

Achieving Optimum Received Power for Discrete-Phase RISs with Elementwise Updates in the Least Number of Steps

DOGAN KUTAY PEKCAN, STUDENT MEMBER, IEEE AND ENDER AYANOGLU, FELLOW, IEEE

¹Center for Pervasive Communications and Computing (CPCC), Department of Electrical Engineering and Computer Science, University of California, Irvine, Irvine, CA 92697 USA

CORRESPONDING AUTHOR: Ender Ayanoglu (e-mail: ayanoglu@uci.edu).

This work is partially supported by NSF grant 2030029.

ABSTRACT The problem of optimizing discrete phases in a reconfigurable intelligent surface (RIS) to maximize the received power at a user equipment is addressed. Necessary and sufficient conditions to achieve this maximization are given. These conditions are employed in an algorithm to achieve the maximization. New versions of the algorithm are given that are proven to achieve convergence in N or fewer steps whether the direct link is completely blocked or not, where N is the number of the RIS elements, whereas previously published results achieve this in KN or $2N$ number of steps where K is the number of discrete phases. Thus, for a discrete-phase RIS, the techniques presented in this paper achieve the optimum received power in the smallest number of steps published in the literature. In addition, in each of those N steps, the techniques presented in this paper determine only one or a small number of phase shifts with a simple elementwise update rule, which result in a substantial reduction of computation time, as compared to the algorithms in the literature. As a secondary result, we define the uniform polar quantization (UPQ) algorithm which is an intuitive quantization algorithm that can approximate the continuous solution with an approximation ratio of $\text{sinc}^2(1/K)$ and achieve low time-complexity, given perfect knowledge of the channel.

INDEX TERMS Intelligent reflective surface (IRS), reconfigurable intelligent surface (RIS), discrete phase configuration, global optimum, linear time discrete beamforming for large IRS/RIS, discrete quadratic program, uniform quantization.

I. INTRODUCTION

IN wireless communications, the performance demand has significantly grown with the advancements in technology, requiring to cope with many challenges in the propagation environment. In dense urban outdoor-indoor areas, the problems of shadowing and fading can become severe. In specific scenarios where the direct link between the base stations and users is blocked, despite the more extensive beamforming capability of the base stations, the performance can be significantly degraded [1]. In order to catch up with the ever-growing quality of service (QoS) and energy efficiency requirements in mobile communications, the challenges due

to blockages must be overcome to prevent performance degradation [2].

A reconfigurable intelligent surface (RIS), also known as intelligent reflective surface (IRS), can manipulate the incident electromagnetic waves to control the propagation environment by varying the phases of the incident signals with its low-cost passive reflecting elements [3]. In recent years, RISs have been studied in many communication systems in the literature, and it is shown that RIS can become a crucial enabler for communication environments, especially when there is a loss in line-of-sight (LOS) between the base station (BS) and user equipment (UE) [4], [5].

For complex communication scenarios, algorithms are developed with the assumption of continuous phase shifts at the RIS, for ease of optimization [6], [7], [8], [9]. In [6], to achieve a globally optimal solution, the authors developed an algorithm which is based on the branch-and-bound method. In [7], the authors developed a secure wireless communication system for a single-user case by adopting the majorization-minimization and block coordinate descent (BCD) techniques. Assuming perfect knowledge of the channels in [8], the authors jointly designed the active and passive beamforming to minimize the transmit power, subject to individual signal-to-interference-ratio constraints of the users. In [9], the authors consider an RIS-aided multi-antenna transmission with the effect of channel fading and phase noise impairments at the RIS, to maximize the signal-to-noise-ratio (SNR) at the receiver.

Practical large RISs can usually employ quantized phase shift as they are more cost-effective [10]. A two-stage approach to address the discrete phase shifts constraints is to project the continuous solution to the closest value in the discrete set [10], [11], [12], [13]. Although, as shown in the present paper, the quantization approach can potentially provide high performance and low computational complexity, it can only give marginal insight on the actual optimum discrete phase shifts selection problem.

When it comes to designing algorithms with discrete phase shift constraints, as the number of possible solutions increases exponentially with the number of RIS elements, exponential search techniques are required [14], and a closed-form solution is practically unavailable, as we discuss in Section IV-A. As an example, the authors in [15] stated that the discrete beamforming problem for the RIS turns out to be a generally NP-hard discrete quadratic program (QP), and a generic K -ary discrete QP remains as an open problem. To guarantee the optimal solution, most of the prior work in single-user scenarios had exponential complexity, sometimes over the number of RIS elements (N) [10], and sometimes over the phase shift levels (K) [16]. In this regard, probabilistic optimization techniques have also drawn attention [17], [18], [19]. In [17], the idea of probabilistic data association (PDA) is used to address general binary quadratic problems (BQPs) for multi-user detection in code division multiple access. Machine-to-machine wireless communication with high reliability is considered in [18], where authors approached BQPs by developing a PDA algorithm, achieving near-optimal results. Recently, authors in [19] developed a comprehensive probabilistic technique to address the quantization error and scalability issues in a variety of discrete RIS optimization problems, where the proposed technique outperforms general approaches such as the closest point projection (CPP) method given in [13].

To the best of our knowledge, there is still a gap in the literature for further research in the optimal discrete beamforming problem. Regarding the previously published work addressing the optimal discrete beamforming, the authors

in [20] pointed out that an N -step search algorithm for K -ary beamforming can be developed. In [13], the authors provided an optimal algorithm for the binary case, where $2N + 2$ steps are required for convergence. Also in [13], the authors provided an approximation algorithm (APX), and a simple quantization algorithm for K -ary beamforming, where the performance is evaluated over the estimated channels. The authors in [21] proposed a K -ary optimal discrete beamforming algorithm with a polynomial search complexity, i.e., $2N(K - 1)$ steps to ensure optimality. Recently, the authors in [15] proposed the first K -ary linear time algorithm with KN steps to converge, which could be reduced to $2N$ steps when the direct link is not completely blocked.

In this work, we address the problem of optimizing discrete phases in an RIS to maximize the received power at a UE for a single-input single-output (SISO) system with full channel state information (CSI). We also extend our solutions to special cases of multiple-input multiple-output (MISO), multiple users, and imperfect CSI scenarios. The main contributions are given as follows:

- We provide necessary and sufficient conditions to achieve this maximization. We employ these conditions to develop a linear time algorithm achieving provable optimality convergence in N or *fewer* steps, whether the direct link between the BS and the UE is blocked or not. Employing the geometric approach in [15], we develop a formal elementwise update rule to be used, so that in each of those N or *fewer* steps, one or a small number of elements are updated, respectively.
- We prove the periodicity in elementwise updates for the optimal beamforming, and exploit this periodicity to both reduce the number of steps, and to provide a simple initialization for the search algorithms. With this, we show the relation between the channel phases and the pattern of elements in which the periodicity must occur.
- Our developed algorithms are shown to give a substantial reduction of computation time, as compared to the algorithms in the literature [13], [14], [15], [20]. Besides providing improved optimal discrete phase shift selection algorithms, this work gives further insights on the discrete beamforming optimization, with the periodicity rule.
- We formally define an intuitive, quantization-based, discrete phase shift selection algorithm, which we will call uniform polar quantization (UPQ). We show that provided full CSI is available, it performs well, in terms of both performance and the computational time complexity. With UPQ, we reveal that the optimal discrete beamforming can be closely approximated as a quantization solution.

The received power maximization problem in this work is connected to generic K -ary discrete QP, where the objective is to maximize $\mathbf{x}^T \mathbf{Q} \mathbf{x}$, with $\text{rank}(\mathbf{Q}) = 1$ [14], [15], where

the single eigenvector is positive. Therefore, the linear time algorithms with the elemental update rules in the present paper can be employed in similar maximization problems, so that the global optimum can be achieved in linear time.

II. SYSTEM MODEL

Consider a point-to-point communication scenario aided by an RIS, where there is a non-line-of-sight (NLOS) channel between the BS and the UE. The RIS has N elements located over N_z rows and N_y columns, in a uniform planar array (UPA) structure, as shown in Fig. 1. We consider K discrete phase shifts for the RIS, i.e., $\theta_n \in \Phi_K$ and $\Phi_K = \{\omega, 2\omega, \dots, K\omega\}$ with $\omega = \frac{2\pi}{K}$ and $j = \sqrt{-1}$. The set Φ_K can equivalently be described as $\{0, \omega, 2\omega, \dots, (K-1)\omega\}$. Hence, the RIS introduces amplitude β_n^R and phase shift θ_n for $n = 1, 2, \dots, N$ in the N -element reflection coefficient vector

$$\mathbf{w} = [\beta_1^R e^{j\theta_1}, \beta_2^R e^{j\theta_2}, \dots, \beta_N^R e^{j\theta_N}] \quad (1)$$

where for practicality, we let $\beta_n^R = 1$ in this paper. Let $s \in \mathbb{C}$ be the transmitted symbol. The received signal at the UE is given as [8]

$$y = (\mathbf{h}_u^H \mathbf{W} \mathbf{h}_b + h_0)s + z, \quad (2)$$

where $\mathbf{W} = \text{diag}(\mathbf{w})$, $h_0 \in \mathbb{C}$ is the direct link between the BS and UE, $\mathbf{h}_u \in \mathbb{C}^{N \times 1}$ and $\mathbf{h}_b \in \mathbb{C}^{N \times 1}$ are the equivalent channels of the RIS-UE and BS-RIS links, respectively, and z is the additive white Gaussian noise (AWGN) at the UE antenna.

Let $\mathbf{h} = \mathbf{h}_u^* \odot \mathbf{h}_b$, where \odot is the elementwise (Hadamard) multiplication of the two vectors, and the additive noise z be a complex Gaussian random variable with variance σ^2 , i.e., $z \sim \mathcal{CN}(0, \sigma^2)$. Assuming a mean power constraint at the BS, i.e., $\mathbb{E}[|s|^2] \leq P$, the achievable ergodic data rate in bps/Hz is given by

$$\gamma = \mathbb{E} \left[\log \left(1 + \frac{P}{\sigma^2} |h_0 + \mathbf{h}^T \mathbf{w}|^2 \right) \right]. \quad (3)$$

Therefore, the maximization of the channel power inside the logarithm in (3) amounts to maximizing the ergodic data rate, which is a commonly used performance metric in wireless communications. With this, the problem of maximizing the overall channel gain, i.e., the received signal power by performing discrete beamforming at the RIS is given as

$$(P1) \quad \max_{\substack{[w]_n \in \Phi_K}} |h_0 + \mathbf{h}^T \mathbf{w}|^2. \quad (4)$$

The received power maximization problem in (P1) is shown to belong to the class of general quadratic programming problems, i.e., K -ary discrete QP, where the objective is to maximize $\mathbf{x}^T \mathbf{Q} \mathbf{x}$, with $\text{rank}(\mathbf{Q}) = 1$, which has been known to be NP-hard [14], [15], [22]. Furthermore, authors in [14] show that the problem (P1) can be reformulated as

$$(P2) \quad \max_{\substack{\bar{w}_1, \dots, \bar{w}_{N+1} \in \Phi_K}} |\tilde{\mathbf{g}}^H \bar{\mathbf{w}}|^2 \quad (5)$$

where $\bar{\mathbf{w}} = [e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_N}, e^{j\theta_{N+1}}]^T$ and $\tilde{\mathbf{g}}$ corresponds to the unique eigenvector of $\tilde{\mathbf{Q}}$ with $\tilde{\mathbf{Q}} =$

$[\mathbf{h}, h_0][\mathbf{h}, h_0]^H$. So, the solution of (P1) can be extracted as $\hat{\mathbf{w}} = \frac{\bar{\mathbf{w}}(1:N)}{\bar{\mathbf{w}}(N+1)}$, where $\bar{\mathbf{w}}(1:N)$ correspond to the first N elements in $\bar{\mathbf{w}}$ and $\bar{\mathbf{w}}(N+1)$ is the element at the last index. Therefore, the problem of interest is the maximization of the inner product either in (P1) or (P2) with discrete phase shift constraints at the RIS¹.

The equivalent maximization problems (P1) and (P2) frequently arise in RIS-aided communications systems. For example, in RIS-assisted single-user multiple-input multiple-output (MIMO) localization systems, in two-stage approaches to refine the localization performance, the objective boils down to solving the discrete beamforming design for the RIS to maximize the passive beamforming gain [23], [24]. In fact, in [23], the authors develop a discrete beamforming approach named Fast Passive Beamforming (FPB) algorithm that can achieve the global optimum for the very same problem in (P2) and used for improving the localization performance in mmWave MIMO. Also, a joint framework for channel estimation and passive beamforming for an RIS is developed with the discrete phase shift constraints at the RIS [12], where the exact problem (P2) needs to be solved for the channel-gain-maximization approach in the progressive passive beamforming refinement.

The special case of single-antenna assumption has importance and draws attention in the literature, as problems (P1) and (P2), especially for analyzing the potential of the RIS in general scenarios [4], practical codebook design and optimization for RISs [25], [26], and experimental setups with an RIS where the system is tested and optimized with discrete phase shifts [5], [27], [28]. Therefore, having an efficient algorithm with provable global optimality as a benchmark is crucial to properly assess the performance of the developed algorithms in experimental setups, as well as to determine how reliable the direct quantization approach can be under different channels.

Beyond the extent of (P1) and (P2), ample research in the context of MISO and MIMO systems with discrete beamforming at the RISs has been conducted in the recent literature [29], [30]. In [29], authors performed joint active and passive beamforming with the generalized Benders decomposition (GBD)-based algorithm with codebook-based passive beamforming at multiple RISs for a multi-user MISO system to minimize the transmit power subject to signal-to-interference-and-noise-ratio (SINR). The results are shown to approximate the global optimum. In [30], the authors consider a RIS-aided MIMO system with transceiver hardware impairments, where the aim is to minimize the total mean squared errors (MSEs) of multiple data streams. They propose a two-tier majorization-minimization (MM) based and a modified Riemannian gradient descent (RGA) algorithm to obtain the sub-optimal solution of the RIS reflection matrix. Thus, regarding the recent work with more

¹Strictly speaking, it is the magnitude squared of a complex-valued inner product, but we will use the informal statement ‘inner product’ as in [14].

comprehensive scenarios and problems than (P1) and (P2), the aim of maximizing the channel gain has not attracted attention due to the non-convex and NP-hard constraints as well as the solutions can only approximate the global optimum solution.

A. Special Case of Achieving Global Optimum with MISO

In this section, we remark on a special case of a point-to-point downlink MISO communication scenario aided by an RIS, by considering a special case similar to [30] to achieve optimality, where the direct BS-UE link is blocked. With this, (2) can be rewritten as

$$y = \mathbf{h}_u^H \mathbf{WGx} + z, \quad (6)$$

where $\mathbf{x} \in \mathbb{C}^M$ is the transmit beamforming vector, $\mathbf{G} \in \mathbb{C}^{N \times M}$ is the equivalent channel of the BS-RIS link, and $\mathbf{W} = \text{diag}(\mathbf{w})$. The channel from RIS to UE is denoted by $\mathbf{h}_u \in \mathbb{C}^N$. Taking the far-field regime into account with a physical model that is based on angle directions [31], the received signal power in equation (6), i.e., $|y|^2$, can be expressed as

$$|y|^2 = |\mathbf{b} \mathbf{WCx}|^2 \quad (7)$$

where $\mathbf{C} = [\mathbf{a}(\vartheta_{b,1}, \varphi_{b,1}), \mathbf{a}(\vartheta_{b,2}, \varphi_{b,2}), \dots, \mathbf{a}(\vartheta_{b,M}, \varphi_{b,M})]$, $\mathbf{b} = \mathbf{a}^H(\vartheta_u, \varphi_u)$, with $\{\vartheta_u, \varphi_u\}$ is the departing direction of the reflected signal and $\{\vartheta_{b,m}, \varphi_{b,m}\}, m = 1, \dots, M$ is the arriving signal direction from BS antenna m . Here, $\mathbf{a}(\vartheta, \varphi)$ is the array response vector defined in (9) in the next section.

Similar to the analysis in the Appendix of the fifth version of [14], equation (7) can be rewritten as

$$\begin{aligned} |y|^2 &= |\mathbf{b} \mathbf{WCx}|^2 \\ &= \mathbf{w}^H (\mathbf{P}^T \odot \mathbf{R}) \mathbf{w} \\ &= \mathbf{w}^H \mathbf{Q} \mathbf{w}, \end{aligned} \quad (8)$$

where $\mathbf{Q} = \mathbf{P}^T \odot \mathbf{R}$, $\mathbf{P} = \mathbf{C} \mathbf{x} \mathbf{x}^H \mathbf{C}^H$, and $\mathbf{R} = \mathbf{b}^H \mathbf{b}$. Since \mathbf{Q} is a semi-positive definite matrix with rank one, maximizing (8) boils down to maximizing $|\mathbf{z}^H \mathbf{w}|$, where \mathbf{z} is the eigenvector associated with the maximum eigenvalue of \mathbf{Q} .

In this paper, one of the primary objectives is to address the problem in (P1), equivalently (P2), with a highly efficient algorithm. In fact, to the best of our knowledge, our proposed algorithms are the fastest converging to the global optimum in the literature to perform discrete beamforming optimization for (P1), and a general K -ary discrete quadratic programming. The extension of this framework with *elementwise updates* for a general MISO scenario is left as a future work. In Section IV, we define the problem formally and then introduce several algorithms to solve it.

III. CHANNEL MODEL

This section describes the channel model we employed. The RIS is placed such that the origin is at the first row and the first column of the RIS, as shown in Fig. 1. We let $\{\vartheta_b, \varphi_b\}$ and $\{\vartheta_u, \varphi_u\}$ pairs be the elevation and azimuth angles for

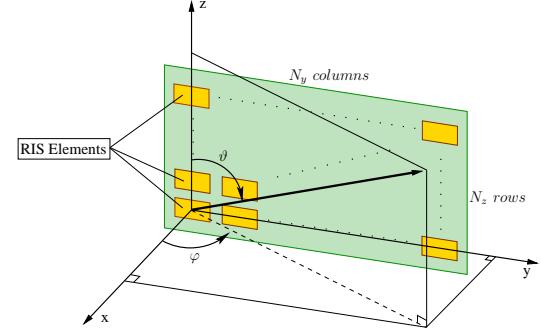


FIGURE 1. RIS structure.

the BS and the UE. The array response vector for the RIS is calculated by considering two uniform linear arrays (ULA) along the y -axis and the z -axis with array response vectors $\mathbf{a}_y(\vartheta, \varphi)$ and $\mathbf{a}_z(\vartheta)$, which are calculated as

$$\begin{aligned} \mathbf{a}_y &= \left[1, e^{-j2\pi \frac{d_y}{\lambda} \sin \vartheta \sin \varphi}, \dots, e^{-j2\pi(N_y-1) \frac{d_y}{\lambda} \sin \vartheta \sin \varphi} \right]^T \\ \mathbf{a}_z(\vartheta) &= \left[1, e^{-j2\pi \frac{d_z}{\lambda} \cos \vartheta}, \dots, e^{-j2\pi(N_z-1) \frac{d_z}{\lambda} \cos \vartheta} \right]^T \end{aligned}$$

where $d_y \leq \lambda/2$ and $d_z \leq \lambda/2$ and where we used the shorthand notation \mathbf{a}_y for $\mathbf{a}_y(\vartheta, \varphi)$. In this formulation, d_y and d_z are the spacing between the elements in respective axes, and λ is the wavelength of the incident signal. Therefore, for the described uniform planar array (UPA) structure, the array response vector can be calculated as

$$\mathbf{a}(\vartheta, \varphi) = \mathbf{a}_y(\vartheta, \varphi) \otimes \mathbf{a}_z(\vartheta) \quad (9)$$

where \otimes is the Kronecker product and $\mathbf{a}(\vartheta, \varphi) \in \mathbb{C}^N$. Note that, in $\mathbf{a}(\vartheta, \varphi)$, elements are ordered column-wise.

We employ the channel model from [10], [32], where the BS-RIS link \mathbf{h}_b undergoes Rician fading:

$$\mathbf{h}_b = \text{PL}_b \left(\sqrt{\frac{\kappa}{1+\kappa}} \mathbf{h}_b^{\text{LOS}} + \sqrt{\frac{1}{1+\kappa}} \mathbf{h}_b^{\text{NLOS}} \right) \quad (10)$$

where $\text{PL}_b = 30 + 22 \log_{10}(d_b)$ is the path loss (in dB) in the BS-RIS link with d_b being the distance between the BS and RIS, the LOS component is calculated with the RIS steering vectors given the angle-of-arrival (AoA) information of the BS, i.e., $\{\vartheta_b, \varphi_b\}$, with equation (9) as $\mathbf{h}_b^{\text{LOS}} = \mathbf{a}(\vartheta_b, \varphi_b)$. The NLOS component, $\mathbf{h}_b^{\text{NLOS}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N \times N})$, consists of circularly symmetric complex Gaussian random variables and κ is the Rician factor. The RIS-UE link \mathbf{h}_u also undergoes Rician fading in (10) with the same parameters except for the distance and the AoA information, i.e., d_u and $\{\vartheta_u, \varphi_u\}$, respectively.

The developed algorithms do not require the direct BS-UE link to be present to converge to the global optimum in the least number of steps. Whenever the direct link is not completely blocked, it follows Rayleigh fading, i.e., $h_0 = 10^{-\frac{PL_0}{20}} \times \rho_0$, where, $PL_0(d_0) = 32.6 + 36.7 \log_{10}(d_0)$ is the path loss (in dB) with d_0 being the BS-UE distance and $\rho_0 \sim \mathcal{CN}(0, 1)$ is a circularly symmetric complex Gaussian random variable.

To compare the performance and time complexity of our proposed algorithms with the existing methods from the literature with discrete phase shift constraints for the RIS, similar to [15], we consider NLOS transmission in BS-RIS and RIS-UE channels with $\kappa = 0$. The transmit power is 30 dBm with -90 dBm background noise power. Three-dimensional coordinate vectors $(-2, -1, 0)$, $(50, -200, 20)$, and $(0, 0, 0)$ are used as the locations of the RIS, BS, and the UE, respectively.

Finally, we remark that all of the developed algorithms can work not only with different Rician factors but also with arbitrary α_n selections. Therefore, the proposed algorithms can be applied to a general K -ary discrete Quadratic Program, where the objective is to maximize $\mathbf{x}^T \mathbf{Q} \mathbf{x}$, with $\text{rank}(\mathbf{Q}) = 1$, or equivalently, the inner product maximization of $|\mathbf{b}^T \mathbf{x}|^2$, as in (P2).

IV. PROBLEM DEFINITION

In this paper, we address the problem of finding the values $\theta_1, \theta_2, \dots, \theta_N$ to maximize $|h_0 + \sum_{n=1}^N h_n e^{j\theta_n}|$ where $\theta_n \in \Phi_K$ and $\Phi_K = \{\omega, 2\omega, \dots, K\omega\}$ with $\omega = \frac{2\pi}{K}$ and $j = \sqrt{-1}$. The set Φ_K can equivalently be described as $\{0, \omega, 2\omega, \dots, (K-1)\omega\}$. The values $h_n \in \mathbb{C}$, $n = 1, 2, \dots, N$ are the channel coefficients and θ_n are the phase values added to the corresponding h_n by a reconfigurable intelligent surface (RIS).

The problem can be formally described as

$$\begin{aligned} & \underset{\boldsymbol{\theta}}{\text{maximize}} \quad f(\boldsymbol{\theta}) \\ & \text{subject to } \theta_n \in \Phi_K, \quad n = 1, 2, \dots, N \end{aligned} \quad (11)$$

where

$$f(\boldsymbol{\theta}) = \frac{1}{\beta_0^2} \left| \beta_0 e^{j\alpha_0} + \sum_{n=1}^N \beta_n e^{j(\alpha_n + \theta_n)} \right|^2, \quad (12)$$

$h_n = \beta_n e^{j\alpha_n}$ for $n = 0, 1, \dots, N$, and $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_N)$. Also, g is defined as

$$g = h_0 + \sum_{n=1}^N h_n e^{j\theta_n^*} \quad (13)$$

where θ_n^* are the outcomes of the optimization in (11) and μ is defined as

$$\mu = \frac{g}{|g|}. \quad (14)$$

Note that, in (11), β_0^2 is a constant and therefore the maximization affects only the numerator. For that reason, the case with $h_0 = 0$ can be taken care of by maximizing the numerator only.

The solution to the problem can be achieved by making use of the following lemma.

Lemma: For an optimal solution $(\theta_1^*, \theta_2^*, \dots, \theta_N^*)$, it is necessary and sufficient that each θ_n^* satisfy

$$\theta_n^* = \arg \max_{\theta_n \in \Phi_K} \cos(\theta_n + \alpha_n - \underline{\mu}) \quad (15)$$

where $\underline{\mu}$ stands for the phase of μ in (14).

Proof: We can rewrite (13) as

$$|g| = \beta_0 e^{j(\alpha_0 - \underline{\mu})} + \sum_{n=1}^N \beta_n e^{j(\alpha_n + \theta_n^* - \underline{\mu})} \quad (16)$$

$$\begin{aligned} &= \beta_0 \cos(\alpha_0 - \underline{\mu}) + j \beta_0 \sin(\alpha_0 - \underline{\mu}) \\ &+ \sum_{n=1}^N \beta_n \cos(\theta_n^* + \alpha_n - \underline{\mu}) \\ &+ j \sum_{n=1}^N \beta_n \sin(\theta_n^* + \alpha_n - \underline{\mu}). \end{aligned} \quad (17)$$

Because $|g|$ is real-valued, the second and fourth terms in (17) sum to zero, and

$$|g| = \beta_0 \cos(\alpha_0 - \underline{\mu}) + \sum_{n=1}^N \beta_n \cos(\theta_n^* + \alpha_n - \underline{\mu}), \quad (18)$$

from which (15) follows as a necessary and sufficient condition for the lemma to hold. ■

Reference [15] attempts to decide a range of μ for which $\theta_n^* = k\omega$ must hold. Towards that end, it first defines a sequence of complex numbers with respect to each $n = 1, 2, \dots, N$ as

$$s_{nk} = e^{j(\alpha_n + (k-0.5)\omega)}, \quad \text{for } k = 1, 2, \dots, K. \quad (19)$$

Then, [15] defines, for any two points a and b on the unit circle C , $\text{arc}(a : b)$ to be the unit circular arc with a as the initial end and b as the terminal end in the counterclockwise direction; in particular, it defines $\text{arc}(a : b)$ as an open arc with the two endpoints a and b excluded. With this definition, [15] states the following proposition holds.

Proposition 1: A sufficient condition for $\theta_n^* = k\omega$ is

$$\mu \in \text{arc}(s_{nk} : s_{n,k+1}). \quad (20)$$

Proposition 1 is compatible with the lemma given above. To see this, assume μ satisfies (20). Then,

$$\underline{\mu} \in \left(\alpha_n + \left(k - \frac{1}{2} \right) \omega, \alpha_n + \left(k + \frac{1}{2} \right) \omega \right). \quad (21)$$

Since $\omega = \frac{2\pi}{K}$,

$$\alpha_n - \underline{\mu} \in \left((-2k-1) \frac{\pi}{K}, (-2k+1) \frac{\pi}{K} \right) \quad (22)$$

considering the reversal of order due to the subtraction of $\underline{\mu}$. Now, let $\theta_n = k\omega = 2k \frac{\pi}{K}$. Then

$$\theta_n + \alpha_n - \underline{\mu} \in \left(-\frac{\pi}{K}, \frac{\pi}{K} \right) \quad (23)$$

and thus $\cos(\theta_n + \alpha_n - \underline{\mu})$ is the largest among all other possibilities for θ_n because the slice $(-\frac{\pi}{K}, \frac{\pi}{K})$ corresponds to the largest values of the cosine function among all slices corresponding to different values of $\theta_n \in \Phi_K$ for $n = 1, 2, \dots, N$.

We remark that, as an open arc, $\text{arc}(a : b)$ does not contain the two endpoints a and b . This implies that we omit the situation where $\mu = s_{nk}$ for some n, k , or similarly when μ lies right in between $\text{arc}(s_{n,k-1} : s_{nk})$ and $\text{arc}(s_{nk} : s_{n,k+1})$. To justify this, assume $\mu = s_{nk}$ for some n, k , i.e., $\underline{\mu} = \alpha_n + (k - \frac{1}{2})\omega$. With this, the lemma in (15) results

in not one but two possible solutions, $\theta_{n,1}^* = k\omega$ and $\theta_{n,2}^* = (k-1)\omega$, for which $\theta_{n,1}^* + \alpha_n - \lfloor \mu \rfloor = \frac{\omega}{2}$ and $\theta_{n,2}^* + \alpha_n - \lfloor \mu \rfloor = -\frac{\omega}{2}$, respectively. Note that $\theta_{n,1}^*$ and $\theta_{n,2}^*$ cannot make $|g|$ in (17) real-valued at the same time, because $\sin(\theta_{n,1}^* + \alpha_n - \lfloor \mu \rfloor) = (-1) \sin(\theta_{n,2}^* + \alpha_n - \lfloor \mu \rfloor)$. Therefore only one of them can be the optimum selection, which we already consider separately for scenarios when $\mu \in \text{arc}(s_{n,k-1} : s_{nk})$ and $\mu \in \text{arc}(s_{nk} : s_{n,k+1})$, i.e., μ is right before and right after s_{nk} , respectively.

We note that g in (13) is defined for optimal phases $(\theta_1^*, \theta_2^*, \dots, \theta_N^*)$. The μ in (15), which comes from the definition in (14) which follows from (13), is the optimum one. On the other hand, in the rest of the paper, when we refer to μ , it is a value we are considering in search of the optimal μ .

A. UNIFORM POLAR QUANTIZATION

To address the discrete constraint on the RIS phase shifts, a straightforward approach is to project the relaxed continuous solution to the closest discrete value in the discrete phase shift set Φ_K . In [13], the authors named the discretization process of the continuous solutions as the closest point projection (CPP). Note that, by its definition, the CPP approach can be employed over any other algorithm that gives the relaxed solutions. Thus, the time-complexity of the discrete beamforming problem with CPP is dependent on the complexity of the algorithm that gives the preliminary continuous solution. For example, in [14], the authors used Discrete Manifold Optimization (Discrete-Manopt), which corresponds to quantization of the continuous phase shifts provided by Manifold Optimization, and the computational time complexity to achieve the quantized solutions is extremely high.

In this section, we will define a practical intuitive algorithm that can achieve a suboptimal solution, similar to the CPP approach in [13], which we call uniform polar quantization (UPQ). For this purpose, we can redefine the received power maximization problem with relaxed continuous phase shifts as follows:

$$\begin{aligned} & \underset{\boldsymbol{\theta}^{\text{cont}}}{\text{maximize}} \quad f_{\text{rx}}(\boldsymbol{\theta}^{\text{cont}}) \\ & \text{subject to } \theta_n^{\text{cont}} \in [0, 2\pi), \quad n = 1, 2, \dots, N \end{aligned} \quad (24)$$

where

$$f_{\text{rx}}(\boldsymbol{\theta}) = \left| \beta_0 e^{j\alpha_0} + \sum_{n=1}^N \beta_n e^{j(\alpha_n + \theta_n)} \right|^2. \quad (25)$$

Note that in the magnitude of the objective term $f_{\text{rx}}(\boldsymbol{\theta})$, we are adding complex numbers, or equivalently two-dimensional vectors on the complex-plane. Therefore, the bounds on the received power in (25) can simply be given as $0 \leq f_{\text{rx}}(\boldsymbol{\theta}) \leq \left(\sum_{n=0}^N \beta_n \right)^2$. In this case, assuming continuous phase shifts, the solution to the maximization problem would be to find θ_n^{cont} such that

$$\alpha_n + \theta_n^{\text{cont}} = \alpha_0, \quad \text{for } n = 1, 2, \dots, N. \quad (26)$$

TABLE 1. Approximation ratio of the UPQ Algorithm to the continuous solution.

	$K = 2$	$K = 3$	$K = 4$	$K = 6$	$K = 8$
$E_{\infty}(K)$	0.4053	0.6839	0.8106	0.9119	0.9496
$1/E_{\infty}(2)$	–	2.27 dB	3.01 dB	3.52 dB	3.70 dB

Therefore, by letting $\theta_n^{\text{cont}} = \alpha_0 - \alpha_n$ for $n = 1, 2, \dots, N$, we can select the discrete phase shifts with the following rule, which we refer to as UPQ in this paper:

$$\theta_n^{\text{UPQ}} = \left\lfloor \frac{\alpha_0 - \alpha_n}{\omega} \right\rfloor \times \omega, \quad \text{for } n = 1, 2, \dots, N, \quad (27)$$

where $\lfloor \cdot \rfloor$ is the rounding function defined as

$$\lfloor x \rfloor = \text{sgn}(x) \lfloor |x| + 0.5 \rfloor. \quad (28)$$

The importance of defining UPQ for the problem in (11), or equivalently for a discrete QP with the rank constraint is to present a fair comparison in the computational complexity results, where UPQ performs surprisingly well in terms of both performance and computational complexity given full CSI. We note that, in terms of full CSI, UPQ only requires α_n for $n = 0, 1, \dots, N$, and in a real scenario, look-up tables can be employed to further simplify the beamforming process.

Similar to UPQ, there are approaches used under different names in the literature. Firstly, as mentioned, [13] uses a similar approach for the CPP result where the authors show performance over estimated channels. Another example is, in [5], the authors derive a measurement based beamforming algorithm that is called the “Greedy Fast Beamforming Algorithm,” which is based on a similar quantization approach employed for the binary case.

We note that, a secondary result of equation (26) is to show why the optimal discrete phase shift selection problem was originally thought to be very difficult to solve. In (26), α_n and α_0 values are continuous. When θ_n^{cont} are constrained by the set Φ_K , it practically prevents a closed-form solution for θ_n to be available, because in (25) each θ_n is related to each other. Yet, in the following sections, our derivations with the optimal algorithm show that the optimum discrete phase shift selection problem turns out to be simpler, actually.

1) Efficiency Calculation for UPQ

In [13], the authors provide a lower bound on the performance of quantization. In this section, similar to the approach in [10], the performance loss in UPQ due to quantization is quantified with respect to K . For this purpose, the expected value of the normalized performance, i.e., E_N , is calculated for asymptotically large N . For large N ,

equation (25) with the UPQ solution can be rewritten as

$$\begin{aligned}
f_{\text{rx}}(\boldsymbol{\theta}^{\text{UPQ}}) &= \left| \beta_0 e^{j\alpha_0} + \sum_{n=1}^N \beta_n e^{j(\alpha_n + \theta_n^{\text{UPQ}})} \right|^2 \\
&= |e^{j\alpha_0}|^2 \left| \beta_0 + \sum_{n=1}^N \beta_n e^{j(\alpha_n + \theta_n^{\text{UPQ}} - \alpha_0)} \right|^2 \\
&= \left| \beta_0 + \sum_{n=1}^N \beta_n e^{j(\theta_n^{\text{UPQ}} - \theta_n^{\text{cont}})} \right|^2 \\
&\approx \left| \sum_{n=1}^N \beta_n e^{j(\theta_n^{\text{UPQ}} - \theta_n^{\text{cont}})} \right|^2,
\end{aligned} \tag{29}$$

where, for asymptotically large N , β_0 from the BS-UE direct link in (29) is practically discarded. Therefore, the received power for large N can be approximated as

$$\begin{aligned}
f_{\text{rx}}(\boldsymbol{\theta}^{\text{UPQ}}) &\approx \left| \sum_{n=1}^N \beta_n e^{j(\theta_n^{\text{UPQ}} - \theta_n^{\text{cont}})} \right|^2 = \sum_{n=1}^N \beta_n^2 \\
&+ 2 \sum_{k=2}^N \sum_{l=1}^{k-1} \beta_k \beta_l \cos((\theta_k^{\text{UPQ}} - \theta_k^{\text{cont}}) - (\theta_l^{\text{UPQ}} - \theta_l^{\text{cont}})).
\end{aligned} \tag{30}$$

Assume that in (30) all $\beta_k, \beta_l, \theta_k$, and θ_l are independent from each other. Let $\delta_i = \theta_i^{\text{UPQ}} - \theta_i^{\text{cont}}$ for $i = 1, \dots, N$, so that in (30), the argument of the cosine is equal to $\delta_k - \delta_l$. From equation (27), it follows that $\delta_i \in [-\frac{\pi}{K}, \frac{\pi}{K}]$. Assume δ_k and δ_l are i.i.d. uniform random variables in $[-\frac{\pi}{K}, \frac{\pi}{K}]$, i.e., $\delta_k, \delta_l \sim \mathcal{U}[-\frac{\pi}{K}, \frac{\pi}{K}]$, which results in $\mathbb{E}[\cos(\delta_m - \delta_n)] = \text{sinc}^2(\frac{1}{K})$. Therefore, the expected value of $f_{\text{rx}}(\boldsymbol{\theta}^{\text{UPQ}})$ can be defined as

$$\mathbb{E}[f_{\text{rx}}(\boldsymbol{\theta}^{\text{UPQ}})] = N\mathbb{E}[\beta_n^2] + N(N-1)\mathbb{E}[\beta_k \beta_l] \text{sinc}^2\left(\frac{1}{K}\right). \tag{31}$$

With this, we normalized the received power by the expected value of the maximum achievable power, i.e., $\mathbb{E}[(\sum_{n=1}^N \beta_n)^2] = N\mathbb{E}[\beta_n^2] + N(N-1)\mathbb{E}[\beta_k \beta_l]$, which gives

$$E_N(K) \triangleq \frac{N\mathbb{E}[\beta_n^2] + N(N-1)\mathbb{E}[\beta_k \beta_l] \text{sinc}^2(\frac{1}{K})}{N\mathbb{E}[\beta_n^2] + N(N-1)\mathbb{E}[\beta_k \beta_l]}. \tag{32}$$

Taking $N \rightarrow \infty$, the expected value of the normalized performance can be represented as a function of K as

$$E_\infty(K) = \text{sinc}^2\left(\frac{1}{K}\right), \tag{33}$$

which quantifies the effect of K , i.e., the number of available discrete phase shift selections, as an approximation to the optimal continuous solution. For example, as $K \rightarrow \infty$, i.e., the continuous case, $E_\infty(\infty) = 1$, therefore $E_\infty(K)$ also serves as an approximation ratio to the upper bound. Two examples for different selections of K are given in Table 1, where it can be observed that gains from using at least $K = 4$ discrete phases over $K = 2$ are significant.

A final emphasis we want to make is to show that the optimal result can be achieved as a quantization of the relaxed solution. In other words, the optimal solution

θ_n^* given μ using UPQ, i.e., $\theta_{n|\mu}^{\text{UPQ}}$, can be achieved by employing the UPQ algorithm with $\theta_n^{\text{cont}} = \lfloor \mu - \alpha_n \rfloor$. To see this, assume μ satisfies (20). From (22), we know that $\lfloor \mu - \alpha_n \rfloor \in ((2k-1)\frac{\pi}{K}, (2k+1)\frac{\pi}{K})$, in which case the quantization in UPQ gives

$$\theta_{n|\mu}^{\text{UPQ}} = \left\lfloor \frac{\lfloor \mu - \alpha_n \rfloor}{\omega} \right\rfloor \times \omega = k\omega, \tag{34}$$

for $n = 1, \dots, N$. Note that the quantization result in (34) is compatible with Proposition 1, which proves that the optimum result is actually a quantization solution. As the two approaches (equation (15) and quantization) are akin, UPQ can provide close-to-optimal solution given full CSI knowledge. However, as μ is unknown and one needs to search for the optimal μ , we present this as an insight on the problem rather than an operational idea.

Finally, although both CPP and UPQ give the quantized continuous solution without modification, they have a basic difference. CPP is applied over any continuous solution and defined by $\theta_n^{\text{CPP}} = \arg \min_{\theta_n \in \Phi_K} |\theta_n - \theta_n^{\text{cont}}|$ [13]. On the other hand, UPQ is specifically defined for the problem in (11) by using the rounding operator in equation (28) to determine θ_n^{UPQ} directly from α_n and α_0 . Using the $\arg \min$ operator not only gives marginal insight on the problem of interest, but it can lead to incorrect results for specific values of $\theta_n \in [0, 2\pi)$ and $\theta_n^{\text{cont}} \in [0, 2\pi)$.

In the next section, we will define our discrete phase shift selection algorithm that guarantees the global optimal solution. We further improve it in the sequel to converge with the least number of steps, and achieve significantly lower computational complexity.

V. A NEW ALGORITHM

Reference [15, Algorithm 1], as published, employs the criterion

$$\theta_n^* = \arg \min_{\theta_n \in \Phi_K} |(\theta_n + \alpha_n - \lfloor \mu \rfloor) \bmod 2\pi| \tag{35}$$

instead of (15)². We now specify Algorithm 1 as an alternative to [15, Algorithm 1]³.

We present the cumulative distribution function (CDF) results for SNR Boost [15] in Fig. 2 for $K = 2$, and in Fig. 3 for $K = 4$. Similarly, the ergodic rate results are presented in Fig. 4 and 5, respectively. The CDF results are presented for $N = 16, 64$, and 256 , using the average of 10,000 realizations of the channel model defined in Section

²In this paper, we define the mod function (the modulus function or the modulo operation) $x \bmod y$ as the remainder after the dividend $x > 0$ is divided by the divisor $y > 0$. We write it as $x \bmod y$, $x \bmod y$, or $\text{mod}(x, y)$. For $x < 0$ and $y > 0$, we use the convention that the remainder should always be the smallest such nonnegative number.

³In Algorithm 1 we define $(\theta \pm \omega) \bmod \Phi_K$ as follows. First note that the two sets $\{0, \omega, 2\omega, \dots, (K-1)\omega\}$ and $\{\omega, 2\omega, 3\omega, \dots, K\omega\}$ have the same members since $\omega = 2\pi/K$. Then, $(\theta + \omega) \bmod \Phi_K$ can be defined as $(\theta \pm \omega) \bmod \Phi_K \triangleq ((k \pm 1) \bmod K) \omega$.

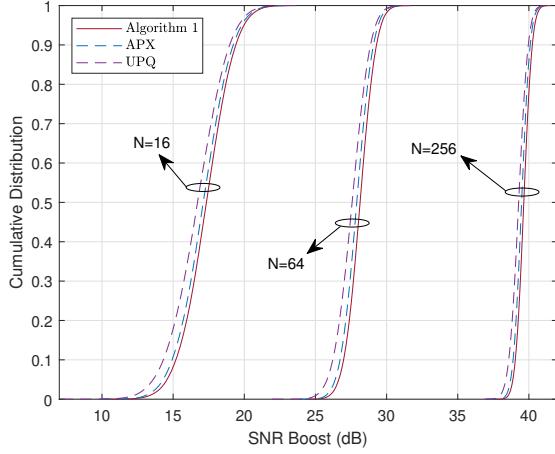


FIGURE 2. CDF plots for SNR Boost [15] with Uniform Polar Quantization (UPQ), Algorithm 1, and Approximation (APX) Algorithm [13], $K = 2$.

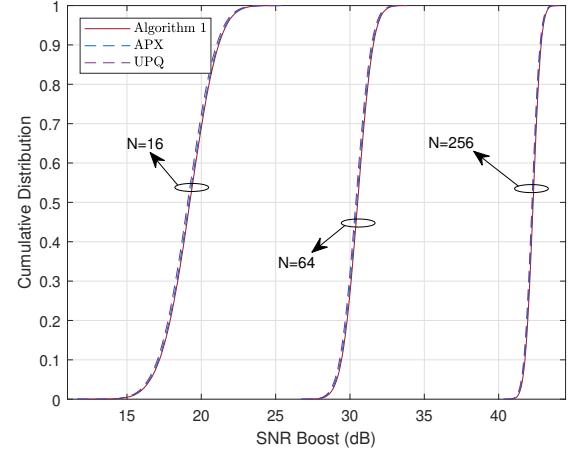


FIGURE 3. CDF plots for SNR Boost [15] with Uniform Polar Quantization (UPQ), Algorithm 1, and Approximation (APX) Algorithm [13], $K = 4$.

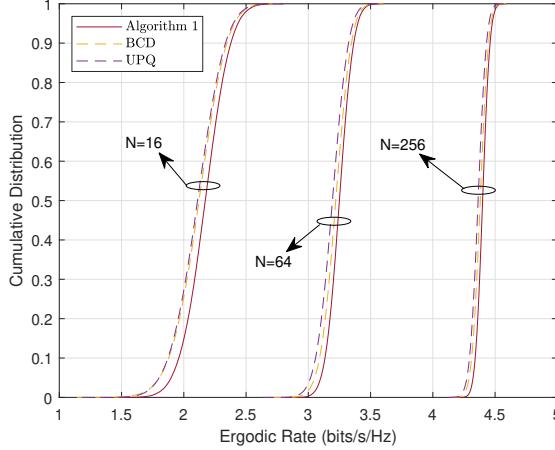


FIGURE 4. CDF plots for Ergodic Rate (3) with Uniform Polar Quantization (UPQ), Algorithm 1, and block coordinate descend (BCD) [33], $K = 2$.

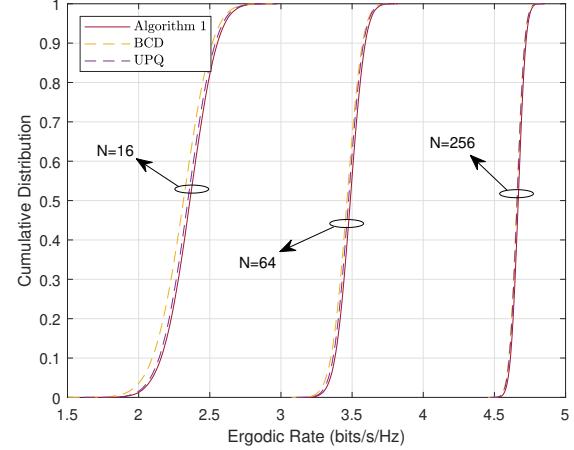


FIGURE 5. CDF plots for Ergodic Rate (3) with Uniform Polar Quantization (UPQ), Algorithm 1, and block coordinate descend (BCD) [33], $K = 4$.

III with $\kappa = 0$, where we employed UPQ, Algorithm 1, and the Approximation (APX) algorithm from [13], where all algorithms ran over the same realization in each step. For the ergodic rate results, instead of APX algorithm, we employed the block coordinate descend (BCD) [33] as a typical benchmark where the phase shifts are optimized for each RIS element at a time. Although the gains are not large, especially with $K = 4$ in Fig. 3, these two figures serve as a verification of the optimality of Algorithm 1, which we already know from the analysis presented in this paper. Besides, in Fig. 6, we focus on the low SNR Boost regime and compare the 1st percentile results. For $K = 2$, while there is only about 1 dB loss of using UPQ, it cannot be recovered by increasing the number of RIS elements. On the other hand, for $K \geq 4$, both APX and UPQ can provide close-to-optimal results, with UPQ being quite efficient in terms of complexity, which we discuss in Section XIII.

Finally, the normalized received power results are calculated by $|\beta_0 e^{j\alpha_0} + \sum_{n=1}^N \beta_n e^{j(\alpha_n + \theta_n)}|^2 / (\sum_{n=0}^N \beta_n)^2$ and plotted in Fig. 7 for $\kappa = 10$. The figure verifies that the expected value approximation for large N , i.e., $E_\infty(K)$ results in Table 1, fall in line with the numerical results.

In the Appendix, we discuss an alternative way to initialize Algorithm 1, which significantly reduces the computational complexity. We will use this technique in initializing Algorithm 2 and Algorithm 3 in the sequel.

VI. ALGORITHM CONVERGENCE: TOWARDS TWO NEW ALGORITHMS

We will now show the periodicity in the update rule in Algorithm 1, i.e., in $\mathcal{N}(\lambda_l)$. With this observation, we will prove that Algorithm 1 takes N or fewer steps to converge, as opposed to the statement that [15, Algorithm 1] takes KN or $2N$ steps on average [15]. This will result in two new versions of the algorithm, i.e., Algorithm 2 and

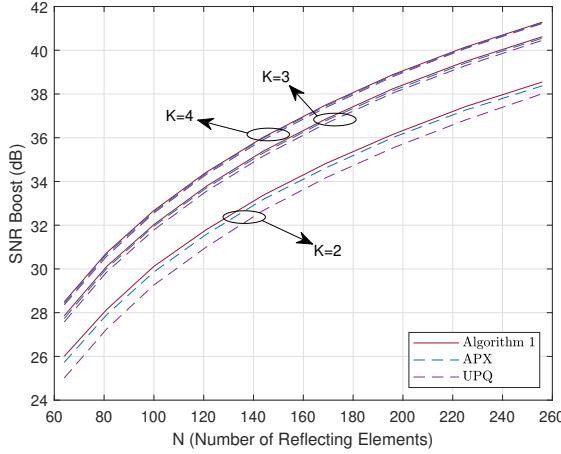


FIGURE 6. 1st percentile SNR Boost results vs. N , for $K \in \{2, 3, 4\}$.

Algorithm 1 Update for [15, Algorithm 1]

```

1: Initialization: Compute  $s_{nk} = e^{j(\alpha_n + (k-0.5)\omega)}$  for  $n = 1, 2, \dots, N$  and  $k = 1, 2, \dots, K$ .
2: Eliminate duplicates among  $s_{nk}$  and sort to get  $e^{j\lambda_l}$  such that  $0 \leq \lambda_1 < \lambda_2 < \dots < \lambda_L < 2\pi$ .
3: Let, for  $l = 1, 2, \dots, L$ ,  $\mathcal{N}(\lambda_l) = \{n | s_{nk} = \lambda_l\}$ .
4: Set  $\underline{\mu} = 0$ . For  $n = 1, 2, \dots, N$ , calculate  $\theta_n = \arg \max_{\theta_n \in \Phi_K} \cos(\theta_n + \alpha_n - \underline{\mu})$ .
5: Set  $g_0 = h_0 + \sum_{n=1}^N h_n e^{j\theta_n}$ ,  $\text{absmax} = |g_0|$ .
6: for  $l = 1, 2, \dots, L-1$  do
7:   For each  $n \in \mathcal{N}(\lambda_l)$ , let  $(\theta_n + \omega \leftarrow \theta_n) \bmod \Phi_K$ .
8:   Let

$$g_l = g_{l-1} + \sum_{n \in \mathcal{N}(\lambda_l)} h_n (e^{j\theta_n} - e^{j(\theta_n - \omega) \bmod \Phi_K})$$

9:   if  $|g_l| > \text{absmax}$  then
10:    Let  $\text{absmax} = |g_l|$ 
11:    Store  $\theta_n$  for  $n = 1, 2, \dots, N$ 
12:   end if
13: end for
14: Read out  $\theta_n^*$  as the stored  $\theta_n$ ,  $n = 1, 2, \dots, N$ .

```

Algorithm 3, with a simple elementwise update rule, in the sequel. Towards this end, we first make the following statement.

Claim 1: As in (19), set $s_{nk} = e^{j(\alpha_n + (k-\frac{1}{2})\frac{2\pi}{K})}$, $n = 1, 2, \dots, N$, $k = 1, 2, \dots, K$, $\alpha_n \in [0, 2\pi)$. Let $\lambda_l = \underline{s}_{nk}$ such that $0 \leq \lambda_1 < \lambda_2 < \dots < \lambda_L < 2\pi$. Let $\mathcal{N}(\lambda_l) = \{n | \lambda_l = \underline{s}_{nk}\}$. Assuming for now that $|\mathcal{N}(\lambda_l)| = 1$, $l = 1, 2, \dots, L = NK$, which we will relax in the sequel, we claim that $\mathcal{N}(\lambda_{l'}) = \mathcal{N}(\lambda_{l'+N})$ for $l' = 1, 2, \dots, N(K-1)$.

To prove *Claim 1*, we will first introduce *Claim 2* and prove it.

Claim 2: Without loss of generality, we can assume that $\alpha_n < \frac{2\pi}{K}$, $n = 1, 2, \dots, N$.

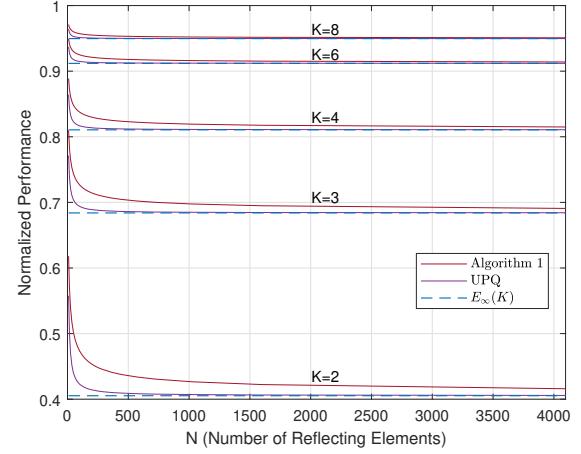


FIGURE 7. Normalized Performance results vs. N , for $\kappa = 10$ and $K \in \{2, 3, 4, 6, 8\}$.

Proof of Claim 2: Suppose that for some n , we have $\frac{2\pi}{K} \cdot m \leq \alpha_n < \frac{2\pi}{K} \cdot (m+1)$, $m = 1, 2, \dots, K-1$. Let $\beta_n \triangleq \alpha_n - m \cdot \frac{2\pi}{K}$, so that $\beta_n < \frac{2\pi}{K}$. We will write below each phase value in (19), $\alpha_n + \frac{(2k-1)\pi}{K}$, for $k = 1, 2, \dots, K$ (note that the $(\bmod 2\pi)$ notation below applies to both sides of the equation).

- $k = 1$:

$$\alpha_n + \frac{\pi}{K} = \beta_n + \frac{(2m+1)\pi}{K} \pmod{2\pi}$$
where $\left(\frac{(2m+1)\pi}{K}\right)_{m=1}^{K-1} = \left\{\frac{3\pi}{K}, \frac{5\pi}{K}, \dots, \frac{(2K-1)\pi}{K}\right\}$,
- $k = 2$:

$$\alpha_n + \frac{3\pi}{K} = \beta_n + \frac{(2m+3)\pi}{K} \pmod{2\pi}$$

$$= \begin{cases} \beta_n + \frac{(2m+3)\pi}{K} \pmod{2\pi}, & m \leq K-2 \\ \beta_n + \frac{\pi}{K} \pmod{2\pi}, & m = K-1 \end{cases}$$
where $\left(\frac{(2m+3)\pi}{K}\right)_{m=1}^{K-2} = \left\{\frac{5\pi}{K}, \frac{7\pi}{K}, \dots, \frac{(2K-1)\pi}{K}\right\}$,
 \vdots
- $k = K$:

$$\alpha_n + \frac{(2K-1)\pi}{K} = \beta_n + \frac{(2m+2K-1)\pi}{K} \pmod{2\pi}$$

$$= \beta_n + \frac{(2m-1)\pi}{K} \pmod{2\pi}$$
where $\left(\frac{(2m-1)\pi}{K}\right)_{m=1}^{K-1} = \left\{\frac{\pi}{K}, \frac{3\pi}{K}, \dots, \frac{(2K-3)\pi}{K}\right\}$.

Thus, if there is an $\alpha_n \geq \frac{2\pi}{K}$ to generate K phase values, there is always a β_n , $\beta_n < \frac{2\pi}{K}$ with which one can generate the same K phase values in a similar fashion. Therefore, in order to prove *Claim 1*, one can work with the assumption that $\alpha_n < \frac{2\pi}{K}$, for $n = 1, 2, \dots, N$. \blacksquare

Proof of Claim 1: Assuming $0 \leq \alpha_1 < \alpha_2 < \dots < \alpha_N < \frac{2\pi}{K}$, without loss of generality, we will now show that $\mathcal{N}(\lambda_{l'}) = \mathcal{N}(\lambda_{l'+N})$ for $l' = 1, 2, \dots, N(K-1)$. For this, there are $N+1$ cases to consider.

Case 0: In this case, we assume $\alpha_n < \frac{\pi}{K}$, $n = 1, 2, \dots, N$. We write all possible values of \underline{s}_{nk} as follows.

$n = 1$:

$$\underline{s}_{1k} \in \left\{ \alpha_1 + \frac{\pi}{K}, \alpha_1 + \frac{3\pi}{K}, \dots, \alpha_1 + \frac{(2K-1)\pi}{K} \right\}, \quad (36)$$

$n = 2:$

$$\underline{s}_{2k} \in \left\{ \alpha_2 + \frac{\pi}{K}, \alpha_2 + \frac{3\pi}{K}, \dots, \alpha_2 + \frac{(2K-1)\pi}{K} \right\}, \quad (37)$$

⋮

$n = N:$

$$\underline{s}_{Nk} \in \left\{ \alpha_N + \frac{\pi}{K}, \alpha_N + \frac{3\pi}{K}, \dots, \alpha_N + \frac{(2K-1)\pi}{K} \right\}. \quad (38)$$

Sorting (36)–(38), we have

$$\begin{aligned} \alpha_1 + \frac{\pi}{K} &< \alpha_2 + \frac{\pi}{K} < \dots < \alpha_N + \frac{\pi}{K} \\ &< \alpha_1 + \frac{3\pi}{K} < \alpha_2 + \frac{3\pi}{K} < \dots < \alpha_N + \frac{3\pi}{K} \\ &\vdots \\ &< \alpha_1 + \frac{(2K-1)\pi}{K} < \alpha_2 + \frac{(2K-1)\pi}{K} < \dots < \alpha_N + \frac{(2K-1)\pi}{K}. \end{aligned}$$

Thus,

$$(\mathcal{N}(\lambda_l))_{l=1}^{L=NK} = \left\{ \underbrace{1, 2, \dots, N}_1, \underbrace{1, 2, \dots, N}_2, \dots, \underbrace{1, 2, \dots, N}_K \right\}.$$

Therefore, for Case 0 and for $l' = 1, 2, \dots, N(K-1)$, we have $\mathcal{N}(\lambda_{l'}) = \mathcal{N}(\lambda_{l'+N})$.

There are N remaining cases. We will discuss these cases as *Case i* where $i = 1, 2, \dots, N$.

Case i, ($i = 1, 2, \dots, N$): In *Case i*, we have i occurrences of $\alpha_n > \frac{\pi}{K}$ as follows.

$$\begin{aligned} 0 \leq \alpha_1 &< \alpha_2 < \dots < \alpha_{N-i} < \frac{\pi}{K} \\ &\leq \alpha_{N-i+1} < \dots < \alpha_N < \frac{2\pi}{K}. \end{aligned} \quad (39)$$

We write all possible values of \underline{s}_{nk} as follows.

$n = 1:$

$$\begin{aligned} \underline{s}_{1k} \in & \left\{ \alpha_1 + \frac{\pi}{K}, \alpha_1 + \frac{3\pi}{K}, \dots, \right. \\ & \left. \alpha_1 + \frac{(2K-3)\pi}{K}, \alpha_1 + \frac{(2K-1)\pi}{K} \right\}, \end{aligned} \quad (40)$$

⋮

$n = N-i:$

$$\begin{aligned} \underline{s}_{(N-i)k} \in & \left\{ \alpha_{N-i} + \frac{\pi}{K}, \alpha_{N-i} + \frac{3\pi}{K}, \dots, \right. \\ & \left. \alpha_{N-i} + \frac{(2K-3)\pi}{K}, \alpha_{N-i} + \frac{(2K-1)\pi}{K} \right\}, \end{aligned} \quad (41)$$

$n = N-i+1:$

$$\begin{aligned} \underline{s}_{(N-i+1)k} \in & \left\{ \alpha_{N-i+1} + \frac{\pi}{K}, \alpha_{N-i+1} + \frac{3\pi}{K}, \dots, \right. \\ & \left. \alpha_{N-i+1} + \frac{(2K-3)\pi}{K}, \alpha_{N-i+1} - \frac{\pi}{K} \right\}, \end{aligned} \quad (42)$$

$n = N-i+2:$

$$\begin{aligned} \underline{s}_{(N-i+2)k} \in & \left\{ \alpha_{N-i+2} + \frac{\pi}{K}, \alpha_{N-i+2} + \frac{3\pi}{K}, \dots, \right. \\ & \left. \alpha_{N-i+2} + \frac{(2K-3)\pi}{K}, \alpha_{N-i+2} - \frac{\pi}{K} \right\}, \end{aligned} \quad (43)$$

⋮

$n = N:$

$$\begin{aligned} \underline{s}_{Nk} \in & \left\{ \alpha_N + \frac{\pi}{K}, \alpha_N + \frac{3\pi}{K}, \dots, \right. \\ & \left. \alpha_N + \frac{(2K-3)\pi}{K}, \alpha_N - \frac{\pi}{K} \right\}. \end{aligned} \quad (44)$$

Sorting (40)–(44), we have

$$\begin{aligned} \alpha_{N-i+1} - \frac{\pi}{K} &< \dots < \alpha_N - \frac{\pi}{K} \\ &< \alpha_1 + \frac{\pi}{K} < \dots < \alpha_{N-i} + \frac{\pi}{K} \\ &< \alpha_{N-i+1} + \frac{\pi}{K} < \dots < \alpha_N + \frac{\pi}{K} \\ &< \alpha_1 + \frac{3\pi}{K} < \dots < \alpha_{N-i} + \frac{3\pi}{K} \\ &< \alpha_{N-i+1} + \frac{3\pi}{K} < \dots < \alpha_N + \frac{3\pi}{K} \\ &< \alpha_1 + \frac{5\pi}{K} < \dots < \alpha_{N-i} + \frac{5\pi}{K} \\ &\vdots \\ &< \alpha_{N-i+1} + \frac{(2K-3)\pi}{K} < \dots < \alpha_N + \frac{(2K-3)\pi}{K} \\ &< \alpha_1 + \frac{(2K-1)\pi}{K} < \dots < \alpha_{N-i} + \frac{(2K-1)\pi}{K}. \end{aligned}$$

Thus,

$$\begin{aligned} (\mathcal{N}(\lambda_l))_{l=1}^{L=NK} = & \left\{ \underbrace{N-i+1, \dots, N, 1, 2, \dots, N-i}_1, \right. \\ & \underbrace{N-i+1, \dots, N, 1, 2, \dots, N-i}_2, \dots, \\ & \left. \underbrace{N-i+1, \dots, N, 1, 2, \dots, N-i}_K \right\} \end{aligned} \quad (45)$$

for $i = 1, 2, \dots, N$. Therefore, for *Case i*, $i = 1, 2, \dots, N$, and for $l' = 1, 2, \dots, N(K-1)$, we have $\mathcal{N}(\lambda_{l'}) = \mathcal{N}(\lambda_{l'+N})$. With this, Claim 1 is proved. \blacksquare

Algorithm 2 Simplified Algorithm 1 with $|\mathcal{N}(\lambda_l)| = 1$ for all l

```

1: Initialization: Set  $\underline{\mu} = \alpha_0 - \frac{\pi}{K}$ 
2: Compute  $\varphi_n = (\alpha_n - \alpha_0) \bmod \frac{2\pi}{K}$ ,  $n = 1, 2, \dots, N$ 
3: Sort  $\varphi_n$  such that  $0 \leq \varphi_{n_1} < \varphi_{n_2} < \dots < \varphi_{n_N} < \frac{2\pi}{K}$ 
4: Set  $\theta_n = \operatorname{argmax}_{\theta'_n \in \Phi_K} \cos(\theta'_n + \alpha_n - \underline{\mu})$ , store  $\theta_n$ ,  $n = 1, 2, \dots, N$ 
5: Set  $g_0 = h_0 + \sum_{n=1}^N h_n e^{j\theta_n}$ ,  $\text{absmax} = |g_0|$ 
6: for  $l = 1, 2, \dots, N$  do
7:   Let  $(\theta_{n_l} + \omega \leftarrow \theta_{n_l}) \bmod \Phi_K$ 
8:   Let
      
$$g_l = g_{l-1} + h_{n_l} \left( e^{j\theta_{n_l}} - e^{j(\theta_{n_l} - \omega) \bmod \Phi_K} \right)$$

9:   if  $|g_l| > \text{absmax}$  then
10:    Let  $\text{absmax} = |g_l|$ 
11:    Store updated  $\theta_{n_l}$ 
12:   end if
13: end for
14: Read out  $\theta_n^*$  as the stored  $\theta_n$ ,  $n = 1, 2, \dots, N$ .

```

VII. N STEPS SUFFICE WHEN $|\mathcal{N}(\lambda_l)| = 1$ FOR ALL l

Given $|\mathcal{N}(\lambda_l)| = 1$ and

$$\mathcal{N}(\lambda_{l'}) = \mathcal{N}(\lambda_{l'+N}), \quad l' = 1, 2, \dots, N(K-1), \quad (46)$$

we want to show that N steps will suffice for convergence. Now, consider the main problem of maximizing $|h_0 + \sum_{n=1}^N h_n e^{j\theta_n}|$, where it is clear that our discrete phase shift selections can only tune the second term in the absolute value. Let

$$g_c \triangleq \sum_{n=1}^N h_n e^{j\theta_n} = \sum_{n=1}^N \beta_n e^{j(\alpha_n + \theta_n)}. \quad (47)$$

In each step of the Algorithm 1, we define

$$g_{c,l} \triangleq g_l - h_0, \quad l = 1, 2, \dots, L. \quad (48)$$

Note that, in (48), h_0 , g_l , and $g_{c,l}$ are complex numbers, with l being a generation index. From Proposition 1, we know that whenever μ is anywhere in $\operatorname{arc}(s_{nk} : s_{n,k+1})$, θ_n does not change. The angle θ_n only changes when μ changes from one arc to another, i.e.,

$$\mu \in \operatorname{arc}(e^{j\lambda_l} : e^{j\lambda_{l+1}}) \rightarrow \mu \in \operatorname{arc}(e^{j\lambda_{l+1}} : e^{j\lambda_{l+2}}) \quad (49)$$

in which case θ_n must be updated as

$$\theta_n \rightarrow \theta_n + \omega, \quad n \in \mathcal{N}(\lambda_{l+1}). \quad (50)$$

With (49)–(50), the naive approach in Algorithm 1 gathers all possibilities for g_l in NK steps by considering all possible arcs that μ can be in. To show that N steps will suffice, we want to point out the redundancy in those NK steps. Consider any consecutive N steps in Algorithm 1. In those steps, the phase shifts will be updated as $\theta_n \rightarrow \theta_n + \omega$ with $n \in (\mathcal{N}(\lambda_l))_{l=l'}^{l'+N-1}$, $l' = 1, 2, \dots, N(K-1) + 1$. Since we have (46), the following must hold

$$(\mathcal{N}(\lambda_l))_{l=l'}^{l'+N-1} = \{1, 2, \dots, N\}, \quad (51)$$

which says that after *any* N consecutive steps in Algorithm 1, $\theta_n \rightarrow \theta_n + \omega$, $n = 1, 2, \dots, N$. To proceed further, we need an intermediate result, which we discuss below.

Remark: Let $g_c = \sum_{n=1}^N \beta_n e^{j(\theta_n + \alpha_n)}$ be the cascaded channel term in (47). For any angle θ' , $g_c(\theta_n + \theta') = g_c(\theta_n) e^{j\theta'}$, resulting in $|g_c(\theta_1, \theta_2, \dots, \theta_N)| = |g_c(\theta_1 + \theta', \theta_2 + \theta', \dots, \theta_N + \theta')|$. This remark illustrates that the rotation by a phase shift of an arbitrary angle θ' does not change the received power. Therefore, if $\theta_n \rightarrow \theta_n + \omega$ with $n = 1, 2, \dots, N$,

$$|g_{c,l'}| = |g_{c,l'+N}|, \quad l' = 1, 2, \dots, N(K-1) \quad (52)$$

must be true. Therefore, among the NK possibilities in Algorithm 1, there are only N unique values of $|g_c|$. Consequently, as the algorithm is tuning g_c to maximize $|g_c + h_0|$, it is sufficient to consider N arcs that are closest to h_0 .

The algorithm to implement when $|\mathcal{N}(\lambda_l)| = 1$ for $l = 1, 2, \dots, N$ is given under Algorithm 2. The initialization technique introduced in the Appendix for Algorithm 1 is employed in Algorithm 2.

VIII. FEWER THAN N STEPS SUFFICE WHEN $|\mathcal{N}(\lambda_l)| > 1$ FOR SOME l

With $\mathcal{N}(\lambda_l)$ given in Claim 1, for $|\mathcal{N}(\lambda_l)| > 1$ to be true for some l , consider a repetition among s_{nk} , i.e., assume there are n_1, n_2, k_1 , and k_2 such that $s_{n_1, k_1} = s_{n_2, k_2}$, i.e.,

$$\begin{aligned} \left(\alpha_{n_1} + \frac{(2k_1 - 1)\pi}{K} \right) \bmod 2\pi = \\ \left(\alpha_{n_2} + \frac{(2k_2 - 1)\pi}{K} \right) \bmod 2\pi. \end{aligned} \quad (53)$$

Equation (53) is possible only if $\beta_{n_1} = \beta_{n_2}$ as β_n are defined in Claim 1. Therefore, all K phase values represented by β_{n_1} and β_{n_2} must be equal, meaning there is an N' such that $M = (N - N')K$. Consequently, the problem of sorting s_{nk} according to their phase values with $0 \leq \alpha_1 < \alpha_2 < \dots < \alpha_N < \frac{2\pi}{K}$ reduces to $0 \leq \gamma_1 < \gamma_2 < \dots < \gamma_{\frac{M}{K} = N - N'} < \frac{2\pi}{K}$ for

$$s_{mk} = e^{j(\gamma_m + (k - \frac{1}{2})\frac{2\pi}{K})}, \quad m = 1, 2, \dots, \frac{M}{K}, \quad k = 1, 2, \dots, K \quad (54)$$

where $\gamma_m = \min\{\alpha_{n_m,1}, \alpha_{n_m,2}, \dots, \alpha_{n_m,G_m}\}$ such that $\beta_{n_m,1} = \beta_{n_m,2} = \dots = \beta_{n_m,G_m}$. So, this time, there are $\frac{M}{K} + 1 = N - N' + 1$ many cases.

For unique s_{mk} , let $\mathcal{M}(\lambda_l) = \{m | \lambda_l = \underline{s}_{mk}\}$. We know from (45) that the following must hold

$$(\mathcal{M}(\lambda_l))_{l=1}^{M=(N-N')K} = \left\{ \underbrace{\frac{M}{K} - i + 1, \dots, \frac{M}{K}}_1, 1, 2, \dots, \underbrace{\frac{M}{K} - i, \dots, \frac{M}{K}}_2, \dots, \underbrace{\frac{M}{K} - i + 1, \dots, \frac{M}{K}}_K, 1, 2, \dots, \frac{M}{K} - i \right\} \quad (55)$$

for $i = 1, 2, \dots, \frac{M}{K}$ where in each one of the K groups there are $\frac{M}{K} = N - N'$ elements. To calculate $\mathcal{N}(\lambda_l)$, we define the following sets

$$\begin{aligned} \mathcal{R}_m &= \{n_{m,1}, n_{m,2}, \dots, n_{m,G_m}\} \\ \gamma_m &= \min\{\alpha_{n_{m,1}}, \alpha_{n_{m,2}}, \dots, \alpha_{n_{m,G_m}}\}, \\ \beta_{n_m,1} &= \beta_{n_m,2} = \dots = \beta_{n_m,G_m} \end{aligned} \quad (56)$$

where $G_m = |\mathcal{R}_m|$ and $\bigcup_{m=1}^M \mathcal{R}_m = \{1, 2, \dots, N\}$ must hold. As a consequence, one can calculate $\mathcal{N}(\lambda_l) = \mathcal{R}_{\mathcal{M}(\lambda_l)}$. Therefore, the “update loop” in Algorithm 1 can be written as

$$(\mathcal{N}(\lambda_l))_{l=1}^{M=(N-N')K} = \left\{ \underbrace{\mathcal{R}_{\frac{M}{K}-i+1}, \dots, \mathcal{R}_{\frac{M}{K}}, \mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_{\frac{M}{K}-i}}_1, \dots, \underbrace{\mathcal{R}_{\frac{M}{K}-i+1}, \dots, \mathcal{R}_{\frac{M}{K}}, \mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_{\frac{M}{K}-i}}_2, \dots, \underbrace{\mathcal{R}_{\frac{M}{K}-i+1}, \dots, \mathcal{R}_{\frac{M}{K}}, \mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_{\frac{M}{K}-i}}_K \right\} \quad (57)$$

where the periodicity in the update rule still holds in (57), i.e., $\mathcal{N}(\lambda_{l'}) = \mathcal{N}(\lambda_{l'+\frac{M}{K}})$, $l' = 1, 2, \dots, \frac{M}{K}(K-1)$. With the new update rule, after any $\frac{M}{K}$ consecutive steps in Algorithm 1, the phase shift selections will be updated such that $\theta_n \rightarrow \theta_n + \omega$, $n = 1, 2, \dots, N$. This will result in

$$|g_{c,l'}| = \left| g_{c,l'+\frac{M}{K}} \right|, \quad l' = 1, 2, \dots, \frac{M}{K}(K-1). \quad (58)$$

Therefore, the sufficiency of $\frac{M}{K} = N - N'$ steps follows from (52) and the text that follows it in Section VII.

Algorithm 3 implements the technique described in this section. The initialization technique introduced in the Appendix for Algorithm 1 is employed in Algorithm 3.

Note that if the BS-UE link is completely blocked, i.e., $h_0 = 0$, the for loop in Step 7 can end at $l = \frac{M}{K} - 1 = N - N' - 1$, which is one fewer step to run Algorithm 3. This is because, we can guarantee in (58) that $|g_{c,l'}| = \left| g_{c,l'+\frac{M}{K}} \right|$,

Algorithm 3 Simplified Algorithm 1 where $|\mathcal{N}(\lambda_l)| > 1$ for some l

```

1: Initialization: Set  $\underline{\mu} = \alpha_0 - \frac{\pi}{K}$ 
2: Find  $\gamma_m$  and  $\mathcal{R}_m$  as in (54) and (56),  $m = 1, 2, \dots, \frac{M}{K}$ 
3: Compute  $\varphi_m = \gamma_m - \alpha_0 \pmod{\frac{2\pi}{K}}$ ,  $m = 1, 2, \dots, \frac{M}{K}$ 
4: Sort  $\varphi_m$  such that  $0 \leq \varphi_{m_1} < \varphi_{m_2} < \dots < \varphi_{m_{\frac{M}{K}}} < \frac{2\pi}{K}$ 
5: Set  $\theta_n = \operatorname{argmax}_{\theta'_n \in \Phi_K} \cos(\theta'_n + \alpha_n - \underline{\mu})$ , store  $\theta_n$ ,  $n = 1, 2, \dots, N$ 
6: Set  $g_0 = h_0 + \sum_{n=1}^N h_n e^{j\theta_n}$ ,  $\text{absmax} = |g_0|$ 
7: for  $l = 1, 2, \dots, \frac{M}{K} = N - N'$  do
8:   Let  $(\theta_n + \omega \leftarrow \theta_n) \pmod{\Phi_K}$   $n \in \mathcal{R}_{m_l}$ 
9:   Let
    
$$g_l = g_{l-1} + \sum_{n \in \mathcal{R}_{m_l}} h_n \left( e^{j\theta_n} - e^{j(\theta_n + \omega) \pmod{\Phi_K}} \right)$$

10:  if  $|g_l| > \text{absmax}$  then
11:    Let  $\text{absmax} = |g_l|$ 
12:    Store updated  $\theta_n$  for  $n \in \mathcal{R}_{m_l}$ 
13:  end if
14: end for
15: Read out  $\theta_n^*$  as the stored  $\theta_n$ ,  $n = 1, 2, \dots, N$ .

```

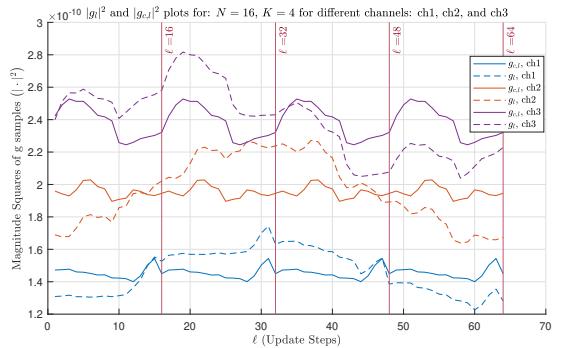


FIGURE 8. Variation of $|g_l|^2$ and $|g_{c,l}|^2$ with l .

whereas we cannot say right away that $|g_{c,l'} + h_0| = |g_{c,l'+\frac{M}{K}} + h_0|$ will be satisfied. In [15], the authors reduce the number of steps from KN to $2N$ exploiting h_0 . In this work, with the periodicity proof, we reduce to N or fewer steps whether the direct link is blocked or not.

Step 7 in Algorithm 2 and Step 8 in Algorithm 3 are such that the phase shifts updates are restricted to just one or just a few. These steps ensure that the running times of these algorithms are much less than those published in the literature, e.g., [14], [20].

Figure 8 shows the behaviors of $|g_l|^2$ and $|g_{c,l}|^2$ against l for a number of channel realizations at $N = 16$ and $K = 4$. The periodicity of $|g_{c,l}|^2$ is clearly observable in this figure. It is this periodicity that we take advantage of in reducing the number of steps for the algorithm to converge to N with a simple elementwise update rule in this paper.

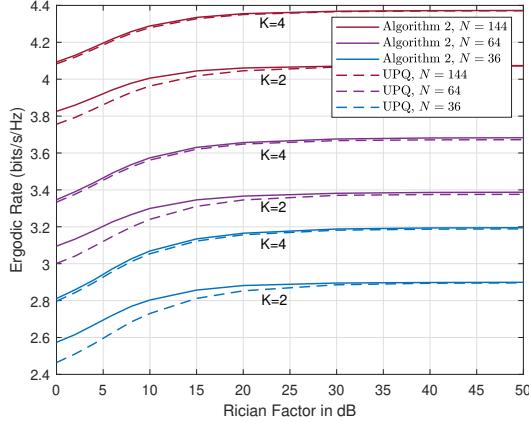


FIGURE 9. 1st Percentile Ergodic Rate (3) results vs. κ , for $K \in \{2, 4\}$

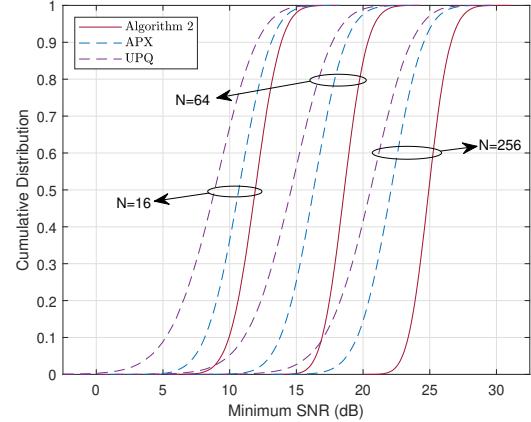


FIGURE 10. CDF of the minimum SNR across $U = 4$ users for $K = 2$ and $N \in \{16, 64, 256\}$.

IX. EFFECT OF RICIAN FACTOR

Using the given definition of the Rician channel model in equation (10), we extend the 1st percentile SNR Boost results in Fig. 6 by showing the 1st percentile ergodic rate performance results for $N \in \{36, 64, 144\}$, with different values of κ , in Fig. 9. The performance of UPQ against Algorithm 2 shows how closely UPQ can approximate the global optimum with increasing LOS gain, i.e., κ . Even with $K = 2$ and lower values of N , the UPQ performance gets significantly closer to Algorithm 2, unveiling the potential of the simple quantization approach, i.e., UPQ. We remark that, as κ increases, variation among each β_n realization decreases. This results in more reliable performance provided by UPQ.

X. EXTENSION TO MULTIUSER SCENARIO

Similar to [13], we extend our Algorithm 2 to optimize a multicast network, assuming perfect CSI with Rayleigh fading. Consider a max-min SNR problem with $U \geq 2$ receivers with a transmit power of $P = 30$ dBm, i.e.,

$$\max_{\theta \in \Phi_K} \min_u \left\{ \frac{P|\beta_{0,u}e^{j\alpha_{0,u}} + \sum_{n=1}^N \beta_{n,u}e^{j(\alpha_{n,u} + \theta_n)}|^2}{\sigma_u^2} \right\}, \quad (59)$$

where $\sigma_u^2 = -90$ dBm is the noise variance at each receive antenna, $h_{0,u} = \beta_{0,u}e^{j\alpha_{0,u}}$ is the direct channel, and $h_{n,u} = \beta_{n,u}e^{j\alpha_{n,u}}$ is the reflected channel through the n -th RIS element for the u -th receiver.

The way we extend our algorithm is as follows. While performing Algorithm 2 for user u , we decide the best possible solution in the for-loop of Algorithm 2 by maximizing the minimum channel gain among all users. Therefore, the for-loop of Algorithm 2 for a multicast network takes $\sum_{l=1}^N \mathcal{O}(U) = \mathcal{O}(NU)$ steps. Then, this process is repeated for each user, to select the best option among U possibilities, which results in $\mathcal{O}(NU^2)$ complexity in total. We remark that the complexity of the APX algorithm in this multicast scenario is $\mathcal{O}(NU)$.

The CDF plots for the minimum SNR performance of the multicast extension are given in Fig. 10 for $K = 2$ and $N \in \{16, 64, 256\}$. It can be seen that Algorithm 2 can provide superior performance compared to both APX and UPQ. The average gain against UPQ is 3.0 dB and against APX is 1.3 dB for $N = 16$. When $N = 64$, the average gain against UPQ is 4.0 dB and against APX is 2.3 dB. Note that these gains with Algorithm 2 get larger as N increases.

XI. PERFECT CSI ASSUMPTION

In this paper, we consider a wireless communications scenario assisted by an RIS to examine the performance of our algorithms. The proposed algorithms are highly efficient compared to the current state of the literature, as shown by numerical results. One of the main contributions is the improved performance that our algorithms can achieve. With this, the proposed algorithms in this paper take $\alpha_n, n = 1, \dots, N$, i.e., the channel phases, together with $\beta_n, n = 1, \dots, N$ as input to give the optimal discrete beamforming solution. We assume perfect CSI while deciding the RIS phase shifts, similar to [15]. One concern can be that, since the RIS elements are passive and there is a lack of signal processing capabilities, perfect CSI for the cascaded BS-RIS-UE link practically may not be available. However, the perfect CSI assumption can relate to some scenarios with certain assumptions. For example, with a two-stage approach in RIS-aided localization systems, the passive beamforming at the RIS is performed with the available channel information from the last localization step [23], [24]. Also, in experimental setups with RISs, the locations of the transmitter and receiver are known and passive beamforming is performed assuming that the location information of the UE is available [5], [14], [20], [21].

As described in Section III, the channels undergo Rician fading. While the performance and numerical results are provided with $\kappa = 0$ to compare with similar algorithms from the recent literature, for scenarios with the dominant

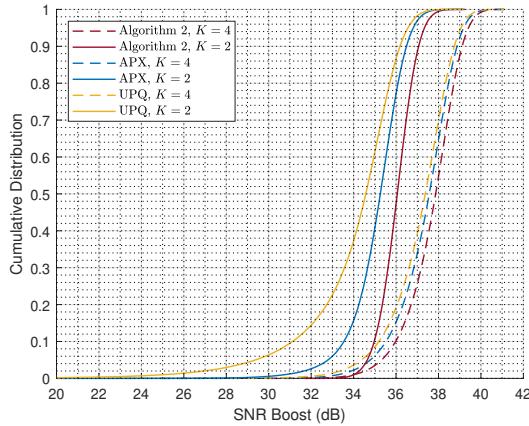


FIGURE 11. CDF plots for SNR Boost [15], $N = 200$, $\sigma_{\text{est}}^2 = -90\text{dBm}$, $\kappa = 0$, and $K \in \{2, 4\}$.

line-of-sight assumption in both BS-RIS and RIS-UE, i.e., large κ , the user's direction-of-arrival (DoA) information would be sufficient to find the channel phases, similar to the geometrical optimal model proposed in [31]. Whereas, if the channels are strictly NLOS, i.e., $\kappa = 0$ corresponding to Rayleigh fading, an example in the literature for estimating the channels in BS-RIS and RIS-UE links is to estimate the channels [34], which is the technique used in [13] for performance analysis. Another example is the joint channel estimation and passive beamforming framework in [12], where both are refined in each step with the CSI assumption for channel-gain-maximization. Therefore, in many scenarios, our perfect CSI assumption does not violate the applicability of our algorithms for future work in the literature.

A. Performance Results with the Estimated CSI

We investigate the performance of our proposed algorithm also considering the effect of imperfect CSI, similar to [13], by using the ON-OFF strategy in [34]. In these results, the APX algorithm is from [13], whereas UPQ and Algorithm 2 are as proposed in this paper. We adopt the same system parameters as in [13] to give a fair comparison, where the background noise power in the channel estimation phase σ_{est}^2 is set to -90 dBm as in the transmission phase. With this, the CDF of the performance results are given in Fig. 11 for $K = 2$ and $K = 4$ where the channels follow Rayleigh fading, i.e., $\kappa = 0$. The performance results show that Algorithm 2 outperforms both UPQ and APX. The average gain against UPQ is 1.5 dB and against APX is 0.8 dB for $K = 2$. When $K = 4$, the average gain against UPQ is 0.5 dB and against APX is 0.3 dB. However, it is important to note that in the worst-case scenario, i.e., the lower SNR Boost regime, the gain of Algorithm 2 over UPQ and APX is much higher, especially when $K = 2$.

XII. CONVERGENCE TO OPTIMALITY

We will now discuss the convergence of Algorithm 2 and Algorithm 3 to the global optimum. This will be given in the *Theorem* below. The proof of our *Theorem* is similar to but actually different than the proof of *Theorem 1* of [15].

Theorem: Algorithm 2 and Algorithm 3 yield the global optimum solution $(\theta_1^*, \theta_2^*, \dots, \theta_N^*)$ to (11) in average time $\mathcal{O}(N)$.

Proof: First, we will discuss the global optimality of our Algorithm 2 and Algorithm 3. This follows because each θ_n is optimally decided from (15) as in *Proposition 1* and all the possible arcs have been considered for μ that yields to all unique values of $|g_c|$, for which the optimality follows from (52) and the text that follows it in Section VII.

Next, we will discuss the complexity of Algorithm 2 and Algorithm 3. Similar to [15], assuming α_n are uniformly distributed in $[0, 2\pi)$, φ_n and φ_m are uniformly distributed in the interval $[0, \frac{2\pi}{K})$. As in [15], it can be argued that sorting in both algorithms will take $\mathcal{O}(N)$ time on average. For Algorithm 2, the for loop from Step 6 to Step 13 takes $\sum_{l=1}^N \mathcal{O}(1) = \mathcal{O}(N)$ steps. For Algorithm 3, the for loop from Step 7 to Step 14 takes $\sum_{l=1}^{N-N'} \mathcal{O}(|\mathcal{R}_{m_l}|) = \mathcal{O}(N)$ steps. As our algorithms are working with elementwise updates, the time complexity of the steps is also linear in N , as discussed in Section XIII. This means the overall complexity is $\mathcal{O}(N)$ on the average for both Algorithm 2 and Algorithm 3. Note that in contrast to [15], when $h_0 = 0$, we do not need to try out all possible arcs, and the average complexity will remain as $\mathcal{O}(N)$, instead of becoming $\mathcal{O}(KN)$. \blacksquare

Although we intend Algorithm 1 as a conceptual stepping stone towards Algorithms 2 and 3, a proof for its convergence to the optimal solution can be deduced in a fashion similar to that for Algorithm 2.

Finally, we remark that our algorithms do not require the uniformity assumption on α_n to be able to assure the global optimum. We assume α_n to be uniform for calculating the complexity, as in [15]. In fact, both Algorithm 2 and Algorithm 3 can be applied to a scenario and assure the global optimum where α_n are arbitrarily selected.

XIII. COMPUTATIONAL TIME RESULTS

We now provide computational complexity figures for our algorithms Algorithm 1 and Algorithm 2 in Fig. 12 and Fig. 13 for $K = 2$ and $K = 4$, respectively. We compare our algorithm with a number of algorithms from the literature, which are Approximation algorithm (APX) [13], and Divide-and-Sort (DaS) algorithm [14]. We also tabulate these results in terms of simulation time on the same computer (Dell XPS 15 9530 employing Intel Core i9-13900H CPU, 2.6 GHz, with 14 cores and 20 logical processors) with implementations carried out in Matlab. We note that all the results provided are obtained with our own implementation of the algorithms in the most efficient way we were able to achieve. In Table 2, we have the simulation time results in seconds

TABLE 2. Execution time [s] comparisons for 1000 channel realizations in non line-of-sight channel (NLOS), $K = 2$.

Method	$N = 10$	$N = 50$	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
DaS	0.1319	0.6268	1.2641	3.0912	18.2086	60.3753	277.2977
Algorithm 1	0.0804	0.1722	0.3874	0.9425	5.2790	16.2042	55.7506
Algorithm 2	0.0192	0.0416	0.0650	0.1107	0.2836	0.5130	1.1875
APX	0.0233	0.0377	0.0596	0.1036	0.2373	0.3732	0.9257
UPQ	0.0007	0.0011	0.0016	0.0024	0.0061	0.0094	0.0211

TABLE 3. Execution time [s] comparisons for 1000 channel realizations in NLOS, $K = 4$.

Method	$N = 10$	$N = 50$	$N = 100$	$N = 200$	$N = 500$	$N = 1000$	$N = 2000$
DaS	0.1552	0.9806	1.9666	6.9802	30.7800	105.6092	491.2088
Algorithm 1	0.0796	0.3384	0.7967	2.2021	10.0270	33.7036	115.7444
Algorithm 2	0.0211	0.0521	0.0934	0.1609	0.3180	0.5718	1.2623
APX	0.0249	0.0506	0.0890	0.1475	0.2648	0.4460	1.0281
UPQ	0.0007	0.0012	0.0016	0.0028	0.0063	0.0090	0.0204

plotted against the number of RIS elements N for $K = 2$. Then, in Table 3, we have the simulation results for $K = 4$. We note that among these algorithms, APX and UPQ do not have optimal performance, whereas DaS is claimed to be optimal. Recall that Algorithm 1 and Algorithm 2 achieve the optimal result.

In Fig. 12 and Fig. 13, the computational complexities of the algorithms are plotted against the number of RIS elements, N . Algorithm 2, APX, and UPQ provides execution times that increase linearly with N . Moreover, it is seen that our Algorithm 2 can achieve comparable computational complexity with APX, for $K = 2$ and $K = 4$, thanks to our practical initialization technique presented in the Appendix. Therefore, Algorithm 2, or Algorithm 3 when applicable, can be considered as benchmark algorithms instead of the APX algorithm, to get the *optimum* result with negligible additional computational costs. DaS is claimed to be optimal, but requires extremely high execution time, which makes it difficult to increase K . Our algorithms Algorithm 1 and Algorithm 2 are optimal in terms of performance. With the substantial reduction in computational complexity from Algorithm 1, Algorithm 2 can achieve substantially lower execution times, which is robust to increasing K , as shown in Tables 2 and 3, still ensuring the optimum result.

A. Remarks on the Complexity

Finally, we note that, to achieve a linear time-complexity as in [15], the main enabling factor is using the incremental update of g_l with $\mathcal{N}(\lambda_l)$ in the for-loop of Algorithm 1, 2, and 3, rather than calculating $\theta_n, n = 1, 2, \dots, N$ for each candidate μ . In order to do the elementwise updates, the sorting of channels in the initialization steps of Algorithm 1, Algorithm 2, and Algorithm 3 are required. It could be possible to avoid sorting of the channels by quitting elementwise updates.

Assume we adopt an approach, similar to [14], [21], where we first gather possible phase shift selections $\theta^l, l =$

$1, \dots, L$, and then perform each of them separately to find the maximum solution, for which the time complexity becomes $O(N^2)$ [15]. With this approach, [14], [21] require the sorting algorithm, in fact, [21] reports the complexity of the sorting requirement as it is the dominant term. However, we claim that, if we were to quit the elementwise approach, we could avoid the requirement of sorting in our algorithms. The main function of sorting is to try out all “arcs” for μ , so, we need to know the sorting of \underline{s}_{nk} . On the other hand, if we were to try $\mu = s_{n'k'}$ for all unique $s_{n'k'}$ and decide $\theta_n, n \in \{1, \dots, N\} \setminus n'$ with our *Lemma*, then setting $\theta_{n'} = (k' - 1)\omega$, we could generate the same set of solutions. Note that, in that case, the information of which s_{nk} should come before or after would be unnecessary. Therefore, we could avoid sorting. However, it still would not perform better than Algorithm 1, as the elementwise update plays a significant role in terms of lower time complexity. Moreover, using our initialization technique with the least number of steps can reduce the overall computational time substantially with the elementwise updates.

XIV. ALGORITHM COMPARISONS

To highlight the novelty and contribution of our paper, we compare our work with the existing literature for the problem defined in (11) in Table 4. There are two main comparisons: The first is the number of search steps to ensure convergence to the global optimum, and the second is the required time complexity to run the algorithms. Note that, although the number of search steps is linear in N , an incremental or elementwise structure is required for trying all the candidate phase shift configurations. Among those algorithms with ensured global optimality, our Algorithm 2 and Algorithm 3 are the only ones that can ensure global optimality in N or *fewer* search steps for any scenario. Regarding the algorithms that *converge to a local optimum*, APX requires three search steps to *approximate* the global optimum. Other algorithms that achieve a local minimum are CPP and our algorithm

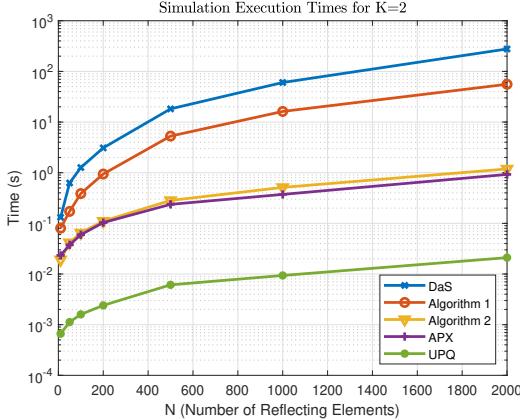


FIGURE 12. Plot of execution time [s] comparisons for 1000 channel realizations in NLOS, $K = 2$. Note vertical scale is logarithmic.

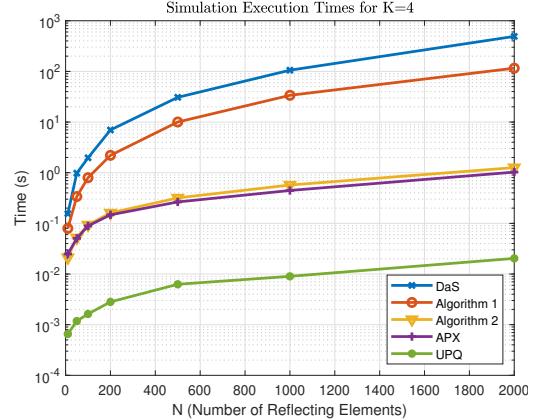


FIGURE 13. Plot of execution time [s] comparisons for 1000 channel realizations in NLOS, $K = 4$. Note vertical scale is logarithmic.

TABLE 4. Comparison of Algorithms 2-3 and UPQ with algorithms from the literature. See the text for definition of $\mathcal{N}(\lambda_l)$.

	Search Steps	Time Complexity	Optimality Guarantee
[13] APX	3	$\mathcal{O}(N)$	Local
[13] CPP	Projection of Each Phase Selection	—	Local
[13] Optimal	$2N + 2$, for $K = 2$	$\mathcal{O}(N)$ ($K = 2$ Only)	Global
[14] DaS	$N, K = 2$ $KN, K > 2$	$\mathcal{O}(N^2)$ (not elementwise)	Global
[15]	$2N, h_0 \neq 0$ $KN, h_0 = 0$	$\mathcal{O}(N)$ $\mathcal{O}(KN)$	Global
[20]	N , for $h_0 = 0$	$\mathcal{O}(N^2)$ (not elementwise)	Global
[21]	$2N + 1, K = 2$ $2N(K - 1), K > 2$	$\mathcal{O}(N^2)$ (not elementwise)	Global
[23] FPB	KN	$\mathcal{O}(N^2)$ (not elementwise)	Global
UPQ	Deterministic	—	Local
Algorithm 2	N , any h_0 $\mathcal{N}(\lambda_l) = 1$, all l	$\mathcal{O}(N)$	Global
Algorithm 3	$< N$, any h_0 $\mathcal{N}(\lambda_l) > 1$, some l	$\mathcal{O}(N)$	Global

UPQ. The detailed comparison of CPP and UPQ is discussed in Sec. IV.

XV. CONCLUSION

In this paper, we provided necessary and sufficient conditions for determination of optimum phase values in order to maximize the received power at a UE which receives its transmission by means of reflections from an RIS, when the phase values are from a discrete-valued set. Algorithms are provided to achieve this in a number of steps equal to N ,

the number of RIS elements, or fewer. In the literature, the number of steps to achieve this maximum is given as KN or $2N$ on the average, e.g., [14], [15]. In conclusion, for a discrete-phase RIS, the techniques in this paper achieve the optimum received power in the smallest number of steps published in the literature with an elementwise update rule. In addition, in each of those N steps, the techniques presented determine only one or a small number of phase shifts, which result in a substantial reduction of computation time, as compared to the algorithms in the literature, e.g., [14], [20], [21].

Finally, we want to make the following important point. In this paper, we addressed the ongoing problem in the literature of finding an *optimal solution* to the problem (11) within the fewest number of steps, or with minimum computational complexity. Our Algorithm 2 (or Algorithm 3 when applicable) achieves this goal. Yet, a secondary result of this paper is that the intuitive UPQ solution, which is based on the independent uniform quantization of θ_n on the unit circle with K points, results in suboptimal but very close to the optimal solution with very small complexity for all practical cases, provided full CSI.

XVI. ACKNOWLEDGMENT

The authors would like to thank Professor Thomas Ketseoglou for discussions on the asymptotic analysis of the UPQ algorithm.

APPENDIX: ELEMENT-BASED SIMPLE UPDATE RULE

We now further simplify Algorithm 1 by employing our periodicity proof, so that there is no need for calculating s_{nk} or λ_l . What we need to have is, given an initial $\underline{\mu}$ selection, say $e^{j\underline{\mu}} \in \text{arc}(e^{j\lambda_{i-1}} : e^{j\lambda_i})$, we want to know the N -step update rule $\mathcal{N}(\lambda_l)$, $l = i, i+1, \dots, i+N-1$ in the for loop of Algorithm 1.

Claim: Let \mathcal{U} be the set to define the N consecutive updates in the for loop of Algorithm 1. For an initial $\underline{\mu}_0$ selection, the update rule in the for-loop of Algorithm 1 will

$$\text{be } \mathcal{U} = \{n_1, n_2, \dots, n_N | 0 \leq \varphi_{n_1} < \varphi_{n_2} < \dots < \varphi_{n_N} < \frac{2\pi}{K}, \varphi_n = (\alpha_n - \underline{\mu}_0 + \frac{\pi}{K}) \bmod \frac{2\pi}{K}, n = 1, 2, \dots, N\}.$$

Proof: First, consider the case when $\underline{\mu} = 0$. We know that the initial arc is $\text{arc}(e^{j\lambda_L} : e^{j\lambda_1})$. Therefore, the update rule must be $\mathcal{U} = (\mathcal{N}(\lambda_l))_{l=1}^N$. We have already calculated this in (45) for any Case i given in (39). Note, from (39) to (45), $(\mathcal{N}(\lambda_l))_{l=1}^{NK}$ follows from the indexes of the sorted values of

$$\varphi_n = \left(\alpha_n + \frac{\pi}{K} \right) \bmod \frac{2\pi}{K}. \quad (60)$$

Now, consider the case when $\underline{\mu} = \underline{\mu}_0$ where $e^{j\underline{\mu}_0} \notin \text{arc}(e^{j\lambda_L} : e^{j\lambda_1})$. In this case, instead of moving μ to a new arc, we can introduce an offset of $-\underline{\mu}_0$ for all λ_l . Note that this corresponds to $\alpha_n \rightarrow \alpha_n - \underline{\mu}_0$, for all n . Therefore (60) will be updated as

$$\varphi_n = \left(\alpha_n - \underline{\mu}_0 + \frac{\pi}{K} \right) \bmod \frac{2\pi}{K}. \quad (61)$$

Thus, the proof is complete. \blacksquare

Now, when $h_0 \neq 0$, to initialize with $\underline{\mu} = \alpha_0 - \frac{\pi}{K}$, we can simply insert $\underline{\mu}_0 = \alpha_0 - \frac{\pi}{K}$ in (61) and get

$$\varphi_n = (\alpha_n - \alpha_0) \bmod \frac{2\pi}{K} \quad (62)$$

to be used in the initialization step. When the BS-UE link is completely blocked, or $h_0 = 0$, initializations can be updated as $\underline{\mu} = 0$ in Step 4 and $\varphi_n = (\alpha_n - \frac{\pi}{K}) \bmod \frac{2\pi}{K}$ for $n = 1, 2, \dots, \frac{M}{K}$ in Step 2.

It is important to note that, the simplification in (61) relieves Algorithm 1 from the burden to calculate NK instances of both s_{nk} and λ_l , and significantly reduces the computational complexity, as shown in Section XIII.

REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [2] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.
- [3] E. Ayanoglu, F. Capolino, and A. L. Swindlehurst, "Wave-controlled metasurface-based reconfigurable intelligent surfaces," *IEEE Wireless Communications*, May 2022.
- [4] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135–158, Feb. 2022.
- [5] X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, and E. Björnson, "RIS-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials," *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8627–8640, Sep. 2021.
- [6] X. Yu, D. Xu, and R. Schober, "Optimal beamforming for MISO communications via intelligent reflecting surfaces," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020, pp. 1–5.
- [7] —, "Enabling secure wireless communications via intelligent reflecting surfaces," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.
- [8] 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.
- [9] X. Qian, M. Di Renzo, J. Liu, A. Kammoun, and M.-S. Alouini, "Beamforming through reconfigurable intelligent surfaces in single-user MIMO systems: SNR distribution and scaling laws in the presence of channel fading and phase noise," *IEEE Wireless Communications Letters*, vol. 10, no. 1, pp. 77–81, 2021.
- [10] 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, Dec. 2019.
- [11] Z. Yigit, E. Basar, and I. Altunbas, "Low complexity adaptation for reconfigurable intelligent surface-based MIMO systems," *IEEE Communications Letters*, vol. 24, no. 12, pp. 2946–2950, 2020.
- [12] 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.
- [13] 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, Aug. 2022.
- [14] R. Xiong, X. Dong, T. Mi, K. Wan, and R. C. Qiu, "Optimal discrete beamforming of RIS-aided wireless communications: An inner product maximization approach," arXiv:2211.04167v6, Nov. 2023.
- [15] S. Ren, K. Shen, X. Lin, 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, Mar. 2023.
- [16] M. Jung, W. Saad, M. Debbah, and C. S. Hong, "On the optimality of reconfigurable intelligent surfaces (RISs): Passive beamforming, modulation, and resource allocation," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4347–4363, 2021.
- [17] J. Luo, K. Pattipati, and P. Willett, "A sub-optimal soft decision PDA method for binary quadratic programming," in *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat.No.01CH37236)*, vol. 5, 2001, pp. 3140–3145 vol.5.
- [18] A. Yellepeddi, K. J. Kim, C. Duan, and P. Orlik, "On probabilistic data association for achieving near-exponential diversity over fading channels," in *2013 IEEE International Conference on Communications (ICC)*, 2013, pp. 5409–5414.
- [19] A. Pradhan and H. S. Dhillon, "A probabilistic reformulation technique for discrete RIS optimization in wireless systems," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.
- [20] J. Sanchez, E. Bengtsson, F. Rusek, J. Flordelis, K. Zhao, and F. Tufvesson, "Optimal, low-complexity beamforming for discrete phase reconfigurable intelligent surfaces," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 01–06.
- [21] I. Vardakas, G. Kotridis, S. Peppas, K. Skyvalakis, G. Vougioukas, and A. Bletsas, "Intelligently wireless batteryless RF-powered reconfigurable surface: Theory, implementation & limitations," *IEEE Transactions on Wireless Communications*, vol. 22, no. 6, pp. 3942–3954, 2023.
- [22] S. Zhang and Y. Huang, "Complex quadratic optimization and semidefinite programming," *SIAM Journal on Optimization*, vol. 16, no. 3, pp. 871–890, 2006. [Online]. Available: <https://doi.org/10.1137/04061341X>
- [23] B. Luo, Y. Deng, M. Dong, Z. Huang, X. Chen, W. Han, and B. Bai, "Localization and discrete beamforming with a large reconfigurable intelligent surface," 2023.
- [24] J. He, H. Wymeersch, T. Sanguanpuak, O. Silven, and M. Juntti, "Adaptive beamforming design for mmWave RIS-aided joint localization and communication," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2020, pp. 1–6.
- [25] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "Deep reinforcement learning based beamforming codebook design for RIS-aided mmWave systems," in *2023 IEEE 20th Consumer Communications Networking Conference (CCNC)*, 2023, pp. 1020–1026.
- [26] W. R. Ghanem, V. Jamali, M. Schellmann, H. Cao, J. Eichinger, and R. Schober, "Optimization-based phase-shift codebook design for large IRSs," *IEEE Communications Letters*, vol. 27, no. 2, pp. 635–639, 2023.
- [27] G. C. Trichopoulos, P. Theofanopoulos, B. Kashyap, A. Shekhawat, A. Modi, T. Osman, S. Kumar, A. Sengar, A. Chang, and A. Alkhateeb, "Design and evaluation of reconfigurable intelligent surfaces in real-world environment," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 462–474, 2022.
- [28] 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.

[29] H. Huang, Y. Zhang, H. Zhang, Z. Zhao, C. Zhang, and Z. Han, “Multi-IRS-aided millimeter-wave multi-user MISO systems for power minimization using generalized Benders decomposition,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 7873–7886, 2023.

[30] J. Wang, S. Gong, Q. Wu, and S. Ma, “RIS-aided MIMO systems with hardware impairments: Robust beamforming design and analysis,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 10, pp. 6914–6929, 2023.

[31] T. Mi, J. Zhang, R. Xiong, Z. Wang, P. Zhang, and R. C. Qiu, “Towards analytical electromagnetic models for reconfigurable intelligent surfaces,” *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.

[32] T. Jiang, H. V. Cheng, and W. Yu, “Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 1931–1945, 2021.

[33] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, “Intelligent reflecting surface: Practical phase shift model and beamforming optimization,” *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5849–5863, 2020.

[34] T. L. Jensen and E. De Carvalho, “An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 5000–5004.



DOGAN KUTAY PEKCAN (Student Member, IEEE) received his B.S. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University, Ankara, Türkiye, in 2018 and 2021, respectively. During his M.S. studies, he worked on research and development of detection and estimation algorithms, for various 5G physical layer channels, specializing in channel estimation and phase noise mitigation in mm-Wave. Currently, he is pursuing his Ph.D. degree in Electrical Engineering and Computer Science at University of California, Irvine, CA, USA. His research interests are wireless communications, reconfigurable intelligent surfaces, and algorithms. Outside the academia and work, his set of hobbies includes playing the piano, swimming, and solving puzzles.



ENDER AYANOGLU (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1986. He was with the Bell Laboratories Communications Systems Research Laboratory, Holmdel, NJ, USA. From 1999 to 2002, he was a Systems Architect with Cisco Systems Inc., San Jose, CA. Since 2002, he has been a Professor with the Department of Electrical Engineering and Computer Science, University of California, Irvine, CA, where he was the Director of the Center for Pervasive Communications and Computing and the Conexant-Broadcom Endowed Chair, from 2002 to 2010. He was a recipient of the IEEE Communications Society Stephen O. Rice Prize Paper Award in 1995, the IEEE Communications Society Best Tutorial Paper Award in 1997, and the IEEE Communications Society Communication Theory Technical Committee Outstanding Service Award in 2014. From 2000 to 2001, he was the Founding Chair of the IEEE-ISTO Broadband Wireless Internet Forum, an industry standards organization. He served on the Executive Committee for the IEEE Communications Society Communication Theory Committee, from 1990 to 2002, and its Chair, from 1999 to 2002. From 1993 to 2014, he was an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS. He was the Editor-in-Chief of IEEE TRANSACTIONS ON COMMUNICATIONS, from 2004 to 2008, and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS-Series on Green Communications and Networking, from 2014 to 2016. He was the Founding Editor-in-Chief of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, from 2016 to 2020. He served as an IEEE Communications Society Distinguished

Lecturer, from 2022 to 2023 and will serve in the same capacity in 2024-2025. In 2023, he received the IEEE Communications Society Joseph L. LoCicero Award for outstanding contributions to IEEE Communications Society journals as Editor, Editor-in-Chief (EiC), and Founding EiC.