

Optimum Beamforming and Grating Lobe Mitigation for Intelligent Reflecting Surfaces

Sai Sanjay Narayanan, Uday K Khankhoje, *Senior Member, IEEE*, Radha Krishna Ganti *Member, IEEE*

Abstract—Ensuring adequate wireless coverage in upcoming communication technologies such as 6G is expected to be challenging. This is because user demands of higher data rate require an increase in carrier frequencies, which in turn reduce the diffraction effects (and hence coverage) in complex multipath environments. Intelligent reflecting surfaces have been proposed as a way of restoring coverage by adaptively reflecting incoming electromagnetic waves in desired directions. This is accomplished by judiciously adding extra phases at different points on the surface. In practice, these extra phases are only available in discrete quantities due to hardware constraints. Computing these extra phases is computationally challenging when they can only be picked from a discrete distribution, and existing approaches for solving this problem were either heuristic or based on evolutionary algorithms. We solve this problem by proposing fast algorithms with provably optimal solutions. Our algorithms have linear complexity, and are presented with rigorous proofs for their optimality. We show that the proposed algorithms exhibit better performance. We analyze situations when unwanted grating lobes arise in the radiation pattern, and discuss mitigation strategies, such as the use of triangular lattices and prephasing techniques, to eliminate them. We also demonstrate how our algorithms can leverage these techniques to deliver optimum beamforming solutions.

Index Terms—Wireless Networks, Surface Reflectance, Intelligent Reflecting Surface, Passive Beamforming, Phase Shift, optimization, beamforming, grating lobe mitigation

I. INTRODUCTION

Intelligent reflecting surface (IRS) is proposed as a key technology to enable the advancement of 6G wireless communications, which demands high data rates and a large spectrum of frequencies to operate [1]. IRS is a planar surface consisting of many sub-wavelength passive reflecting elements, where each element is equipped with a controller that decides the phase shift it can impart to the incoming beam. By carefully tuning these phase shifts, one can make the reflected beams from each element constructively interfere in a given direction, and destructively interfere in other directions. On account of being a passive surface, an IRS exhibits energy efficiency and improves the spectral efficiency of the system [2].

The required phase shifts of the elements are obtained by solving an appropriate beamforming optimization problem.

Corresponding author: Uday Khankhoje (e-mail: uday@ee.iitm.ac.in).

The authors are with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India

Authors acknowledge the following research grants: (i) “6G: Sub-THz Wireless Communication with Intelligent Reflecting Surfaces (IRS),” R-23011/3/2022-CC&BT-MeitY by the Ministry of Electronics and Information Technology (MeitY), Government of India, and (ii) “Intelligent approaches to designing and building Intelligent Reflecting Surfaces for 6G applications,” by the Qualcomm 6G University Research India Program.

Several beamforming problems have been posed in the IRS literature [3]. Minimum variance beamforming (MVB) involves selecting the weights so as to minimize the variance of the output signal’s energy in a given direction while maintaining unity gain in that direction [4]–[6]. In transmit beamforming and broadcast beamforming, the goal is to minimize transmit power while maintaining a signal-to-noise ratio (SNR) threshold for each user [7]–[9]. The above beamforming problems are typically framed as continuous optimization problems, where the weights (and hence the phase shifts) belong to a continuous set of values. However, continuous phase shifters are difficult to implement on the IRS due to the limited size of each element and associated hardware complexity. In practice, each element has a discrete set of phase shifts (e.g., 1-bit or 2-bit), which gives rise to discrete beamforming problems [10]–[12]. While 1-bit phase shifters in particular are easy to implement, the price to pay is the emergence of quantization lobes, which we refer to as the grating lobes induced due to the 1-bit nature of each weight. A common technique to mitigate these lobes (called prephasing) is to perturb the baseline phases of some or all elements by some predetermined strategy [13]–[17], an idea that has evolved from mitigation strategies in phased antenna arrays [18] and reconfigurable reflectarray antennas [19].

Discrete optimization problems are generally hard to solve efficiently. However, for certain classes of problems there exist techniques that yield approximate solutions. In the context of beamforming, two commonly used methods for solving discrete optimization problems are semi-definite relaxation (SDR), which involves relaxing the problem into a convex optimization problem that can be efficiently solved using interior point methods [20], [21], and quantization, which involves solving the continuous version of the problem and then discretizing the solution. Quantization in particular has been used and studied extensively [19], [22]–[25]. The problem with these methods is that SDR has a high computational complexity, while quantization is only a heuristic method that does not provide any mathematical guarantees of optimality. The beamforming problem is posed as a mixed integer non-linear program in [26] and an optimal algorithm is proposed; however, this has exponential worst-case complexity due to which a suboptimal relaxation is proposed via a successive refinement algorithm. Machine learning [27] and quantum algorithms [28] have also been explored for IRS beamforming. A recent review [29] comprehensively covers the various strategies employed in the literature for continuous and discrete variable beamforming as well as related issues such as channel estimation, transmission design, and localization.

More recently, a linear complexity algorithm was proposed [30], [31] for solving the discrete beamforming problem of maximizing the power of the reflected signal in a given direction when a plane wave impinges on the IRS; while optimal, it assumes the IRS elements to have unit reflectivity and the phase shifts to be equiangular. Our work lifts the above assumptions and proposes a series of new algorithms that are both, provably optimal, as well as linear in computational complexity. Building on our previous work [32], we first present a novel $\mathcal{O}(n)$ complexity algorithm for the 1-bit case where all phase shifts of the individual elements are either 0° or 180° (relatively). This is already a vast improvement over heuristic techniques or evolutionary algorithms. The 1-bit algorithm is then generalized to the case where each weight is constrained to its own *arbitrary* discrete two-element set, i.e. the elements of this set need not have unit modulus, nor a phase difference of 180° . We further generalize the algorithm to the case where the weights are constrained to a k -element set ($k > 2$), and the generalization is once again proven to be efficient and optimal. Finally, we propose yet another generalization for the case where beamforming in multiple directions is desired. To the best of our knowledge, ours is the first provably optimal, linear complexity algorithm to solve the beamforming problem under the constraint of arbitrary discrete weights. Using our algorithms to obtain the desired weights, we perform theoretical analyses and simulations for the behavior of sidelobe levels (SLL) for the 1-bit case. We identify the emergence of grating lobes due to a symmetry in the IRS array factor. To this end, we propose two mitigation strategies. In the first approach we propose the use of a triangular lattice to realize the IRS and show supporting numerical results. In the second approach, we demonstrate how our algorithm can be used to incorporate the various grating-lobe mitigation strategies proposed in earlier works.

The paper is structured as follows. In Section II, we formulate the problem statement and present the novel beamforming algorithm along with its generalization to the k -element and multiple beams case. In Section III, we present numerical results to establish the superiority of our approach over heuristic approaches. Using our generalized algorithm, we study the trade-off between the number of bits of allowed phase shifts and beamforming error. In Section IV, we present theoretical analyses and detailed results pertaining to sidelobe levels and identify the issue of grating lobes. This is followed by a discussion on grating lobe mitigation strategies and their integration with our algorithm. Finally, we conclude in Section V.

II. OPTIMUM BEAMFORMING ALGORITHM & GENERALIZATIONS

A. Problem Statement

The IRS is modeled as an $M \times N$ rectangular grid of equi-spaced unit cells in the $x - y$ plane. It is illuminated by an incident electromagnetic plane wave whose coordinates according to the forward scatter alignment (FSA) convention [33] are given by (θ_{in}, ϕ_{in}) , as seen in Fig. (1). As per the FSA convention, the incident beam wavevector is given by

$\hat{k}_{in} = \sin \theta_{in} \cos \phi_{in} \hat{x} + \sin \theta_{in} \sin \phi_{in} \hat{y} - \cos \theta_{in} \hat{z}$, while the reflected wavevector in the direction (θ_0, ϕ_0) is given by $\hat{k}_0 = \sin \theta_0 \cos \phi_0 \hat{x} + \sin \theta_0 \sin \phi_0 \hat{y} + \cos \theta_0 \hat{z}$ (note the difference in the sign of the \hat{z} component between incident and reflected wavevectors). Each cell imparts a complex valued reflection coefficient (called the cell *weight*), $w_{m,n}$, to the incident beam. In order to capture realistic IRS implementations we assume that the weights take on discrete values, and therefore are modeled as belonging to a discrete alphabet, \mathcal{S} (for e.g., in many practical implementations, $\mathcal{S} = \{-1, 1\}$).

The normalized array factor of the IRS, G , is given by:

$$G(\theta, \phi) = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N w_{m,n} e^{j\varphi_{m,n}(\theta, \phi)}, \quad (1)$$

where

$$\varphi_{m,n}(\theta, \phi) = \frac{2\pi d}{\lambda} \left(-m \sin \theta \cos \phi - n \sin \theta \sin \phi + m \sin \theta_{in} \cos \phi_{in} + n \sin \theta_{in} \sin \phi_{in} \right), \quad (2)$$

with d as the inter-cell spacing, and λ being the wavelength of the incident wave. The array factor gives the far-zone electric field of the array when multiplied by the radiation pattern of the constitutive array cell [34] (ignoring mutual coupling).

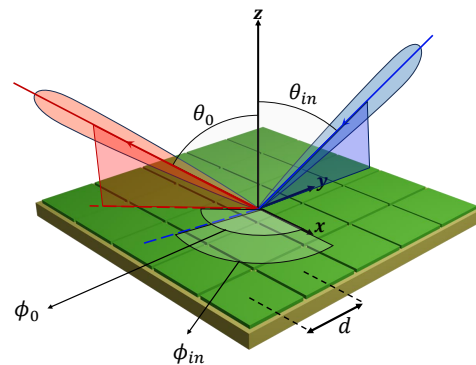


Fig. 1. A schematic diagram of a 5×5 IRS, with the incident beam wavevector (FSA coordinates (θ_{in}, ϕ_{in})) represented in blue and the reflected beam wavevector (FSA coordinates (θ_0, ϕ_0)) represented in red.

We now formally state the problem of IRS beam forming in a certain direction (θ_0, ϕ_0) as:

$$\begin{aligned} \max_{w_{m,n}} & |G(\theta_0, \phi_0)|^2, \\ \text{s.t.} & w_{m,n} \in \mathcal{S}, \quad \forall m, n. \end{aligned} \quad (3)$$

This is a discrete optimization problem since the weights $w_{m,n}$ belong to a finite set. We note that the problem of beamforming can take on many more sophisticated forms as compared to (3) above (see [3] for an extensive review); in this work we focus on the simpler, single beam, noise-free analysis.

B. Thresholding Method

Thresholding (also known as quantization) is the most common discrete beamforming algorithm used in the literature. Consider the case where the allowed values for each cell are: $\mathcal{S} = \{1, -1\}$, i.e., a 1-bit IRS, which corresponds to each unit cell having phase shift of 0° or 180° .

In the thresholding method, the main idea is to first obtain the optimal solution to the continuous version of the problem (i.e., solving (3) with the constraint as $w_{m,n} \in \mathbb{C}$, $|w_{m,n}| = 1$), as done in [35, Thm. 1.2.5]. The solution is unique up to multiplication by a unit-magnitude complex number. Once the continuous solution is obtained, we discretize the weights by mapping them to whichever among $\{1, -1\}$ they are closest to. The complete methodology is presented in Algorithm 1.

Algorithm 1 Thresholding Algorithm

Require: $\theta_{in}, \phi_{in}, \theta_0, \phi_0$

- 1: **for** $1 \leq m \leq M$, $1 \leq n \leq N$ **do**
- 2: **Compute** $w_{m,n} = \exp(-j\varphi_{m,n}(\theta_0, \phi_0))$
- 3: **if** $\angle w_{m,n} \in [-\frac{\pi}{2}, \frac{\pi}{2})$ **then**
- 4: $\tilde{w}_{m,n} \leftarrow 1$
- 5: **else**
- 6: $\tilde{w}_{m,n} \leftarrow -1$
- 7: **end if**
- 8: **end for**
- 9: **return** discrete weights $\tilde{w}_{m,n}$

The thresholding method is popular on account of being intuitive and easy to implement [19], [25]. However, due to its heuristic nature it does not have any guarantees of optimality. Further, it is not clear how to generalize the algorithm if the elements of \mathcal{S} have unequal amplitude. These observations pave the way for our proposed method, which we present next.

C. Optimal Partitioning Algorithm

The proposed method is termed the Optimal Partitioning Algorithm (OPA), which is based on the following intuition. The sum of complex exponentials in (1) can be re-written as $G = \sum_i w_i z_i$, where the index i runs over all tuples (m, n) , w_i represents the weight of the i^{th} unit cell, and $z_i = e^{j\varphi_i(\theta, \phi)}$. To build an intuitive understanding of the algorithm about to be presented, we consider the following two examples.

(i) Consider the z_i 's as shown in Fig. 2(a). To maximize $|G|$ as per (3), the weight w_i associated with each z_i must be chosen appropriately. Since the w_i s are restricted to $\{-1, 1\}$, it follows that $|G|$ is maximized when all the z_i s in the left-half of the plane are assigned $w_i = -1$ (indicated via the red arrows in the Figure) and those in the right-half plane are assigned $w_i = 1$ (or vice-versa). In effect, we have found a partitioning line (in this case the y -axis) that separates the assignment of the weights. The resulting weight assignment is equivalent to the thresholding algorithm.

(ii) Next consider the z_i 's as shown in as shown in Fig. 2(b). As can be seen, the y -axis no longer partitions the weights for maximizing $|G|$; instead, a different weight assignment gives the optimal sum, as shown by the red arrow in the Figure. Thus, the thresholding algorithm would have given a suboptimal weight assignment.

We note that in both examples, the optimal assignment can be conceived of as the result of an optimal partitioning line as indicated by the dashed blue lines in Fig. 2.

The above intuition of an optimal partitioning is formalized via the following theorem.

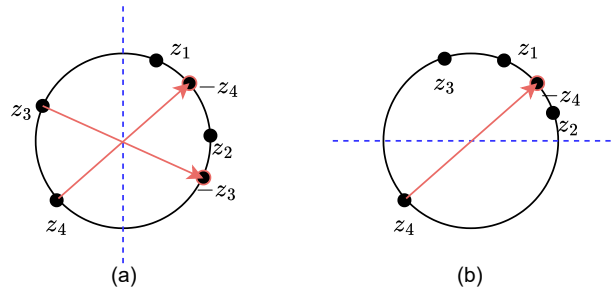


Fig. 2. Complex argand plane with z_i s on the unit circle; in (a) the assignment of weights -1 to z_3, z_4 maximizes $|G|$, while in (b) the assignment of weight -1 to z_4 maximizes $|G|$. The dashed lines indicate the optimal partitioning for weight assignments in each case.

Theorem 1. Let z_1, \dots, z_n be arbitrary non-zero complex numbers, and let $\tilde{w}_1, \dots, \tilde{w}_n$ be an optimal solution to the optimization problem

$$\begin{aligned} \max_{w_i} \quad & |w_1 z_1 + \dots + w_n z_n|, \\ \text{s.t.} \quad & w_i \in \{1, -1\}, \quad \forall i = 1, \dots, n. \end{aligned} \quad (4)$$

Then, there exists a line ℓ passing through the origin with the following properties:

- 1) ℓ does not contain any z_i .
- 2) $\tilde{w}_i = 1$ for all z_i lying on one side of the line, and $\tilde{w}_i = -1$ for all z_i on the other side.

Proof. The proof is presented in Appendix A. □

The above theorem asserts that if an optimal set of weights exists, then there exists a *separating line* passing through the origin, which partitions the set $\{z_1, \dots, z_n\}$ into two sets \mathcal{A} and \mathcal{B} , where $\tilde{w}_i = 1$ for each $z_i \in \mathcal{A}$, and $\tilde{w}_i = -1$ for each $z_i \in \mathcal{B}$. However, we know that an optimal solution exists because the feasible set is finite. Therefore, an optimal separating line definitely exists, and our goal should be to find one such line. Once the partition $\{\mathcal{A}, \mathcal{B}\}$ is obtained, the assignment of 1 and -1 to each set is arbitrary, as replacing \tilde{w}_i with $-\tilde{w}_i$ for each $i = 1, \dots, n$ will yield the same value of $|\sum_i w_i z_i|$.

The natural question that arises is: how many partitions should we check? We must check all partitions that arise due to separating lines passing through the origin. However, multiple separating lines can lead to the same partition, as shown in Fig. 3. In particular, one has the freedom to rotate a separating line anticlockwise about the origin without changing the partition, until it hits some z_i . Therefore, the partition obtained by *any* separating line is equivalent to one obtained by a separating line whose angle with the x axis is infinitesimally lesser than the argument of some z_i , and it can be easily seen that one of the subsets of the partition must be of the form $\{z_j \mid \arg z_i \leq \arg z_j < \arg z_i + \pi\}$ for some i (see Fig. 3). Since there are n complex numbers z_i , only n such separating lines (and therefore only n such partitions) need to be investigated. So, our optimal partitioning seeking algorithm has linear time complexity.

Theorem 1 can now be used to formulate the Optimal Partitioning Algorithm (OPA), which computes the optimal

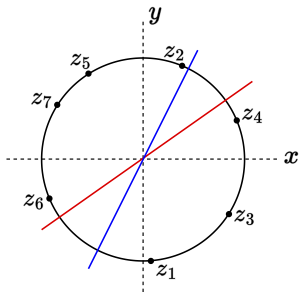


Fig. 3. An example where $n = 7$ and z_1, \dots, z_7 all lie on the unit circle. The red and blue separating lines both give rise to the same partitions, $\mathcal{A} = \{z_1, z_3, z_4\}$ and $\mathcal{B} = \{z_2, z_5, z_6, z_7\}$. Further, the set $\{z_j \mid \arg z_2 \leq \arg z_j < \arg z_2 + \pi\}$ is precisely \mathcal{B} . The blue line can be thought of as having angle infinitesimally lesser than $\arg z_2$.

weights for the optimization problem in (3) with $\mathcal{S} = \{-1, 1\}$. The algorithm is as follows: first, compute the $\{z_i\}$ using (2). Then, choose any arbitrary line passing through the origin, and obtain all possible partitions of $\{z_i\}$ by rotating the line anticlockwise about the origin. When rotating the line, the partition can only change when the line crosses some z_i , and hence we will obtain at most $n = MN$ unique partitions. Next, for each partition, assign $w_i = 1$ for all z_i on one side of the partition, and $w_i = -1$ for all z_i on the other side. Compute $|\sum_i w_i z_i|$ for each partition, and choose the partition that corresponds to the largest value of $|\sum_i w_i z_i|$. Theorem 1 guarantees that the weight assignment corresponding to the resulting partition is optimal.

The complete method is displayed in Algorithm 2; since we only need to check $n = MN$ partitions, OPA will find the optimal weights in time $\mathcal{O}(MN)$. The optimality of the obtained solution is preserved up to a change of sign, *i.e.*, if $\{\tilde{w}_i\}$ are optimal, then so are $\{-\tilde{w}_i\}$, since both sets give the same value of $|G|$.

Algorithm 2 Optimal Partitioning Algorithm (OPA)

Require: Non-zero complex numbers z_1, \dots, z_n

- 1: $\max \leftarrow 0$
 - 2: **for** $i = 1$ to n **do**
 - 3: Compute $\delta_i = \arg z_i$.
 - 4: Set $w_j = 1$ for all j such that $\arg z_j \in [\delta_i, \delta_i + \pi)$ and $w_j = -1$ for others (*corresponds to assigning opposite weights to either side of the partition*)
 - 5: Compute $g = |w_1 z_1 + \dots + w_n z_n|$.
 - 6: **if** $g > \max$ **then**
 - 7: $\max \leftarrow g$
 - 8: optimal weights $\leftarrow [w_1 \ \dots \ w_n]$.
 - 9: **end if**
 - 10: **end for**
 - 11: **return** optimal weights
-

We now discuss a generalization of the above algorithm for the case where each weight is constrained to its own binary set, *i.e.*, $w_i \in \{a_i, b_i\}$ for all $i = 1, \dots, n$. In order to solve the resulting optimization problem, we introduce the new scaled variables y_i , such that $w_i = \left(\frac{a_i + b_i}{2}\right) + y_i \left(\frac{a_i - b_i}{2}\right)$, and rewrite the sum in terms of the variables y_i . Since the constraints

$w_i \in \{a_i, b_i\}$ are mapped to $y_i \in \{1, -1\}$, OPA can be used to obtain the optimal solution \tilde{y}_i , which in turn is used to compute the optimal weights \tilde{w}_i . This generalization is termed as Generalized OPA (gOPA), and the full details are displayed in Algorithm 3.

Algorithm 3 Generalized OPA (gOPA)

Require: Non-zero complex numbers z_1, \dots, z_n , and the constraint sets $\{a_1, b_1\}, \dots, \{a_n, b_n\}$ for each of the weights (where $a_i \neq b_i$ for all i).

- 1: Compute $z'_1, \dots, z'_n, z'_{n+1}$ as:

$$z'_i = \left(\frac{a_i - b_i}{2}\right) z_i \quad \forall i = 1, \dots, n,$$

$$z'_{n+1} = \sum_{i=1}^n \left(\frac{a_i + b_i}{2}\right) z_i.$$

- 2: Apply OPA to the set z'_1, \dots, z'_{n+1} , and get optimal weights $[\tilde{y}_1 \ \dots \ \tilde{y}_n \ \tilde{y}_{n+1}]$.
- 3: **if** $\tilde{y}_{n+1} = -1$ **then**
- 4: **for** $i = 1$ to n **do**
- 5: $\tilde{y}_i \leftarrow -\tilde{y}_i$.
- 6: **end for**
- 7: **end if**
- 8: Compute the optimal weights $\tilde{w}_1, \dots, \tilde{w}_n$ as

$$\tilde{w}_i = \left(\frac{a_i + b_i}{2}\right) + \tilde{y}_i \left(\frac{a_i - b_i}{2}\right).$$

- 9: **return** optimal weights $[\tilde{w}_1 \ \dots \ \tilde{w}_n]$.
-

D. k -element Generalization

We now present yet another generalization of OPA called kOPA, where each weight is constrained to a k -element set with $k > 2$, *i.e.*, $w_{m,n} \in \mathcal{S} = \{a_1, \dots, a_k\}$ for all m, n , and a_1, \dots, a_k are arbitrary complex numbers.

As in the previous case, it turns out that the optimal solution is once again obtained by partitioning the z_i 's into k subsets with a specific geometric structure, where each z_i in a given subset gets the same weight. In OPA, we saw that the optimal partition arises from considering a separating line, which can be thought of as two diametrically opposite rays emerging from the origin. A similar intuition carries forward to kOPA, where the optimal partition arises due to a configuration of k such rays.

Let $\mathcal{S} = \{a_1, \dots, a_k\}$ be the discrete weight set from which the weights $w_{m,n}$ have to be picked. We then define the following regions:

$$A_i = \{z \in \mathbb{C} \mid \operatorname{Re}((a_i - a_j)z) > 0, \forall j \neq i\}. \quad (5)$$

For a given j , the inequality $\operatorname{Re}((a_i - a_j)z) > 0$ represents a half-space in \mathbb{C} (or equivalently \mathbb{R}^2), and hence A_i is an intersection of $k - 1$ half-spaces. Therefore, the collection $\{A_1, \dots, A_k\}$ can be easily seen as a radial partition \mathcal{P} of the complex plane with A_i 's as the cones of the radial partition. Let $\theta \in [0, 2\pi]$, and let $\mathcal{P}(\theta)$ denote the rotation of the partition \mathcal{P} by an angle θ anticlockwise, with the corresponding rotated

cones denoted by $A_i(\theta)$. An illustration of \mathcal{P} and $\mathcal{P}(\theta)$ is provided in Fig 4.

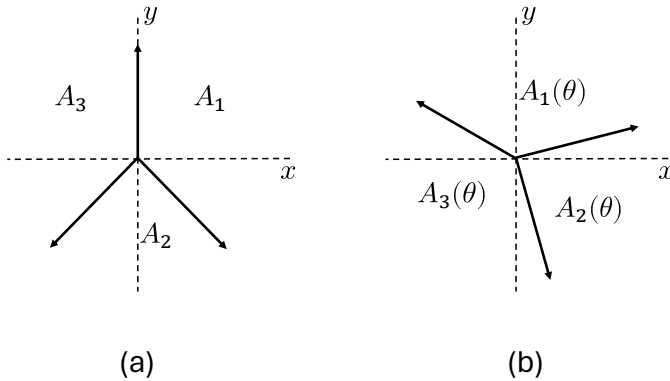


Fig. 4. (a) Depiction of $\mathcal{P} = \{A_1, A_2, A_3\}$ for the constraint set $\mathcal{S} = \{1, j, -1\}$ (b) Rotated partition $\mathcal{P}(\theta) = \{A_1(\theta), A_2(\theta), A_3(\theta)\}$, where $\theta = 60^\circ$.

The idea behind kOPA is that the optimal partition is obtained by a suitable rotation of the partition \mathcal{P} , as described in the following theorem.

Theorem 2. Let z_1, \dots, z_n be arbitrary non-zero complex numbers, and let $\tilde{w}_1, \dots, \tilde{w}_n$ be an optimal solution to the optimization problem

$$\begin{aligned} \max_{w_i} \quad & |w_1 z_1 + \dots + w_n z_n|, \\ \text{s.t.} \quad & w_i \in \mathcal{S} = \{a_1, \dots, a_k\}, \quad i = 1, \dots, n. \end{aligned} \quad (6)$$

Then, there exists $\delta \in [0, 2\pi]$ such that

- 1) None of the z_i lie on the edges of the cones of the rotated partition $\mathcal{P}(\delta)$.
- 2) If z_i belongs to the cone $A_j(\delta)$ in the rotated partition $\mathcal{P}(\delta)$, then $\tilde{w}_i = a_j$.

Proof. The proof is presented in Appendix B. \square

We can now formulate kOPA as follows. Firstly, compute \mathcal{P} . Then, rotate \mathcal{P} by δ (sweeping from 0 to 360°), to obtain various partitions. For each of these partitions, assign weights according to Theorem 2, and compute $|\sum w_i z_i|$. Finally, choose the partition that corresponds to the largest value of $|\sum w_i z_i|$. The details are formalized in Algorithm 4. Theorem 2 guarantees that the chosen weight assignment is optimal. An example of optimal weight assignment is shown in Fig 5.

We make the following observations.

- The number of unique partitions obtained in kOPA is at most nk . This is because when rotating \mathcal{P} , the partition only changes when one of the k edges crosses some z_i , and this can happen in at most nk unique ways. As a consequence, kOPA has time complexity $O(nk)$ (where $n = MN$ in the context of IRS beamforming), making it extremely fast.
- In practical implementations, k is usually equal to 2^b where b denotes the number of bits, and the allowed phase shifts are equispaced by $\frac{2\pi}{k}$, which corresponds to $\mathcal{S} = \{1, e^{j\frac{2\pi}{k}}, \dots, e^{j\frac{2\pi(k-1)}{k}}\}$. In this case, it can be seen that the cones A_1, \dots, A_k , each subtend an angle of $\frac{2\pi}{k}$ at the origin.

Algorithm 4 k-Optimal Partitioning Algorithm (kOPA)

Require: Non-zero complex numbers z_1, \dots, z_n , and the constraint set $\mathcal{S} = \{a_1, a_2, \dots, a_k\}$.

- 1: Compute the radial partition $\mathcal{P} = \{A_1, \dots, A_k\}$ as described.
- 2: $\max \leftarrow 0, \delta \leftarrow 0$.
- 3: **while** $\delta \leq 2\pi$ **do**
- 4: Find the smallest $\psi \in (0, 2\pi]$ such that some z_i lies on one of the edges of the cones of $\mathcal{P}(\delta + \psi)$.
- 5: Rotate the current partition $\mathcal{P}(\delta)$ by some $\psi' \in (0, \psi)$ to obtain $\mathcal{P}(\delta + \psi')$.
- 6: Assign $w_i = a_j$ whenever $z_i \in A_j(\delta + \psi')$, for all i .
- 7: Compute $g = |w_1 z_1 + \dots + w_n z_n|$.
- 8: **if** $g > \max$ **then**
- 9: $\max \leftarrow g$, optimal weights $\leftarrow [w_1 \dots w_n]$.
- 10: **end if**
- 11: $\delta \leftarrow \delta + \psi$.
- 12: **end while**
- 13: **return** optimal weights

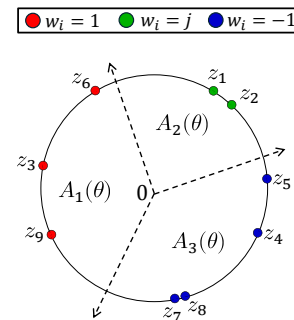


Fig. 5. The optimal weight assignment for a set of 9 randomly generated unit modulus complex numbers z_1, \dots, z_9 (with $\mathcal{S} = \{1, j, -1\}$), which can be seen as arising from a rotated partition $\mathcal{P}(\theta)$.

E. Beamforming in multiple directions

We now discuss the problem of forming simultaneous beams in multiple directions. Suppose $(\theta_1, \phi_1), \dots, (\theta_l, \phi_l)$ are the directions in which we want to beamform, and denote the array factor along these directions as G_1, \dots, G_l , where $G_j = G(\theta_j, \phi_j)$ for all $j \in \{1, \dots, l\}$. Note that we can write $G_j = \sum_i w_i z_i^{(j)}$ where i runs over the tuples (m, n) , and hence each G_j is a function of the weights $\{w_i\}$. The optimization problem that we wish to solve is

$$\begin{aligned} \max_{w_i} \quad & |G_1| + |G_2| + \dots + |G_l|, \\ \text{s.t.} \quad & w_i \in \mathcal{S}, \quad \forall i. \end{aligned} \quad (7)$$

We now construct an equivalent optimization problem as follows. Firstly, for a given set of weights, we have the following due to triangle inequality:

$$|G_1| + |G_2| + \dots + |G_l| \geq |G_1 + \alpha_2 G_2 + \dots + \alpha_l G_l|, \quad (8)$$

where $\alpha_j \in \mathbb{C}$ with $|\alpha_j| = 1$, and equality occurs iff the following co-phasing relation holds: $\arg(G_1) = \arg(\alpha_2 G_2) =$

$\dots = \arg(\alpha_l G_l)$. Hence, we have

$$\begin{aligned} & \max_{w_i} |G_1| + |G_2| + \dots + |G_l| \\ &= \max_{w_i} \max_{|\alpha_j|=1} |G_1 + \alpha_2 G_2 + \dots + \alpha_l G_l| \\ &= \max_{|\alpha_j|=1} \max_{w_i} |G_1 + \alpha_2 G_2 + \dots + \alpha_l G_l|. \end{aligned} \quad (9)$$

The advantage of the equivalent problem is that for a given tuple $(\alpha_2, \dots, \alpha_l)$, the inner maximization problem takes the form $\max_{w_i} |\sum_i w_i y_i|$, because $G_1 + \alpha_2 G_2 + \dots + \alpha_l G_l = \sum_i w_i (z_i^{(1)} + \alpha_2 z_i^{(2)} + \dots + \alpha_l z_i^{(l)})$, and thus it can be solved optimally using OPA. Therefore, our goal is to find the optimal tuple $(\alpha_2, \dots, \alpha_l)$ for the outer maximization.

A naive approach to solve this problem would be to choose a collection \mathcal{D} of tuples $(\alpha_2, \dots, \alpha_l)$, compute $\max_{w_i} |G_1 + \alpha_2 G_2 + \dots + \alpha_l G_l|$ for each tuple in $(\alpha_2, \dots, \alpha_l) \in \mathcal{D}$, and choose the $\{w_i\}$ that corresponds to the largest value.

Instead, we present an iterative approach to solve this problem, based on the following theorem:

Theorem 3. Let $\{\alpha^{(k)}\}_{k \geq 1} = \{(\alpha_2^{(k)}, \dots, \alpha_l^{(k)})\}_{k \geq 1}$ be a sequence of tuples, and $\{w^{(k)}\}_{k \geq 1}$ be a sequence of weights that satisfy the following relations ($k \in \mathbb{N}$ denotes iteration number):

- $w^{(k)} = \arg \max_{w_i} |G_1 + \sum_{j=2}^l \alpha_j^{(k)} G_j|$.
- $\alpha_j^{(k+1)} = e^{j(\arg(G_1) - \arg(G_j))}$ for $j = 2, \dots, l$, where each G_j is computed using the weights $w^{(k)}$.

Let $c_k = \max_{w_i} |G_1 + \sum_{j=2}^l \alpha_j^{(k)} G_j|$, and $d_k = |G_1(w^{(k)})| + \dots + |G_l(w^{(k)})|$. Then, the sequences $\{c_k\}_{k \geq 1}$ and $\{d_k\}_{k \geq 1}$ are both non-decreasing.

Proof. The proof is presented in Appendix D. \square

Thus, performing the iteration steps as specified in the above theorem for each $(\alpha_2, \dots, \alpha_l) \in \mathcal{D}$ will yield better weights than the naive approach. The multiple beamforming method is summarized in Algorithm 5.

Algorithm 5 Multiple Beamforming

Require: Directions $(\theta_1, \phi_1), \dots, (\theta_l, \phi_l)$ in which we desire to beamform, and a discrete collection \mathcal{D} of tuples $(\alpha_2, \dots, \alpha_l)$, where $|\alpha_j| = 1$ for all $j \in \{2, \dots, l\}$.

- 1: **for** each $(\alpha_2, \dots, \alpha_l) \in \mathcal{D}$ **do**
 - 2: Set $(\alpha_2^{(1)}, \dots, \alpha_l^{(1)}) = (\alpha_2, \dots, \alpha_l)$, and $k \leftarrow 1$.
 - 3: **while** convergence is not reached **do**
 - 4: Compute $w^{(k)} = \arg \max_{w_i} |G_1 + \sum_{j=2}^l \alpha_j^{(k)} G_j|$.
 - 5: Set $\alpha_j^{(k+1)} \leftarrow e^{j(\arg(G_1) - \arg(G_j))}$ for $j = 2, \dots, l$, where each G_j is computed using the weights $w^{(k)}$.
 - 6: **end while**
 - 7: Compute $|G_1| + \dots + |G_l|$ using the last computed weights.
 - 8: **end for**
 - 9: **return** weights that corresponds to largest $|G_1| + \dots + |G_l|$.
-

We remark that while the above algorithm is not guaranteed to return the optimal weights, in practice the resulting weights

are found to give very good performance. Typically, a good choice for \mathcal{D} is the collection of all tuples $(\alpha_2, \dots, \alpha_l)$ where each α_j is uniformly sampled on the unit circle. We discuss these issues and numerical results for multiple beamforming in Section III. We note in passing that all the algorithms that have been discussed so far can also be applied for beamforming from a basestation in case it also employs discrete weights; the only modification required is to eliminate the incident field related terms in (2), which is conveniently achieved by setting $\theta_{in} = 0$.

III. RESULTS FOR OPTIMAL BEAMFORMING

In this Section, we provide various simulation results in support of the efficacy of our proposed method. In all simulations, we take the inter-element spacing as $d = \frac{\lambda}{2}$.

A. Optimality of OPA

We first compare the proposed OPA with the existing thresholding method. The hypothesis is that the thresholding method does not always return the optimal weights, and we provide an example that shows exactly that. Consider a 3×3 array, with incident direction $(\theta_{in}, \phi_{in}) = (-45^\circ, 215^\circ)$ and the desired direction $(\theta_0, \phi_0) = (-30^\circ, 35^\circ)$. All weights are constrained to the set $\{1, -1\}$. From the weights returned by the thresholding method, we obtain $|G(\theta_0, \phi_0)|^2 = -3.86$ dB, whereas from the weights obtained by OPA, we obtain $|G(\theta_0, \phi_0)|^2 = -2.95$ dB. Fig. 6 shows the scatter plots the weight assignment for each method. The OPA results are clearly superior, and this example confirms our hypothesis.

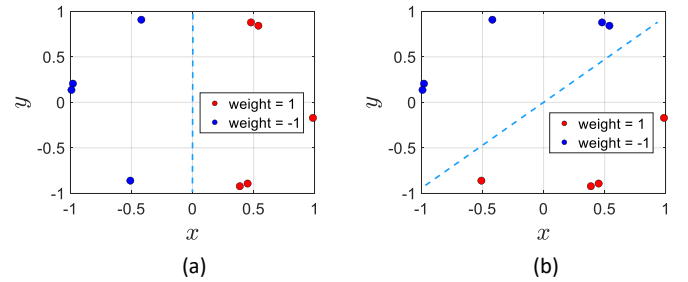


Fig. 6. Scatter plots of $z_{m,n}$ along with assignment of weights due to (a) thresholding method (b) OPA. The dashed lines indicate the separating line in each case.

Next, we establish the optimality of OPA by comparing it with brute force combinatorial search of the discrete weights. For the same incident and desired directions as above, we compare thresholding, OPA and combinatorial search for various array sizes, as shown in Fig. 7. In Fig. 7(a), all weights are constrained to $\{-1, 1\}$, while in Fig. 7(b), the amplitudes and phase difference for the constraint set of each weight are independently and randomly sampled from $U(0.75, 1)$ and $U(160^\circ, 180^\circ)$ respectively, where U denotes the uniform distribution (*i.e.*, for each weight $w_{m,n}$, we have $a_{m,n} \sim U(0.75, 1)$, $|b_{m,n}| \sim U(0.75, 1)$ and $\arg b_{m,n} \sim U(160^\circ, 180^\circ)$), in order to capture realistic scenarios (for e.g. on account of fabrication related imperfections). Note that in case (b), the thresholding algorithm is modified slightly,

such that the continuous solution $w_{m,n}$ is mapped to whichever among $\{a_{m,n}, b_{m,n}\}$ it is closest to, for all m, n . We observe that OPA and combinatorial search yield the same values for $|G|^2$ in both cases, confirming the optimality of OPA. Further, the gap between the thresholding and OPA solutions is on average larger in (b) than in (a), which signifies a preference of OPA over thresholding in realistic scenarios.

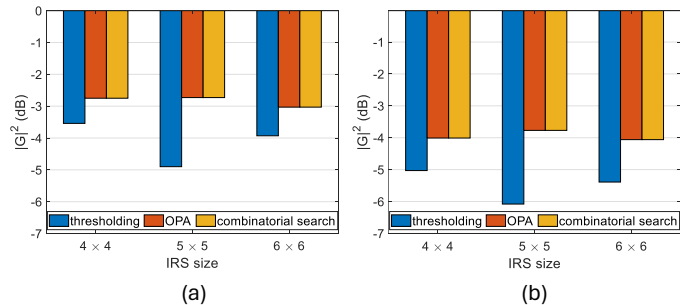


Fig. 7. Comparison of thresholding, OPA and combinatorial search for various array sizes, with constraint sets $\{a_{m,n}, b_{m,n}\}$ equal to (a) $\{1, -1\}$, and (b) $a_{m,n}, |b_{m,n}| \sim U(0.75, 1)$ and $\arg b_{m,n} \sim U(160^\circ, 180^\circ)$, where U denotes uniform distribution

B. Visualization of weights

We now visualize the output of the OPA by plotting the weight distribution across the IRS, and the resulting beam patterns for a couple of sample cases. We fix the array size to be 30×30 , and choose $\theta_{in} = -30^\circ$, $\theta_0 = -15^\circ$ and $\phi_0 = 45^\circ$, $\phi_{in} = 225^\circ$. OPA is run in MATLAB, on a system with Intel Core i7 processor and 16 GB RAM, and the optimal weights are obtained in 0.4 seconds.

As seen in Fig. 8(a), the distribution consists of diagonal strips of 1 and -1 , whose angle with the y axis (vertical axis) is precisely ϕ_{in} . Intuitively, this makes sense because the strips are precisely the locus of points on the IRS that get the same initial phase due to the incident beam.

Next, we consider the case of $\theta_{in} = -45^\circ$, $\theta_0 = -30^\circ$ and $\phi_0 = 0^\circ$, $\phi_{in} = 180^\circ$. As seen in Fig. 8(a), the distribution consists of vertical strips of 1 and -1 as expected. However, in addition to the expected reflected beam at $\theta_0 = -30^\circ$, we also observe a second peak of equal strength around $\theta_0 = -5^\circ$. Understanding the existence of this second peak is of great importance, and we do so in Section IV.

C. Beamforming error vs number of bits

An important consideration in discrete beamforming is the selection of number of bits of allowed phase shifts. While increasing the number of bits increases the quality of beamforming, such as directivity and power in the desired direction, it also makes it significantly harder to implement in hardware. This leads to a natural tradeoff between the number of bits and the quality of beamforming, and is explored in [22], [25]. Of particular interest to us is the effect of number of bits on the *beamforming error* (BE), which we define as the angle between the desired mainlobe direction, (θ_0, ϕ_0) ,

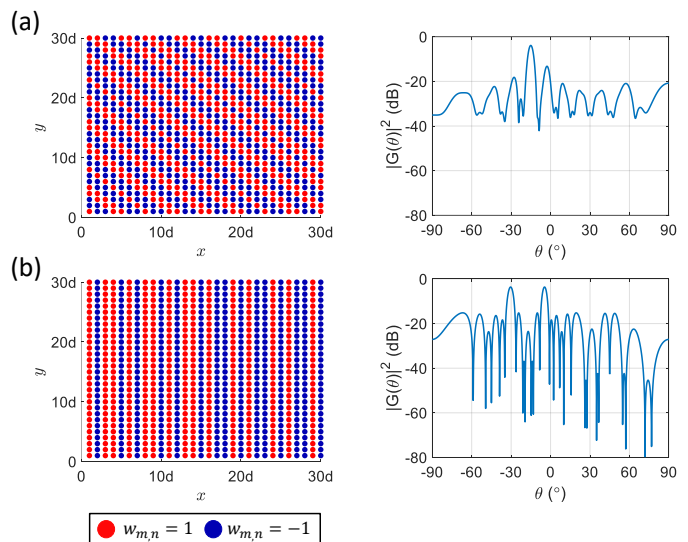


Fig. 8. Distribution of optimal weights on IRS and the corresponding normalized radiation pattern, for (a) $(\theta_{in}, \phi_{in}) = (-30^\circ, 225^\circ)$ and $(\theta_0, \phi_0) = (-15^\circ, 45^\circ)$ (b) $(\theta_{in}, \phi_{in}) = (-45^\circ, 180^\circ)$ and $(\theta_0, \phi_0) = (-30^\circ, 0^\circ)$

and the actual mainlobe direction, (θ'_0, ϕ'_0) , achieved by our algorithms, and is computed using the formula,

$$\cos(\text{BE}) = \sin \theta_0 \sin \theta'_0 \cos(\phi_0 - \phi'_0) + \cos \theta_0 \cos \theta'_0. \quad (10)$$

Fig. 9 shows the variation of beamforming error with the number of bits, for different IRS sizes. We fix $\theta_{in} = 0$ (normal incidence), $\phi_0 = 30^\circ$, and compute the average beamforming error over $\theta_0 \in [-45^\circ, 45^\circ]$ spaced by intervals of 5° . The optimal weights are obtained using kOPA. We observe that the beamforming error decreases if we increase the number of bits and the IRS size. In particular, for 15×15 and 30×30 IRS, we observe negligible beamforming error even with a 1-bit implementation. Additionally, the 3-dB beamwidths and mainlobe gains for various array sizes are shown in Table. I, where we define the gain in a direction (θ, ϕ) as $20 \log(MN|G(\theta, \phi)|)$, with $G(\theta, \phi)$ given by (1). We observe that larger array sizes correspond to larger gains and smaller beamwidth. For example, a 30×30 array exhibits sub- 5° beamwidth. This, when combined with its low beamforming error, makes 30×30 a good practical choice of array size.

TABLE I
3-DB BEAMWIDTH AND MAINLOBE GAIN FOR VARIOUS ARRAY SIZES

	7×7	10×10	15×15	30×30
3-dB beamwidth	15.8°	10.3°	6.8°	3.4°
$20 \log(MN G)$	31	36.4	43.7	55.3

D. Beamforming in multiple directions

We now illustrate an example of multiple beamforming for various array sizes. Fig. 10(a) and 10(b) show the radiation patterns for various array sizes, for the cases where we desire two beams and three beams, respectively. The weights are constrained to $\mathcal{S} = \{1, -1\}$, and are obtained using Algorithm 5. For the two-beam case in Fig. 10(a), we choose $\mathcal{D} = \{e^{j \frac{2\pi k}{30}} | k = 1, \dots, 30\}$, and for the three-beam case in

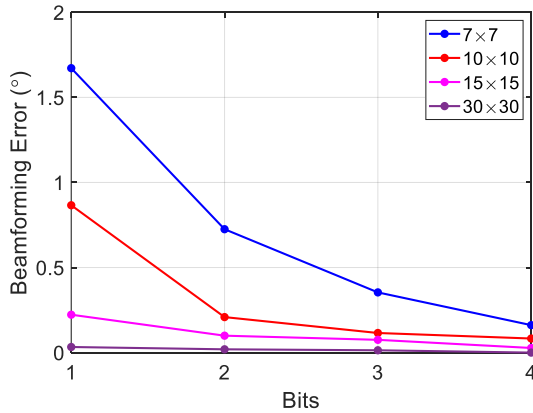


Fig. 9. Variation of beamforming error with the number of bits of allowed phase shifts, under normal incidence with $\phi_0 = 30^\circ$, for 7×7 , 10×10 , 15×15 and 30×30 IRS. The errors are averaged over $\theta_0 \in [-45^\circ, 45^\circ]$ spaced by intervals of 5° .

Fig. 10(b), we choose $\mathcal{D} = \{(e^{j\frac{2\pi k}{30}}, e^{j\frac{2\pi l}{30}}) | k, l = 1, \dots, 30\}$. For each tuple $\alpha \in \mathcal{D}$, the iteration process in Algorithm 5 is observed to converge very quickly (within 10 iterations).

We observe that for smaller array sizes such as 10×10 , the ability to form multiple beams is limited, as evident from the poor sidelobe levels. Additionally, smaller array sizes perform poorly when the desired angles are close to each other. This issue is not faced by larger array sizes such as 30×30 , where we can achieve a sidelobe level of roughly 10 dB as seen in Fig. 10, providing yet another justification to use larger array sizes.

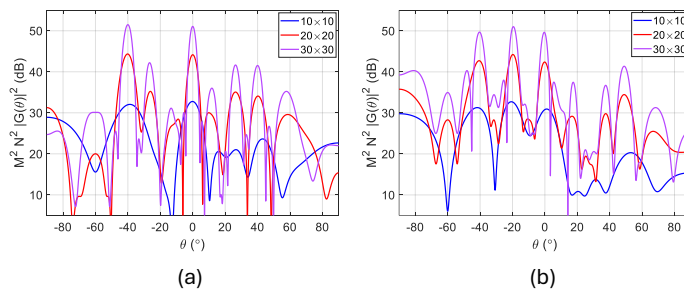


Fig. 10. Radiation patterns illustrating multiple beamforming for various array sizes, with incident direction $(\theta_{in}, \phi_{in}) = (60^\circ, 210^\circ)$, for (a) two beams, where $(\theta_1, \phi_1) = (0^\circ, 30^\circ)$ and $(\theta_2, \phi_2) = (-40^\circ, 30^\circ)$ (b) three beams, where $(\theta_1, \phi_1) = (0^\circ, 30^\circ)$, $(\theta_2, \phi_2) = (-20^\circ, 30^\circ)$ and $(\theta_3, \phi_3) = (-40^\circ, 30^\circ)$.

IV. GRATING LOBE MITIGATION

A. Analysis of grating lobes

We now investigate a critical issue in beamforming concerning the presence of grating lobes in the radiation pattern. A grating lobe is an undesired lobe with the same intensity as the main lobe (*i.e.*, has sidelobe level (SLL) equal to 0 dB). Such a lobe is seen in Fig. 8(b); these lobes can lead to interference in the unwanted directions, and security issues in the context of wireless communications and are undesirable.

We now derive theoretical conditions under which a grating lobe is guaranteed to exist, for the 1-bit case where each

weight is constrained to $\{1, -1\}$. By examining the array factor expression in (1), we can see that given a main lobe at (θ_0, ϕ_0) , a grating lobe will appear at (θ^*, ϕ^*) if the following condition holds for all m, n :

$$\varphi_{m,n}(\theta^*, \phi^*) = -\varphi_{m,n}(\theta_0, \phi_0) + \pi am + \pi bn, \quad (11)$$

where a and b are even integers, because (11) would then imply $|G(\theta^*, \phi^*)| = |G(\theta_0, \phi_0)|$. Further, using the phase expression from (2) and comparing the coefficients of m, n on both sides, (11) gives:

$$\underbrace{-2 \sin \theta_{in} \cos \phi_{in} + \sin \theta_0 \cos \phi_0 + \frac{\lambda a}{2d}}_A = -\sin \theta^* \cos \phi^* \quad (12a)$$

$$\underbrace{-2 \sin \theta_{in} \sin \phi_{in} + \sin \theta_0 \sin \phi_0 + \frac{\lambda b}{2d}}_B = -\sin \theta^* \sin \phi^* \quad (12b)$$

If (12) holds for some choice of even integers a and b , then we have

$$A^2 + B^2 = \sin^2(\theta^*), \text{ and } \tan \phi^* = \frac{B}{A}. \quad (13)$$

Therefore, if there exist even integers a, b such that $A^2 + B^2 \leq 1$, then we can compute $\theta^* = \pm \arcsin \sqrt{A^2 + B^2}$ and $\phi^* = \arctan \frac{B}{A}$. If it turns out that $(\theta^*, \phi^*) \neq (\theta_0, \phi_0)$, then a grating lobe exists at (θ^*, ϕ^*) .

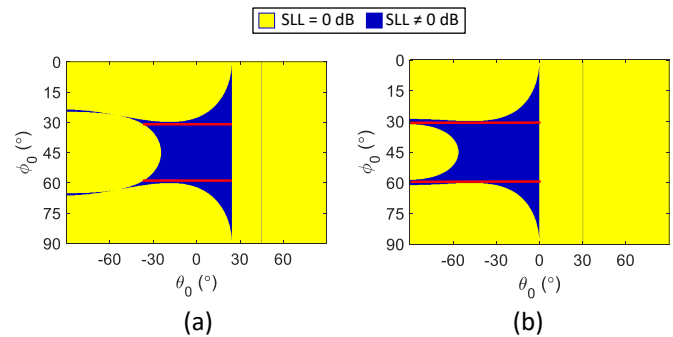


Fig. 11. Plots showing the existence of grating lobes according to (12) for varying main lobe directions (θ_0, ϕ_0) , when the incident and reflected beams lie on the same vertical plane ($\phi_0 = \phi_{in} - 180^\circ$), for (a) $\theta_{in} = -45^\circ$, and (b) $\theta_{in} = -30^\circ$. Yellow regions indicate the presence of a grating lobe at (θ^*, ϕ^*) computed from (12), different from the given main lobe direction (θ_0, ϕ_0) . The red lines indicate the values of ϕ_0 corresponding to largest range of non-zero SLL.

In order to graphically visualize the algebraic conditions for the existence of a grating lobe (*i.e.*, (12)), we consider several cases. For ease of visualization, we fix the incidence beam angle and sweep the required output beam angle, while maintaining the incident and required output wavevectors to be in the same vertical plane. The plots in Fig. 11 show the region (colored yellow) of (θ_0, ϕ_0) for which a grating lobe exists at some $(\theta^*, \phi^*) \neq (\theta_0, \phi_0)$, according to (12). The blue regions indicate the absence of grating lobes. Note that $(-\theta, \phi)$ is equivalent to $(\theta, \phi + \pi)$; thus the left half of the plots, where $\theta_0 < 0$, denote the reflected beam going in the backward

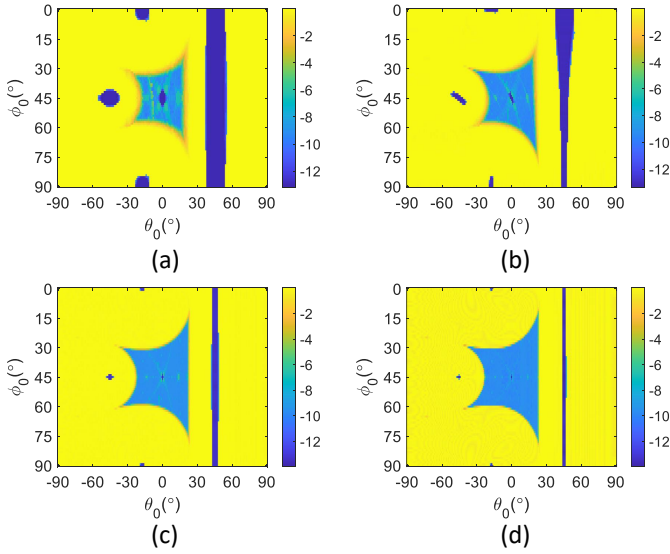


Fig. 12. Plots indicating the simulated SLL (in dB) for varying desired reflected directions (θ_0, ϕ_0) , when the incident and desired reflected directions lie on the same vertical plane ($\phi_0 = \phi_{in} - 180^\circ$) and $\theta_{in} = -45^\circ$, for array sizes: (a) 10×10 , (b) 10×30 , (c) 30×30 , and (d) 50×50 .

scattering direction. We emphasize that the conditions in (12) are irrespective of the size of the IRS; it gives the condition for the existence of a grating lobe if the main lobe is formed *exactly* at (θ_0, ϕ_0) .

We verify the theoretical plots in Fig. 11 by performing simulations for various arrays, where the optimal weights are obtained using OPA. Fig. 12 shows the simulated SLL plots for various array sizes, and they increasingly agree with the theoretical plot in Fig. 11(a) as the array sizes grow. We observe bands around $\theta_0 = 45^\circ$ in Fig. 12, in addition to certain other small regions of non-zero SLL. This occurs due to beamforming error (defined in section III), which decreases with increasing array size. This can be seen by the bands becoming narrower.

Having identified the existence of grating lobes and confirmed them via simulations, we now seek to propose strategies to mitigate these lobes, as their presence is detrimental to the intended operation of an IRS. As we will demonstrate, the proposed OPA and gOPA algorithms are well suited to implement these mitigation strategies. For the sake of definiteness, we only consider 1-bit IRS implementations in this discussion. We remark that for $k > 2$ phase shifts, (11) does not necessarily imply a grating lobe at (θ^*, ϕ^*) , and in general a grating lobe does not occur.

We consider a vertically oriented IRS (*i.e.*, with its surface normal parallel to the ground), and an incident beam emitted by a transmitter with the wavevector parallel to the ground. It is desired to beamform to a mobile user in the same horizontal plane, as depicted in Fig. 13. There arise two scenarios:

Array x -axis parallel with respect to the ground

This amounts to $\phi_0 = 0^\circ$ and $\phi_{in} = 180^\circ$. As we can see from Fig. 11 for $\phi_0 = 0$, a grating lobe always exists,

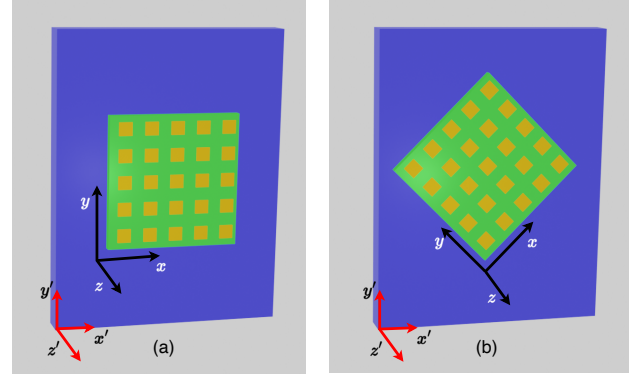


Fig. 13. Two different alignment scenarios of a vertically deployed 5×5 IRS ($x - y$ plane) with respect to the ground ($x' - z'$ plane): (a) \hat{x}, \hat{x}' are parallel, and (b) \hat{x}, \hat{x}' are at an angle.

regardless of the value of θ_{in} . It is easy to see this theoretically, as shown below. (12) simplifies to:

$$2 \sin \theta_{in} + \sin \theta_0 + a = -\sin \theta^* \cos \phi^* \quad (14a)$$

$$b = -\sin \theta^* \sin \phi^*. \quad (14b)$$

Clearly if $b = 0$ and a is chosen to be $2, 0$ or -2 , depending on whether $(2 \sin \theta_{in} + \sin \theta_0)$ lies in $[-3, -1]$, $[-1, 1]$ or $[1, 3]$ respectively, then a grating lobe exists for $\phi^* = 0$ and $\theta^* = \arcsin(2 \sin \theta_{in} + \sin \theta_0 + a)$. The lobe doesn't exist only when $\theta^* = \theta_0$, which happens for two particular angles (one of them being $\theta_0 = -\theta_{in}$). Not only does there exist a grating lobe for almost all values of θ_0 , the grating lobe also lies along the xz plane, as $\phi^* = 0$.

Array x -axis tilted with respect to the ground

When $\phi_{in} \neq 180^\circ$, the issue of grating lobes can be circumvented for some user locations. For e.g. the red lines in Fig. 11 indicate the maximum possible range of grating-lobe free operation for a few different values of θ_{in} . In particular, we see that if $\theta_{in} = -30^\circ$, a user can move in the entire half-range $\theta_0 = [0, 90^\circ]$ for a titled IRS with $\phi_0 = 30^\circ$ and $\phi_{in} = 210^\circ$. Of course, this range is curtailed depending on the size of the array as seen in Fig. 12.

It is often not convenient to tilt an IRS from an implementation perspective, because most deployments are likely to be on walls or sides of buildings, where a horizontal form-factor is more suitable. Therefore, we outline two strategies that mitigating the issue of grating lobes while aligning the IRS axis with the ground.

B. Triangular lattice for IRS layout

In all the examples considered so far, the layout of the constitutive elements was assumed to be rectangular (e.g. see Fig. 1). Instead, if we construct an IRS by arraying its elements in a triangular lattice, it can be shown that $\phi_0 = 0^\circ$ and $\phi_{in} = 180^\circ$ becomes viable for beamforming without encountering the issue of grating lobes.

The array factor is now modified as:

$$G(\theta, \phi) = \frac{1}{MN} \sum_{m,n} w_{m,n} e^{j \frac{2\pi}{\lambda} r_{m,n}^T (-\hat{u}(\theta, \phi) + \hat{u}(\pi - \theta_{in}, \phi_{in}))}, \quad (15)$$

where $\hat{u}(\theta, \phi) = [\sin \theta \cos \phi \ \sin \theta \sin \phi \ \cos \theta]^T$ and $r_{m,n} = md_1 + nd_2$, with $0 \leq n \leq N$ and $-\lfloor \frac{n}{2} \rfloor \leq m \leq M - \lfloor \frac{n}{2} \rfloor$. Here, $d_1 = d[1 \ 0 \ 0]^T$ and $d_2 = d[\frac{1}{2} \ \frac{\sqrt{3}}{2} \ 0]^T$ are the basis vectors for the lattice.

Analogous to the analysis in the case of the rectangular lattice, we can work out the conditions under which a grating lobe appears. The relations for a grating lobe at (θ^*, ϕ^*) are (details in Appendix C):

$$-2 \sin \theta_{in} \cos \phi_{in} + \sin \theta_0 \cos \phi_0 + \frac{\lambda a}{2d} = -\sin \theta^* \cos \phi^*, \quad (16a)$$

$$-2 \sin \theta_{in} \sin \phi_{in} + \sin \theta_0 \sin \phi_0 + \frac{\lambda}{2d} \left(\frac{-a + 2b}{\sqrt{3}} \right) = -\sin \theta^* \sin \phi^*, \quad (16b)$$

where a, b are even integers, as in (12).

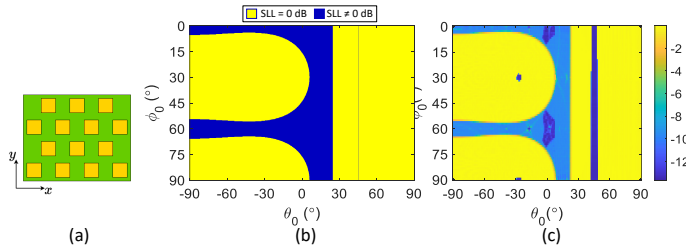


Fig. 14. (a) A schematic of a 3×3 triangular lattice (b) Plot showing the existence of grating lobes according to (16) for varying main lobe directions (θ_0, ϕ_0) , when incident and reflected directions lie on the same vertical plane ($\phi_{in} = \phi_0 + 180^\circ$) and $\theta_{in} = -45^\circ$ (c) Corresponding plot indicating the simulated SLL (measured in dB) for varying desired reflected directions (θ_0, ϕ_0) , for a 30×30 triangular IRS.

Substituting $\phi_{in} = \phi_0 + \pi$ in these relations gives us the theoretical plot as shown in Fig. 14(a), while Fig. 14(b) shows the sidelobe levels after the application of OPA for a finite size IRS. A major difference between the SLL plots of the two different lattices is that $\phi_0 = 0^\circ$ gives a wide range of grating-lobe free operation for triangular lattices, which is not the case in rectangular arrays. A detailed theoretical analysis in support is provided in Appendix C. Thus, the triangular lattice enables the alignment of the IRS axis with the ground while mitigating the grating lobe issue.

C. Prephasing technique for elimination of grating lobes

In the grating lobe considerations considered so far, a non-zero incidence angle was assumed. However, there may be deployment scenarios where keeping $\theta_{in} = 0$ is desired, since this opens up the possibility of steering the reflected beam in either quadrant (*i.e.*, $\theta_0 > 0$ or $\theta_0 < 0$).

However, this presents a fundamental problem, because when: (a) $\theta_{in} = 0$, and (b) the weights are constrained to $\{1, -1\}$, we can show that $|G(\theta, \phi)| = |G(-\theta, \phi)|$ for all

(θ, ϕ) , irrespective of the type of lattice:

$$\begin{aligned} G(-\theta, \phi) &= \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N w_{m,n} e^{j\varphi_{m,n}(-\theta, \phi)} \\ &= \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N w_{m,n}^* \cdot (e^{j\varphi_{m,n}(\theta, \phi)})^* = G(\theta, \phi)^*, \end{aligned} \quad (17)$$

where $*$ denotes complex conjugation. The key argument we used is that $w_{m,n} = w_{m,n}^*$, since each $w_{m,n} \in \mathbb{R}$. This symmetry means that there will always be a grating lobe for normal incidence. As earlier work [13] has identified, if the primary assumption of $w_{m,n} \in \mathbb{R}$ is attacked (for *e.g.* by constraining some of the weights to $\{j, -j\}$), we arrive at a technique of symmetry breaking. This technique is called *prephasing*.

There are several techniques to implement prephasing. For instance, adding random phase perturbations to each element of the IRS also helps break the symmetry [14]. Another approach [16] involves shifting subsets of the IRS elements in the normal direction by a small amount. Yet another possibility is to choose a fraction κ of tuples from $\{1, \dots, M\} \times \{1, \dots, N\}$ (denoting this set as \mathcal{Q}), and to then solve the following optimization problem (indeed, [13] does so with $\kappa = 0.5$):

$$\begin{aligned} \max_{w_{m,n}} \quad & |G(\theta_0, \phi_0)| \\ \text{s.t.} \quad & w_{m,n} \in \{j, -j\}, \quad \forall (m, n) \in \mathcal{Q}, \\ & w_{m,n} \in \{1, -1\}, \quad \forall (m, n) \notin \mathcal{Q}. \end{aligned} \quad (18)$$

In general, there can be more than two prephases [17]. Formally, if the prephase set is $\{\psi_1, \dots, \psi_p\}$, and if $\{\mathcal{Q}_i\}_{i=1}^p$ is a partition of the IRS such that each unit cell in \mathcal{Q}_i gets the prephase ψ_i , then the optimization problem becomes

$$\begin{aligned} \max_{w_{m,n}} \quad & |G(\theta_0, \phi_0)| \\ \text{s.t.} \quad & w_{m,n} \in \{e^{j\psi_i}, e^{j(\psi_i + \pi)}\}, \quad \forall (m, n) \in \mathcal{Q}_i, \quad \forall i = 1, \dots, p. \end{aligned} \quad (19)$$

Remarkably, this problem can be solved by the proposed gOPA, as is evident by a comparison of the above equation with the gOPA formulation in Algorithm 3. Even with the earlier techniques of random phase perturbations [14], the problem of determining the optimum weights for beamforming can be solved by the gOPA.

We implement the prephasing method as captured in (18), and implement the gOPA for various values of the prephasing fraction κ . For each value of κ , the set \mathcal{Q} is chosen randomly such that $|\mathcal{Q}| = \kappa MN$.

Fig. 15 shows the radiation pattern in the xz plane for several values of κ , with the highest SLL suppression of -10.9 dB achieved for $\kappa = 0.5$ in a 30×30 IRS. We also note that while SLL reduces with increasing κ , the magnitude of the mainlobe remains the same. Thus, prephasing is a viable means of grating lobe mitigation while not sacrificing the mainlobe magnitude.

Next, we confirm the feasibility of scanning a single beam across space for (fixed) normal incidence. Fig. 16 shows the radiation patterns along the xz -plane, of a 30×30 IRS in the scanning range $[-30^\circ, 30^\circ]$ spaced by intervals of 10° . The set

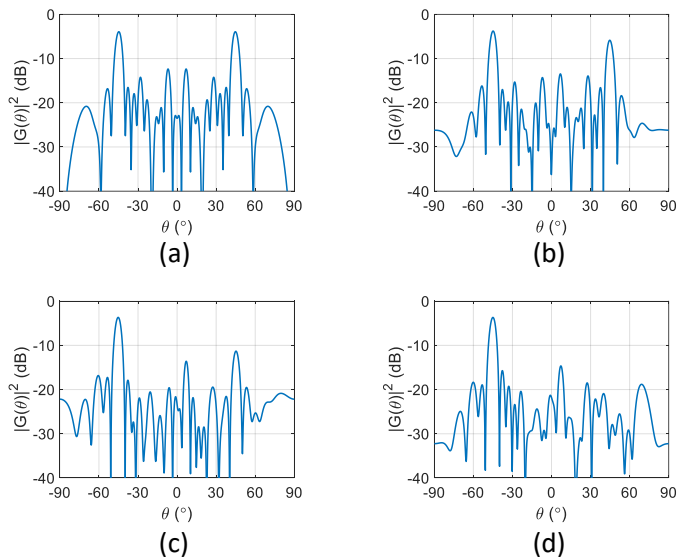


Fig. 15. Radiation patterns for $\theta_{in} = 0^\circ$, $\theta_0 = -45^\circ$, $\phi_0 = 0^\circ$, $\phi_{in} = 180^\circ$ in a 30×30 IRS: (a) $\kappa = 0$ (SLL = 0 dB) (b) $\kappa = 0.1$ (SLL = -2 dB) (c) $\kappa = 0.3$ (SLL = -7.6 dB) (d) $\kappa = 0.5$ (SLL = -10.9 dB)

\mathcal{Q} is chosen randomly, with $\kappa = 0.5$. The worst-case SLL is observed to be -8.6 dB. We note that the SLL can be further improved by optimizing the set \mathcal{Q} itself by using evolutionary algorithms [15], however at a significant computational cost compared to our proposed method.

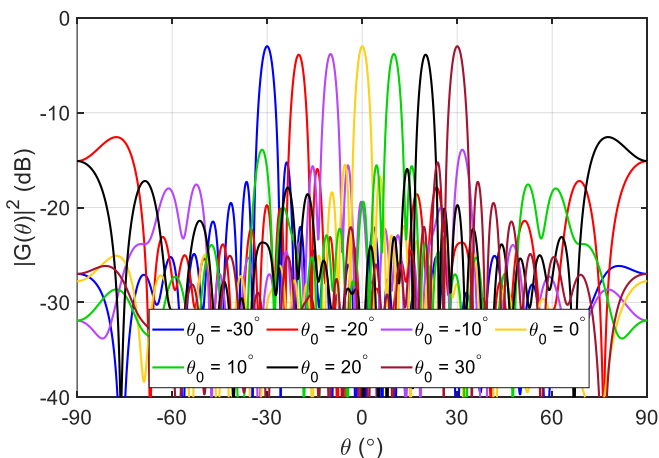


Fig. 16. Radiation patterns along the xz plane corresponding to 30×30 IRS operating at normal incidence, for various θ_0

Finally, we comment on another outcome of the prephasing approach. While originally designed to eliminate the grating lobes issue at normal incidence, it turns out that it offers the opportunity to deploy the IRS at *both* normal or non-normal incidence angle without any issue of grating lobes. We demonstrate this by applying the gOPA on a 30×30 IRS, with the incident and required output wavevectors lying on the same vertical plane ($\phi_{in} = \phi_0 + 180^\circ$) and investigate the behavior w.r.t. incidence angle after implementing prephasing with $\kappa = 0.5$. The results are shown in Fig. 17. We see that there are no grating lobes when operating at normal and non-normal incidence angle (with the exception for small regions

at the corners), without having to tilt the IRS axis with respect to the ground.

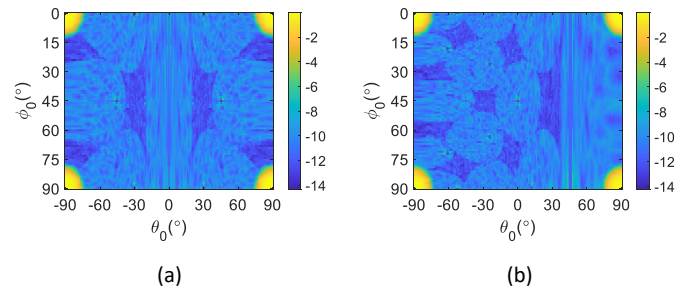


Fig. 17. Sidelobe levels (in dB) on an IRS with prephasing on a 30×30 rectangular IRS for different reflected beam angles (θ_0, ϕ_0), when incident and required wavevectors lie on the same vertical plane ($\phi_{in} = \phi_0 + 180^\circ$): (a) normal incidence, $\theta_{in} = 0$, and (b) non-normal incidence, $\theta_{in} = -45^\circ$.

V. CONCLUSION

In this work, we have presented optimal beamforming algorithms in the context of intelligent reflecting surfaces realized with discrete phase shifts, both for the 1-bit case and the general case of k phase shifts ($k > 2$). Our 1-bit algorithm can be adapted to allow each element of the IRS to select its phase from its own set of possible phases, thus being compatible with imperfect hardware realizations. They also outperform naive thresholding-based implementations. We also demonstrate beamforming in multiple directions. Further, our algorithms come with rigorous proofs of optimality as demonstrated in Theorems 1 and 2 – a feature critically missing in heuristic algorithms. Due to the linear computational complexity of the proposed algorithms, they are extremely fast to deploy in practice, possibly even in real time applications. This makes them an attractive alternative to time-consuming evolutionary algorithms.

We extensively study the issue of grating lobes and the various situations under which they arise via theory and numerical simulations. We discuss two alleviation strategies to mitigate these grating lobes. The first is the use of a triangular lattice in the IRS, and the second is the well known prephasing technique. We demonstrate the versatility of our algorithms by showing how beamforming can easily be accomplished in the presence of prephasing via extensive numerical studies. That said, further work is required to adapt them to more complex objective functions, such as those incorporating sidelobe level minimization, null beamforming directions, and beamforming over a certain bandwidth. We also assume planewave incidence on the IRS in our formulation, and the approach will require modification in the case of nearfield or spherical wave excitation. Our approach currently ignores the presence of mutual coupling between the reflecting elements, and incorporating this is an item of future research.

APPENDIX A PROOF OF THEOREM 1

Suppose that $\tilde{w} = [\tilde{w}_1 \ \dots \ \tilde{w}_n]$ is an optimal solution to the problem in (4), and $\tilde{s} = \sum_{i=1}^n \tilde{w}_i z_i$. We claim that the

line $\ell = \{z \in \mathbb{C} \mid \text{Re}(\tilde{s}z^*) = 0\}$ (i.e., the line perpendicular to \tilde{s} and passing through the origin) satisfies properties 1 and 2 stated in the theorem. We will require the following lemma.

Lemma 1. Suppose a and b are non-zero complex numbers, such that $\text{Re}(ab^*) = \text{Re}(ba^*) \geq 0$. Then, $|a+b|^2 > |a|^2$.

Proof. We have $|a+b|^2 = |a|^2 + |b|^2 + 2\text{Re}(ab^*) > |a|^2$. \square

To prove that ℓ satisfies property 1 (i.e. ℓ does not contain any z_i), assume for contradiction that it does not. Then, there exists some j for which $z_j \in \ell$. Now, by switching $w_j = \tilde{w}_j \rightarrow w_j = -\tilde{w}_j$, we obtain the new sum $s = \sum_{i=1}^n w_i z_i = \tilde{s} - 2\tilde{w}_j z_j$. Noting that $\text{Re}(\tilde{s} \cdot (2\tilde{w}_j z_j)^*) = 2\tilde{w}_j \text{Re}(\tilde{s} z_j^*) = 0$, by Lemma 1 we have $|s|^2 > |\tilde{s}|^2$, which contradicts the optimality of \tilde{w} . Therefore, property 1 has to be satisfied by ℓ .

To prove that ℓ satisfies property 2 (i.e. $\tilde{w}_i = 1$ for all z_i lying on one side of the line, and $\tilde{w}_i = -1$ for all z_i on the other side), we note that ℓ separates the complex plane into two half-planes given by:

$$\mathcal{H}_1 = \{z \in \mathbb{C} \mid \text{Re}(\tilde{s}z^*) > 0\}, \quad \mathcal{H}_2 = \{z \in \mathbb{C} \mid \text{Re}(\tilde{s}z^*) < 0\}.$$

We claim that $\tilde{w}_j = 1$ for every $z_j \in \mathcal{H}_1$. To prove this, suppose the contrary that $\tilde{w}_j = -1$ for some $z_j \in \mathcal{H}_1$. Then, by switching $w_j = \tilde{w}_j = -1 \rightarrow w_j = 1$, we obtain the new sum $s = \sum_{i=1}^n w_i z_i = \tilde{s} + 2z_j$, and noting that $\text{Re}(\tilde{s}z_j^*) > 0$, we have $|s|^2 > |\tilde{s}|^2$, contradicting the optimality of \tilde{w}_j . A similar line of argument can be used to show that $\tilde{w}_j = -1$ for every $z_j \in \mathcal{H}_2$. Hence, property 2 also holds.

APPENDIX B PROOF OF THEOREM 2

Let $\tilde{s} = \sum_{i=1}^n \tilde{w}_i z_i$, where $\{\tilde{w}_i\}$ represent the optimal solution to the problem in (6), and define $\delta = \arg \tilde{s}$. We will show that this δ satisfies the required properties.

To prove that property 1 holds (i.e. none of the z_i lie on the edges of the cones of the rotated partition $\mathcal{P}(\delta)$), suppose for contradiction that some z_i lies on the edge of two cones $A_j(\delta)$ and $A_l(\delta) \iff e^{-j\delta} z_i$ lies on the edge of A_j and A_l . Being a limit point of both A_j and A_l , we have $\text{Re}((a_j - a_m)z_i e^{-j\delta}) \geq 0 \forall m \neq j$ (for $A_j(\delta)$), and similarly $\text{Re}((a_l - a_m)z_i e^{-j\delta}) \geq 0 \forall m \neq l$ (for $A_l(\delta)$). Together, this implies: $\text{Re}((a_j - a_l)e^{-j\delta} z_i) = 0$. We now have two cases:

- 1) Suppose $\tilde{w}_i \neq a_j$. Then, by switching from $w_i = \tilde{w}_i$ to $w_i = a_j$, we obtain the new sum $s = \tilde{s} + (a_j - \tilde{w}_i)z_i$. Since $\text{Re}((a_j - \tilde{w}_i)z_i \tilde{s}^*) = |\tilde{s}| \text{Re}((a_j - \tilde{w}_i)z_i e^{-j\delta}) \geq 0$, by Lemma 1 we have $|s|^2 > |\tilde{s}|^2$, contradicting the optimality of \tilde{w} .
- 2) Suppose $\tilde{w}_i = a_j$. Then, by switching from $w_i = a_j$ to $w_i = a_l$, we obtain the new sum $s = \tilde{s} + (a_l - a_j)z_i$. Since $\text{Re}((a_l - a_j)z_i \tilde{s}^*) = |\tilde{s}| \text{Re}((a_l - a_j)z_i e^{-j\delta}) = 0$, by Lemma 1 we have $|s|^2 > |\tilde{s}|^2$, once again contradicting the optimality of \tilde{w} .

To prove property 2 holds (i.e. if z_i belongs to the cone $A_j(\delta)$ in the rotated partition $\mathcal{P}(\delta)$, then $\tilde{w}_i = a_j$), let $z_i \in A_j(\delta) \iff e^{-j\delta} z_i \in A_j$ and suppose for contradiction that $\tilde{w}_i \neq a_j$. Then, by switching from $w_i = \tilde{w}_i$ to $w_i = a_j$, we obtain the new sum $s = \tilde{s} + (a_j - \tilde{w}_i)z_i$. Since $\text{Re}((a_j - \tilde{w}_i)z_i \tilde{s}^*) = |\tilde{s}| \text{Re}((a_j - \tilde{w}_i)z_i e^{-j\delta}) > 0$

from the definition of A_j , by Lemma 1 we have $|s|^2 > |\tilde{s}|^2$, contradicting the optimality of \tilde{w} . Thus, $\tilde{w}_i = a_j$, and hence property 2 holds.

APPENDIX C DERIVATION OF GRATING LOBE CONDITION FOR TRIANGULAR LATTICE

Define the matrix $D = \begin{bmatrix} d_1 & d_2 \end{bmatrix}$. Then, we can write $x_{m,n} = D \begin{bmatrix} m & n \end{bmatrix}^T$ for all m, n . Noting that $\varphi_{m,n}(\theta, \phi) = \frac{2\pi}{\lambda} x_{m,n}^T (-\hat{u}(\theta, \phi) + \hat{u}(\pi - \theta_{in}, \phi_{in}))$, the grating lobe condition (11) becomes

$$\frac{2\pi}{\lambda} \begin{bmatrix} m & n \end{bmatrix} D^T (-\hat{u}(\theta^*, \phi^*) - \hat{u}(\theta_0, \phi_0) + 2\hat{u}(\pi - \theta_{in}, \phi_{in})) = \pi \begin{bmatrix} m & n \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix}^T. \quad (20)$$

Since the above condition holds for all m, n , we must necessarily have:

$$\frac{2\pi}{\lambda} D^T (-\hat{u}(\theta^*, \phi^*) - \hat{u}(\theta_0, \phi_0) + 2\hat{u}(\pi - \theta_{in}, \phi_{in})) = \pi \begin{bmatrix} a & b \end{bmatrix}^T. \quad (21)$$

Define $v = \begin{bmatrix} -\sin \theta_0 \cos \phi_0 - \sin \theta^* \cos \phi^* + 2 \sin \theta_{in} \cos \phi_{in} \\ -\sin \theta_0 \sin \phi_0 - \sin \theta^* \sin \phi^* + 2 \sin \theta_{in} \sin \phi_{in} \end{bmatrix}$. Then, (21) becomes

$$\frac{2\pi d}{\lambda} \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} v = \pi \begin{bmatrix} a \\ b \end{bmatrix} \implies v = \frac{\lambda}{2d} \begin{bmatrix} a \\ -\frac{a+2b}{\sqrt{3}} \end{bmatrix}. \quad (22)$$

Comparing the components on both sides of (22) gives us the desired (16).

Analysis for $\phi_0 = 0^\circ, \phi_{in} = 180^\circ$

The grating lobe condition now reduces to the following system of equations

$$2 \sin \theta_{in} + \sin \theta_0 + a = -\sin \theta^* \cos \phi^*, \quad (23a)$$

$$\frac{-a+2b}{\sqrt{3}} = -\sin \theta^* \sin \phi^*. \quad (23b)$$

Since $\sin \theta^* \sin \phi^* \in [-1, 1]$ and $-a+2b$ is even, (23b) forces $a = 2b$. Since b is even, a is a multiple of 4. However, (23a) forces $a \in \{-2, 0, 2\}$, because $\sin \theta^* \cos \phi^* \in [-1, 1]$ and $2 \sin \theta_{in} + \sin \theta_0 \in (-3, 3)$. Therefore, we must have $a = b = 0$, and hence (23a) and (23b) reduce to $2 \sin \theta_{in} + \sin \theta_0 = -\sin \theta^* \cos \phi^*$ and $0 = \sin \theta^* \sin \phi^*$ respectively.

Now, if $-2 \sin \theta_{in} - \sin \theta_0 > 1$ (equivalently, $\theta_0 < \arcsin(-2 \sin \theta_{in} - 1)$), then clearly (23a) cannot be satisfied by any (θ^*, ϕ^*) , and hence a grating lobe will not exist. In particular, for $\theta_{in} = -45^\circ$, a grating lobe will not exist when $\theta_0 < \arcsin(\sqrt{2} - 1) \approx 24.5^\circ$, which is observed in Fig. 14.

APPENDIX D PROOF OF THEOREM 3

We recall the following relations mentioned in the theorem:

- $w^{(k)} = \arg \max_{w_i} |G_1 + \sum_{j=2}^l \alpha_j^{(k)} G_j|$.
- $\alpha_j^{(k+1)} = e^{j(\arg(G_1) - \arg(G_j))}$ for $j = 2, \dots, l$, where each G_j is computed using the weights $w^{(k)}$.

We defined $c_k = \max_{w_i} |G_1 + \sum_{j=2}^l \alpha_j^{(k)} G_j|$, and $d_k = |G_1(w^{(k)})| + \dots + |G_l(w^{(k)})|$. Our goal is to show that the sequences $\{c_k\}_{k \geq 1}$ and $\{d_k\}_{k \geq 1}$ are both non-decreasing.

Now, we have $c_k = \max_{w_i} |G_1 + \sum_{j=2}^l \alpha_j^{(k)} G_j| = |G_1(w^{(k)}) + \sum_{j=2}^l \alpha_j^{(k)} G_j(w^{(k)})|$ by definition of $w^{(k)}$. It follows from the triangle inequality that $c_k \leq d_k$.

Next, by definition of $\alpha_j^{(k+1)}$, we have $\arg(G_1(w^{(k+1)})) = \arg(\alpha_j^{(k+1)} G_j(w^{(k+1)}))$ for all $j \in \{2, \dots, l\}$, and hence we have

$$|G_1(w^{(k+1)}) + \sum_{j=2}^l \alpha_j^{(k+1)} G_j(w^{(k+1)})| = \sum_{j=1}^l |G_j(w^{(k+1)})| = d_{k+1}. \quad (24)$$

However, by the definition of c_{k+1} , we have

$$c_{k+1} \geq |G_1(w^{(k+1)}) + \sum_{j=2}^l \alpha_j^{(k+1)} G_j(w^{(k+1)})|, \quad (25)$$

and hence, from (24) we have $c_{k+1} \geq d_k$. Once again, we have $c_{k+1} \leq d_{k+1}$ due to the triangle inequality. Therefore, we have the chain of inequalities $c_k \leq d_k \leq c_{k+1} \leq d_{k+1}$ for all $k \geq 1$, which in particular implies that both the sequences $\{c_k\}_{k \geq 1}$ and $\{d_k\}_{k \geq 1}$ are non-decreasing, as desired.

ACKNOWLEDGMENT

The authors acknowledge useful discussions on beamforming with Dr. Hanumantha Rao, Director General of SAMEER, India, Prof. S. Aniruddhan, Electrical Engineering, IIT Madras, and Mr. Amitesh Kumar of SAMEER, India.

REFERENCES

- [1] W. Mei, B. Zheng, C. You, and R. Zhang, "Intelligent Reflecting Surface-Aided Wireless Networks: From Single-Reflection to Multireflection Design and Optimization," *Proceedings of the IEEE*, vol. 110, no. 9, pp. 1380–1400, 2022.
- [2] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, "Increasing Indoor Spectrum Sharing Capacity Using Smart Reflect-Array," in *2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2016.
- [3] A. M. Elbir, K. V. Mishra, S. A. Vorobyov, and R. W. Heath, "Twenty-Five Years of Advances in Beamforming: From Convex and Nonconvex Optimization to Learning Techniques," *IEEE Signal Processing Magazine*, vol. 40, no. 4, pp. 118–131, 2023.
- [4] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex Optimization-Based Beamforming," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 62–75, 2010.
- [5] R. Lorenz and S. Boyd, "Robust Minimum Variance Beamforming," *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1684–1696, 2005.
- [6] Y. Huang, M. Zhou, and S. A. Vorobyov, "New Designs on MVDR Robust Adaptive Beamforming Based on Optimal Steering Vector Estimation," *IEEE Transactions on Signal Processing*, vol. 67, no. 14, pp. 3624–3638, 2019.
- [7] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit Beamforming for Physical-Layer Multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, 2006.
- [8] Ö. T. Demir and T. E. Tuncer, "Alternating Maximization Algorithm for the Broadcast Beamforming," in *2014 22nd European Signal Processing Conference (EUSIPCO)*, pp. 1915–1919, 2014.
- [9] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.
- [10] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid Beamforming for Reconfigurable Intelligent Surface based Multi-User Communications: Achievable Rates With Limited Discrete Phase Shifts," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1809–1822, 2020.
- [11] C. You, B. Zheng, and R. Zhang, "Channel Estimation and Passive Beamforming for Intelligent Reflecting Surface: Discrete Phase Shift and Progressive Refinement," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2604–2620, 2020.
- [12] S. Gong, Z. Yang, C. Xing, J. An, and L. Hanzo, "Beamforming Optimization for Intelligent Reflecting Surface-Aided SWIPT IoT Networks Relying on Discrete Phase Shifts," *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 8585–8602, 2021.
- [13] J. Yin, Q. Wu, Q. Lou, H. Wang, Z. N. Chen, and W. Hong, "Single-Beam 1 Bit Reflective Metasurface Using Prephased Unit Cells for Normally Incident Plane Waves," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 7, pp. 5496–5504, 2020.
- [14] B. G. Kashyap, P. C. Theofanopoulos, Y. Cui, and G. C. Trichopoulos, "Mitigating Quantization Lobes in mmWave Low-Bit Reconfigurable Reflective Surfaces," *IEEE Open Journal of Antennas and Propagation*, vol. 1, pp. 604–614, 2020.
- [15] J. Yin, Q. Wu, H. Wang, and Z. N. Chen, "Prephase-Based Equivalent Amplitude Tailoring for Low Sidelobe Levels of 1-Bit Phase-Only Control Metasurface Under Plane Wave Incidence," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 11, pp. 10604–10613, 2022.
- [16] D. Vabichevich, A. Belov, and A. Sayanskiy, "Suppression of Quantization Lobes in 1-Bit Reconfigurable Intelligent Surfaces," *IEEE Antennas and Wireless Propagation Letters*, vol. 22, no. 12, pp. 2808–2811, 2023.
- [17] M. Cheng, Q. Wu, C. Yu, H. Wang, and W. Hong, "A Prephased Electronically Steered Phased Array That Uses Very-Low-Resolution Phase Shifters and a Hybrid Phasing Method," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 9, pp. 7310–7322, 2023.
- [18] M. Smith and Y. Guo, "A Comparison of Methods for Randomizing Phase Quantization Errors in Phased Arrays," *IEEE transactions on antennas and propagation*, vol. 31, no. 6, pp. 821–828, 1983.
- [19] H. Yang, F. Yang, S. Xu, M. Li, X. Cao, J. Gao, and Y. Zheng, "A Study of Phase Quantization Effects for Reconfigurable Reflectarray Antennas," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 302–305, 2017.
- [20] S. Zhang, "Quadratic Maximization and Semidefinite Relaxation," *Mathematical Programming*, vol. 87, no. 3, pp. 453–465, 2000.
- [21] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Society for Industrial and Applied Mathematics, 2001.
- [22] B. Wu, A. Sutinjo, M. E. Potter, and M. Okoniewski, "On the Selection of the Number of Bits to Control a Dynamic Digital MEMS Reflectarray," *IEEE Antennas and Wireless Propagation Letters*, vol. 7, pp. 183–186, 2008.
- [23] S. Ebadi, R. V. Gatti, and R. Sorrentino, "Linear Reflectarray Antenna Design Using 1-Bit Digital Phase Shifters," in *2009 3rd European Conference on Antennas and Propagation*, pp. 3729–3732, 2009.
- [24] F. Sohrabi and W. Yu, "Hybrid Digital and Analog Beamforming Design for Large-Scale Antenna Arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, 2016.
- [25] H. Zhang, B. Di, L. Song, and Z. Han, "Reconfigurable Intelligent Surfaces Assisted Communications With Limited Phase Shifts: How Many Phase Shifts Are Enough?," *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 4498–4502, Apr. 2020.
- [26] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1838–1851, 2019.
- [27] W. Yan, G. Sun, W. Hao, Z. Zhu, Z. Chu, and P. Xiao, "Machine learning-based beamforming design for millimeter wave irs communications with discrete phase shifters," *IEEE Wireless Communications Letters*, vol. 11, no. 12, pp. 2467–2471, 2022.
- [28] Q. J. Lim, C. Ross, A. Ghosh, F. W. Vook, G. Gradoni, and Z. Peng, "Quantum-Assisted Combinatorial Optimization for Reconfigurable Intelligent Surfaces in Smart Electromagnetic Environments," *IEEE Transactions on Antennas and Propagation*, vol. 72, no. 1, pp. 147–159, 2024.
- [29] C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu, Y. Pan, H. Ren, M. Di Renzo, A. L. Swindlehurst, R. Zhang, et al., "An overview of signal processing techniques for ris/irs-aided wireless systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 883–917, 2022.
- [30] Y. Zhang, K. Shen, S. Ren, X. Li, X. Chen, and Z.-Q. Luo, "Configuring intelligent reflecting surface with performance guarantees: Optimal beamforming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 967–979, 2022.
- [31] S. Ren, K. Shen, X. Li, X. Chen, and Z.-Q. Luo, "A linear time algorithm for the optimal discrete irs beamforming," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 496–500, 2022.

- [32] S. S. Narayanan, R. K. Ganti, and U. K. Khankhoje, "Discrete Variable Beamforming for Intelligent Reflecting Surfaces," in *2024 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI)*, 2024 [to appear].
- [33] C. Elachi, Y. Kuga, K. McDonald, K. Sarabandi, F. Ulaby, M. Whitt, H. Zebker, and J. van Zyl, *Radar polarimetry for geoscience applications*. Norwood, MA (USA); Artech House Inc., 1990.
- [34] C. A. Balanis, *Antenna theory: analysis and design*. John Wiley & sons, 2016.
- [35] A. K. Bhattacharyya, *Phased array antennas: Floquet analysis, synthesis, BFNs and active array systems*, vol. 179. John Wiley & Sons, 2006.

Sai Sanjay Narayanan is currently an undergraduate student in the Engineering Physics program at the Indian Institute of Technology Madras, Chennai.

Uday K Khankhoje (Senior Member, IEEE) received the B.Tech. degree in Electrical Engineering from the Indian Institute of Technology Bombay, Mumbai, in 2005, and the M.S. and Ph.D. degrees in Electrical Engineering from the California Institute of Technology, Pasadena, in 2010. He was a Caltech Postdoctoral Scholar at the Jet Propulsion Laboratory (NASA/Caltech) from 2011-12, a Postdoctoral Research Associate in the Department of Electrical Engineering at the University of Southern California, Los Angeles, USA, from 2012-13, and an Assistant Professor of Electrical Engineering at the Indian Institute of Technology Delhi, India from 2013-16. Since 2016 he has been at the Department of Electrical Engineering in the Indian Institute of Technology Madras, Chennai, where he is currently an Associate Professor and leads the numerical electromagnetics and optics (NEMO) research group which focuses on solving electromagnetics inspired inverse problems.

Radha Krishna Ganti (Member, IEEE) received the B.Tech. and M.Tech. degrees in Electrical Engineering from the Indian Institute of Technology Madras, Chennai, India in 2004, and the Master's degree in Applied Mathematics and the Ph.D. degree in Electrical Engineering from the University of Notre Dame in 2009. He is currently a Professor of Electrical Engineering at the Indian Institute of Technology Madras, Chennai. He is a coauthor of the monograph, *Interference in Large Wireless Networks* (NOW Publishers, 2008). His doctoral work focused on the spatial analysis of interference networks using tools from stochastic geometry. He received the 2014 IEEE Stephen O. Rice Prize, the 2014 IEEE Leonard G. Abraham Prize, and the 2015 IEEE Communications Society Young Author Best Paper Award. He was also awarded the 2016–2017 Institute Research and Development Award (IRDA) by IIT Madras. In 2019, he was awarded the TSDSI Fellow for Technical Excellence in standardization activities and contribution to LMLC use case in ITU. He was the Lead PI from IITM involved in the development of 5G base stations for the 5G Testbed Project funded by the Department of Telecommunications, Government of India.