequencies. This provides the opportunity for cost-effective synthesis of large aperture antenna arrays using sparse placement of such compact "subarrays". In this paper, we investigate the design of *arrays of subarrays* constructed in this fashion, constraining the number of subarrays and the size of the area over which they are placed, with a view to mitigating the grating lobes created by multi-wavelength separation of the subarrays. We consider multi-objective optimization of directivity, beamwidth, maximum sidelobe level and beam eccentricity. While exhaustive search for the space of optimal solutions is computationally infeasible, we obtain intuitively plausible configurations using simple search strategies for finding local minima. We compare example configurations based on their performance for Direction of Arrival (DoA) estimation.**

I. Introduction

Large aperture antenna arrays with densely spaced elements can form sharp beams, which provide accurate DoA estimates. However, such antennas are expensive in terms of both cost and power consumption. On the other hand, compact antenna arrays with a moderately large number of elements can be realized at relatively low cost, especially as the carrier frequency increases. Thus, a cost-effective approach to creating large apertures (and hence sharp beams) is via sparse placement of a number of such compact arrays, henceforth termed "subarrays," optimizing the placement (and controlling the phases) so as to reduce unwanted grating lobes. As a running example throughout this paper, we consider the design of a 60 GHz array of subarrays created by placing 8 subarrays over an aperture size of 10 cm by 10 cm ($20\lambda \times 20\lambda$ for $\lambda = 0.5cm$), where each subarray has 4×4 elements arranged in uniform rectangular grid with 0.5λ horizontal spacing and 0.6λ vertical spacing. The latter parameters are associated with a prototype subarray. Each subarray unit occupies extra physical area on the plane which must also be accounted for.

There is a significant body of prior work on reducing grating lobes in one-dimensional sparse arrays, including minimum redundancy arrays [1], genetic optimization[2], joint Cramér Rao Bound and sidelobe level optimization [3]. Most of these works either focused on uniform arrays or non-uniform arrays where all elements have flexibility of being placed freely.

Design of linear arrays of two and three subarrays was investigated in [4]. However extension of these techniques for planar arrays is not straightforward. Recent approaches like Nested 2D arrays [5] and H-arrays [6] utilize the idea of difference co-array to reduce the number of redundant spacings and maximize the randomness of element positions, so that the number of spatial frequencies being sampled by the array is maximized. These approaches allow for a richer design space but complicate the problem of identifying the optimal configurations, as the number of possibilities grows exponentially and becomes intractable. In addition, they do not apply when the element placement within subarrays is constrained, as in our setting.

Arrays of subarrays can be used to create spatial degrees of freedom if the subarrays are spaced far enough apart, where the required spacing increases with range and decreases with carrier frequency [7]. This is not the regime of interest to us here (we are interested in relatively small form factors and larger ranges), and our focus here is on beamforming performance. We employ the performance of DoA estimation as a concrete means of comparing different configurations, which implicitly captures the effects of the strengths of main lobe and grating/side lobes: a tight main lobe improves DoA estimation, while strong grating and side lobes degrade performance.

Contributions and Organization: In Section II, we formulate a multi-objective optimization problem for choice of configuration. In Section III, we describe our approach and algorithms for optimization, where the main effort is in mitigating the combinatorial explosion in the number of possible of configurations to search over. Example configurations and beamsteering results are discussed in Seciton IV. Section V contains comparison of these configurations for DoA estimation via simulations of maximum likelihood (ML) DoA estimation and the corresponding Cramer-Rao bounds. Section VI contains our conclusions.

II. Problem Formulation

The design of the array involves jointly optimizing multiple cost functions in a manner that is tailored to the application. For DoA estimation, for instance, we would like to minimize

beamwidth while controlling ambiguities caused by sidelobes. We consider the following performance metrics,

- BW: Approximating the beam shape for a non-uniform configuration as elliptical, we define beamwidth as the product of the 3-dB bandwidth along the major and minor axes:$\texttt{BW} = BW_{max} \times BW_{min}$
- MSLL: Maximum sidelobe level is the power level of the strongest sidelobe generated in the beampattern.
- G_D: Directivity of planar array which is the ratio of mainlobe power to average power [8], $G_D = 4\pi \frac{R_{max}}{R_{Total}}$
- $\texttt{ecc} = \sqrt{1 - (BW_{min}/BW_{max})^2}$ is the eccentricity of the 3-dB mainbeam ellipse.

The preceding metrics have to be computed numerically in general.

Our goal is to find the subarray centers and beamsteering weights which minimize the overall cost function, which we define as

$$f(\boldsymbol{C}, \boldsymbol{w}) = \alpha \texttt{BW}(\boldsymbol{C}, \boldsymbol{w}) + \beta \texttt{MSLL}(\boldsymbol{C}, \boldsymbol{w}) + \gamma \texttt{ecc}(\boldsymbol{C}, \boldsymbol{w}) \quad (1)$$

where $\boldsymbol{C}$ is a $N_{subarray} \times 2$ Subarray center position matrix, $\boldsymbol{w}$ is an $M \times 1$ steering weight vector and $[\alpha, \beta, \gamma]$ are parameters for sweeping through the optimal surface for this multi-objective optimization. The maximum sidelobe level metric is especially relevant for sparse arrays, since they exhibit significant grating lobes. We therefore add additional constraints to encourage a maximum sidelobe of no more than -10 dB relative to the main lobe, and allow 3 dB slack for directivity. This restricts the ambiguity caused by a distant sidelobe to a certain level (10 dB below the mainlobe in our case), which reduces the probability of large errors in low-SNR regimes for DoA estimation, for example. For large non-uniform planar arrays, the location and level of grating lobes is difficult to predict and control, hence we resort to empirical evaluations to evaluate the cost function for a given configuration.

Sidelobe reduction can be achieved by amplitude tapering using, for example, the Dolph Chebychev beamformer for uniform arrays, and more recently by convex optimization for arbitrary arrays [9]. However, we only consider unit magnitude weights (no amplitude tapering) to allow maximum possible transmit power in the desired direction i.e. $w_i = e^{j\psi_i}$. Hence the phase optimization problem is reduced to finding appropriate phases ψ_i for all array elements.

III. Optimization Approach

The cost function in (1) is evaluated using the beampattern $R(u, v)$ of the array which in turn depends on the beamsteering direction (u_0, v_0). For ideal steering phases,

$$R(u, v) = \frac{1}{M^2} \left\| \sum_{i=1}^{M} e^{jk\left((u-u_0)d_i^x + (v-v_0)d_i^y\right)} \right\|^2 \quad (2)$$

where M is number of array elements; $k = \frac{2\pi}{\lambda}$ is the wavenumber; $u = \sin(\theta)\cos(\phi), v = \sin(\theta)\sin(\phi)$ are directional cosines; θ, ϕ are the elevation and azimuth angles,

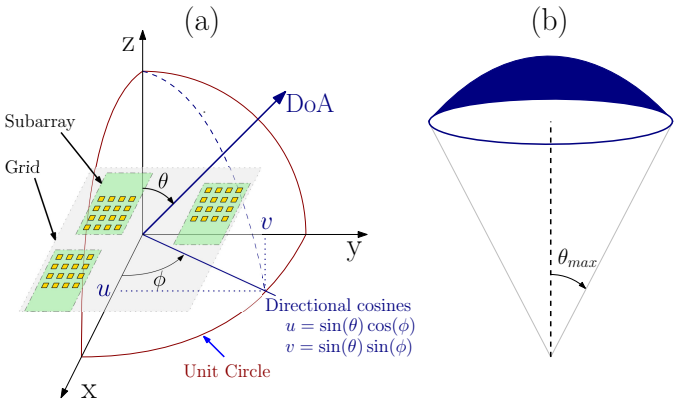

Fig. 1. (a) 2D Array Geometry and Spherical coordinate system. (b)ROI: Uniform distribution of DoA in spherical cap with half angle θ_{max}

respectively; and $[d_i^x, d_i^y]^T \triangleq \boldsymbol{d}_i$ are the 2D coordinates of array elements. Since the elements in the subarray have fixed relative locations, the individual element location matrix is obtained from subarray centers as $\boldsymbol{D} = \boldsymbol{C} \otimes \mathbb{1}_{N_e} + \boldsymbol{D}_e \otimes \mathbb{1}_{N_{sub}}$, where $\boldsymbol{D}_e$ is the $2 \times N_e$ matrix containing the subarray element coordinates with respect to its center, N_{sub} is the number of Subarrays, N_e is number of elements in individual subarrays, $\mathbb{1}_n$ is $n \times 1$ column vector of ones, and $\otimes$ denotes the Kronecker product.

It would be prohibitively expensive to evaluate the cost functions over all such beampatterns for finding the optimal $(\boldsymbol{C}, \boldsymbol{w})$. However, for a Region of Interest (ROI) with maximum steering angle θ_{max}, we can define a worst-case model via an expanded beampattern that subsumes the beampatterns of all steering directions in the ROI, as follows [10]:

$$R_\rho(\tilde{u}, \tilde{v}) = \frac{1}{M^2} \left\| \sum_{i=1}^{M} e^{jk\rho\left(\tilde{u}d_i^x + \tilde{v}d_i^y\right)} \right\|^2$$
$$\rho\tilde{u} = u - u_0, \rho\tilde{v} = v - v_0$$
$$\rho = 1 + \sin(\theta_{max})$$

Figure 2 shows the expanded beampattern $R_{1.5}(\tilde{u}, \tilde{v})$ for ROI with $\theta_{max} < 30°$, which incorporates the actual beampattern information of scan angles $(0.3, 0.4)$. This allows us to calculate the worst case value of MSLL over all steering direction in the ROI from the expanded beampattern. Therefore, assum-

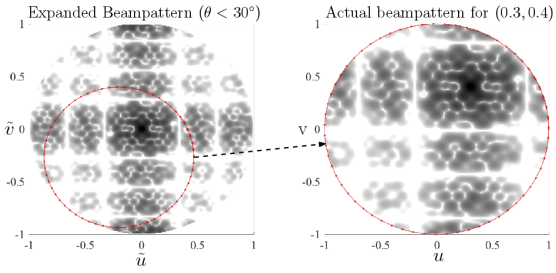

Fig. 2. Expanded Beampattern for ROI ($\theta_{max} < 30°$)

ing ideal beamforming, placement optimization can be made invariant to steering weights, and can therefore be decoupled from weight optimization.

Our approach is to first search for the optimal placement $C^*$of subarrays by fixing the steering phases to broadside i.e. $\boldsymbol{w} = \mathbb{1}_M$. The constrained Multi-objective optimization is expressed as:

$$\begin{aligned} \min_{\boldsymbol{C}\in\mathcal{C}} \quad & f(\boldsymbol{C}, \mathbb{1}_M) \\ \text{subject to} \quad & \texttt{MSLL}(\boldsymbol{C}, \mathbb{1}_M) \leq -10, \\ & \texttt{G}_D(\boldsymbol{C}, \mathbb{1}_M) \geq \texttt{G}_{D,max} - 3 \end{aligned} \tag{3}$$

The domain of optimization $\mathcal{C}$ is evaluated by constructing a dictionary of distinct configurations using the covariance of subarray centers as a similarity measure. We discuss the details of search procedure in Section III-A.

For Phase optimization, we consider quantized phase shifters at each element. The resulting beampattern obtained from the array deviates from the ideal beampattern. In order to diminish the impact of quantization, we do a greedy sequential search to obtain beamsteering phases ψ^* for the array C^*.

$$\begin{aligned} \min_{\boldsymbol{\psi}\in[0,2\pi]} \quad & f(\boldsymbol{C^*}, \boldsymbol{\psi}) \\ \text{subject to} \quad & \texttt{MSLL}(\boldsymbol{C^*}, \boldsymbol{\psi}) \leq -10, \\ & \texttt{G}_D(\boldsymbol{C^*}, \boldsymbol{\psi}) \geq \texttt{G}_{D,max} - 3 \end{aligned} \tag{4}$$

A. Subarray Placement Algorithm

A naive placement approach is to sequentially place subarrays over a grid by searching for the position which minimizes the overall cost function at each step. While this approach has low complexity, it severely restricts the solution space (the N^{th} Subarray position is constrained by the past $N-1$ subarray positions), and we have found it to yield suboptimal results. We therefore adopt a combinatorial search to improve the sampling of the solution space.

The number of possible locations for subarray placement depends on the form factor constraint and the size of the subarray module. However, an exhaustive search is combinatorially explosive: for example, the number of configurations for a discrete grid of size 20×20 is of the order 10^{20}. We therefore employ a heuristic approach to reduce the search space to a reasonable size.

We construct a prefix tree dictionary to find feasible solutions using a breadth first search based enumeration technique. Each node in the tree stores a subarray center position, and the path from root to a node at the n^{th} layer of prefix tree represents a unique configuration of n subarrays. We constrain the subarray centers to lie on a fixed set of discrete grid points G. The algorithm is described in Algorithm 1. We briefly discuss the key steps below.

- INITIALIZE: We start with some initial seed points C_i^1 as the root of the prefix tree, in order to allow sufficient exploration. For example, circular configurations cannot be obtained if a subarray is fixed at the center. A possible alternative is to expand the grid size, but this approach is computationally expensive. We avoid this by multiple initializations of the root node with different locations on the grid.
- LIST: Define operator $\mathcal{T} : G^{|C_i^n|} \rightarrow G^{\kappa}$ which maps the set of subarrays centers $C_i^n \in \mathcal{C}^n$ to the set of κ gridpoints in G covered (blocked for placement of subarray) by n subarray modules placed at C_i^n. This operator also accounts for additional surface area occupied by the subarray PCB apart from the physical antenna elements. The gridpoints available for placement of next subarray is obtained by subtracting from G.
- APPEND: The $(n+1)^{th}$ subarray configuration is constructed from the vacancies, $C_i^n \times \boldsymbol{\nu}_i$ where $\times$ denotes cartesian product of sets. A temporary dictionary $\hat{\mathcal{C}}$ is formed by inserting $|\boldsymbol{\nu}_i| = |G| - \kappa$ child nodes for each node in the n^{th} layer.
- PRUNE: Nodes corresponding to "similar" configurations are deleted based on their subarray center covariance matrix.
  1) Find eigenvalues (λ_1, λ_2) of the subarray center covariance matrix, $\Sigma(C)$.
  2) Enumerate unique configurations by binning all (λ_1, λ_2) pairs over a rectangular grid with resolution τ and randomly picking a few configurations from each bin.

  While pruning is necessarily suboptimal, the choice of covariance as pruning mechanism is motivated by the Cramer-Rao bound on the accuracy of DoA estimation, which is inversely proportional to the covariance of the element positions (see (5)).

Algorithm 1 Prefix Tree Dictionary Search

1: INITIALIZE: $\mathcal{C}^1 = \left\{ C_i^{(n=1)} \right\}; i = [1, N_{init}]$
2: **while** $n < N_{Sub}$ **do**
3: LIST all vacant Gridpoints $\boldsymbol{\nu}_i = G - \mathcal{T}(C_i^n); i \in \{1, |\mathcal{C}^n|\}$
4: APPEND subarray at vacancies $\boldsymbol{\nu}_i$, $\hat{\mathcal{C}} = \bigcup_{i=1}^{|\mathcal{C}^n|} C_i^n \times \boldsymbol{\nu}_i$
5: PRUNE: $\mathcal{C}^{n+1} \leftarrow$ Prune $\left(\hat{\mathcal{C}}\right)$
6: $n = n + 1$
7: **end while**
8: Return $\mathcal{C}^{N_{sub}}$

B. Sequential Phase Optimization

After optimizing the position of elements we perform a greedy sequential search to obtain beamsteering phases $\boldsymbol{\psi}$. An initial feasible set of beamsteering weights is set by rounding the ideal weights to nearest quantized phase value. We then do a local search around this solution by sequentially checking for improvement in cost function by switching steering phase at each element. Repeating this operation over multiple rounds with randomized order of search across elements in each round improves the probability of finding optimal solution.

IV. Results

Subarray Placement: We use Algorithm 1 with a discrete grid, G of size 100×100 and pruning resolution $\tau = 0.05$ to construct a dictionary with $|\mathcal{C}| \sim 2 \times 10^5$ atoms. For our running example (eight 4×4 subarrays placed in an aperture bounded to 10 cm by 10 cm), we obtain four sample array configurations by tuning the weights in 3

1) Triangular: Equally weighting of beamwidth, MSLL and eccentricity ($\alpha = 0.5, \beta = 0.5, \gamma = 0.5$).
2) Circular: Higher emphasis on beamwidth and eccentricity to achieve sharpest beam ($\alpha = 1, \beta = 0.5, \gamma = 1$).
3) Quasi-linear: Only minimize maximum sidelobe level ($\alpha = 0, \beta = 1, \gamma = 0$).
4) Benchmark: Uniform, compact arrangement of subarrays (over an area smaller than the maximum aperture size).

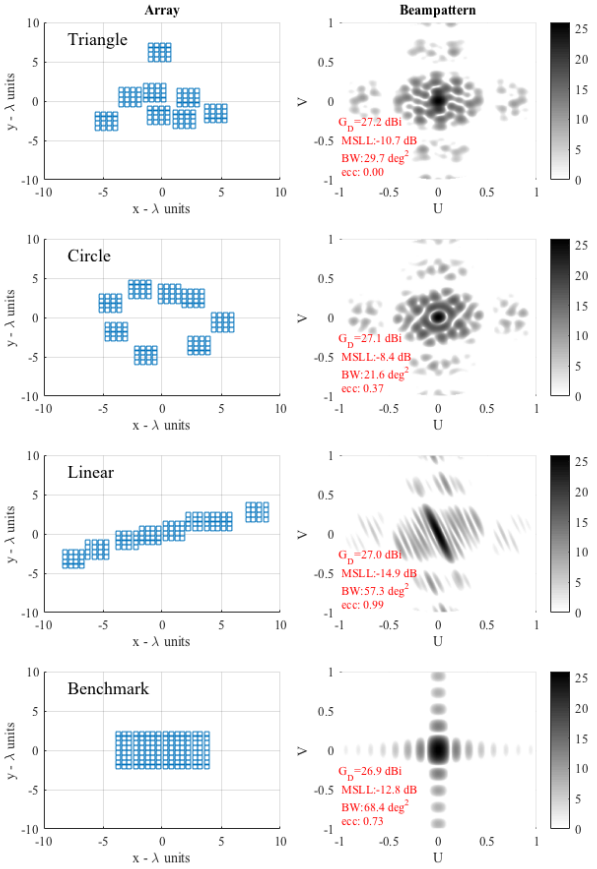

Fig. 3. Sample Subarray configurations & Beampatterns and Costs.

Fig. 3 shows these four sample configurations and their beam patterns while the beam is pointed towards the broadside. The triangular configuration is promising in terms of balancing the various tradeoffs: MSLL=-10.9 dB below mainlobe, $G_D = 27$ dBi and BW=30 deg^2.

Beamsteering Weights: We apply the optimization approach in III-B to the triangular configuration with 3-bit phase quantization $w_i \in \left\{0, \frac{\pi}{4}, \ldots, \frac{7\pi}{4}\right\}; i \in [1, 128]$. Fig. 4 shows the minimization of cost function in 4 for a randomly chosen steering direction $\boldsymbol{u_0} = (0.3, 0.1)$. The search procedure converges to a locally optimal solution for the overall cost function in a few rounds.

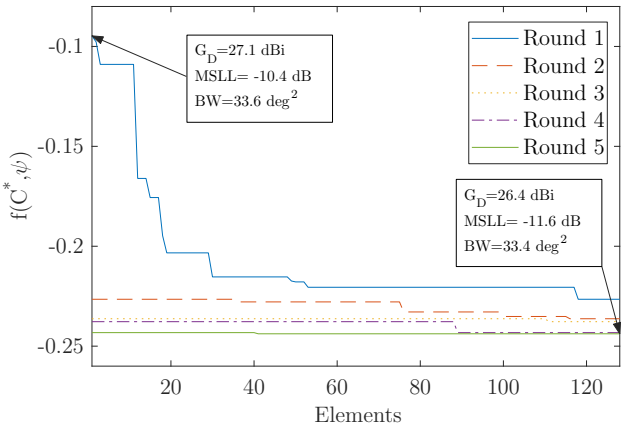

Fig. 4. Sequential search for Phase optimization, $\boldsymbol{u_0} = (0.3, 0.1)$.

V. Comparison via DoA Estimation

In order to compare the array configuration obtained from our algorithm, we analyze their 2D direction of arrival (DoA) estimation performance. Consider a deterministic AWGN model for the DoA estimation problem, the received signal from direction $\boldsymbol{u} \triangleq [u, v]^T$ is given by:

$$\boldsymbol{X} = \alpha \boldsymbol{S}(\boldsymbol{u}) + \boldsymbol{N}$$

where $\boldsymbol{S}(\boldsymbol{u}) = \left[e^{jk\boldsymbol{u}^T\boldsymbol{d_1}}, \ldots, e^{jk\boldsymbol{u}^T\boldsymbol{d}_M}\right]^T$ is the signal vector, $\boldsymbol{N} = [n_1, \ldots, n_M]^T$ is complex white noise such that $\mathbb{E}(NN^H) = \sigma^2 I_M$, and α is an unknown complex gain. The joint maximum likelihood estimator of $\boldsymbol{u}$ and α yields a noncoherent estimator for $\boldsymbol{u}$ as follows:

$$\boldsymbol{u}_{ML} = arg\max_{\boldsymbol{u}} \left|\boldsymbol{S}(\boldsymbol{u})^{\boldsymbol{H}}\boldsymbol{X}\right|^2$$

The covariance of the estimation error is defined as

$$\boldsymbol{R}_\epsilon = \mathbb{E}\left(\boldsymbol{u} - \boldsymbol{u_{ML}}\right)\left(\boldsymbol{u} - \boldsymbol{u_{ML}}\right)^T)$$

Setting $\alpha = 1$ to focus on performance of DoA estimation, the Bayesian Cramér Rao Bound in our setting is given by

$$CRB = (\boldsymbol{J_F} + \boldsymbol{J_P})^{-1}$$

where, $\boldsymbol{J_F}, \boldsymbol{J_P}$ denote the Fisher Information Matrix contributions from the observation and the prior distribution of DoA respectively.

$$(J_F)_{ij} = -\mathbb{E}_{\boldsymbol{X},\boldsymbol{u}}\left[\frac{\partial^2 p(\boldsymbol{X}|\boldsymbol{u})}{\partial u_i \partial u_j}\right], (J_P)_{ij} = -\mathbb{E}_{\boldsymbol{u}}\left[\frac{\partial^2 p(\boldsymbol{u})}{\partial u_i \partial u_j}\right]$$

For our underlying model, $\boldsymbol{J_F}$ is given by[10]:

$$\boldsymbol{J_F} = \frac{2k^2}{\sigma^2}\boldsymbol{D}^T\left[\boldsymbol{I}_M - \frac{1}{M}\mathbb{1}_M\mathbb{1}_M^T\right]\boldsymbol{D} \tag{5}$$

Assuming the DoA prior to be uniformly distributed in the ROI ($\theta \leq 30°$), the prior FIM simplifies to $\boldsymbol{J_P} = 1.343\boldsymbol{I}_2$.

The error covariance matrix $\boldsymbol{R}_\epsilon$ can be geometrically interpreted by its (1) Trace $\sqrt{tr(R_\epsilon)}$, which is the overall expected MSE and (2) Determinant $\sqrt{|R_\epsilon|}$ which is area of the error ellipse. Fig. 5 show the MLE and Cramér Rao Bound on Root Mean Square Error (RMSE) for sample configurations.

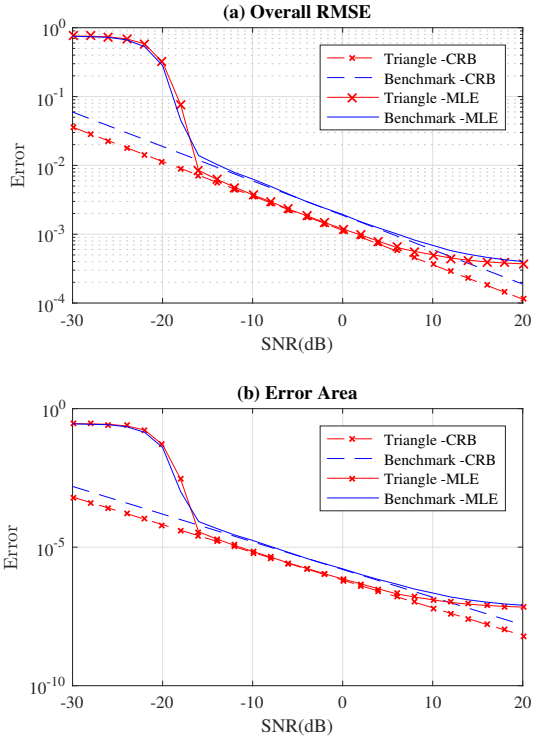

Fig. 5. Error Covariance Matrix measures.

Both measures show similar patterns. Since the Cramér Rao Bound is a local bound it only accounts for estimation errors due to near ambiguities in DoA estimation, thus it depends on the curvature or sharpness of the beam at the mainlobe. The sparse array has a tighter main lobe due to low beamwidth which improves estimation performance. At higher SNR, the MLE tends to approach the Cramér Rao Bound as the estimation errors are mostly due to near ambiguities but diverges at low SNR because estimation errors due to far-ambiguities from the side-lobes dominate the MSE. The threshold point for both configurations occurs at same SNR, indicating that our design for the sparse array is successful in controlling large sidelobes. Note that the divergence from CRB at high SNR is due to limitation on resolution of $\boldsymbol{u}$ in our simulation.

VI. Conclusions

The results in this paper demonstrate that, with appropriate design, sparse arrays of subarrays can provide a tight main beam while controlling the sidelobes and grating lobes generated by the large spacing (relative to wavelength) between subarrays. While the beamforming regime which we explore here differs from prior work [7] on point-to-point spatial multiplexing using arrays of subarrays (the latter requires larger form factors at moderately large ranges), it is of interest to explore point-to-multipoint settings in our present framework. For communication and radar applications, we conjecture that sparsely populated large apertures should be more effective than compact configurations in smaller apertures for separating out widely separated transceivers and targets, respectively, assuming that grating and side lobes are appropriately controlled. Finally, detailed algorithm design and system-level evaluations using sparse arrays, for both communication and radar applications, are an important topic for future work.

Acknowledgments

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA, and by the National Science Foundation under grants CNS-1518812 and CNS-1518632.

References

[1] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Transactions on antennas and propagation*, vol. 16, no. 2, pp. 172–175, 1968.

[2] T. Birinci and Y. Tanık, "Optimization of nonuniform planar array geometry for direction of arrival estimation," in *Signal Processing Conference, 2005 13th European*. IEEE, 2005, pp. 1–4.

[3] V. Roy, S. P. Chepuri, and G. Leus, "Sparsity-enforcing sensor selection for doa estimation," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE 5th International Workshop on*. IEEE, 2013, pp. 340–343.

[4] F. Athley, C. Engdahl, and P. Sunnergren, "On radar detection and direction finding using sparse arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, 2007.

[5] P. Pal and P. P. Vaidyanathan, "Nested arrays in two dimensions, part i: Geometrical considerations," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4694–4705, Sept 2012.

[6] E. Keto, "Hierarchical configurations for cross-correlation interferometers with many elements," *Journal of Astronomical Instrumentation*, vol. 1, no. 01, p. 1250007, 2012.

[7] E. Torkildson, U. Madhow, and M. Rodwell, "Indoor millimeter wave mimo: Feasibility and performance," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4150–4160, 2011.

[8] M. J. Lee, L. Song, S. Yoon, and S. R. Park, "Evaluation of directivity for planar antenna arrays," *IEEE Antennas and Propagation Magazine*, vol. 42, no. 3, pp. 64–67, 2000.

[9] H. Lebret and S. Boyd, "Antenna array pattern synthesis via convex optimization," *IEEE transactions on signal processing*, vol. 45, no. 3, pp. 526–532, 1997.

[10] O. Lange and B. Yang, "Antenna geometry optimization for 2d direction-of-arrival estimation for radar imaging," in *Smart Antennas (WSA), 2011 International ITG Workshop on*. IEEE, 2011, pp. 1–8.