

Low-Complexity IRS Beamforming Based on Sphere Decoding and Tabu Search

Seraphin Kimaryo and Kyungchun Lee

Abstract—The real-time reconfiguration of intelligent reflecting surface (IRS) is one of the crucial, yet challenging tasks required to exploit the benefits of IRS. In this paper, we study an IRS-aided multiple-input single-output (MISO) wireless network, where a coarsely quantized (few resolution bits) IRS is reconfigured to help the point-to-point communication between a multi-antenna access point (AP) and a single-antenna user. Specifically, the transmit precoders at the AP and the reflecting discrete phase shifters at the IRS are jointly optimized to maximize the signal-to-noise ratio (SNR) at the user. The optimization is studied in two practical wireless propagation environments. First, we consider an environment where the direct AP–user link is in deep fade, which enables the relaxation of the optimization problem such that it can be optimally solved by the proposed sphere-decoder (SD) algorithm. Second, a normal propagation environment for all links is considered, wherein we propose the tabu search (TS) algorithm to solve the joint active and passive beamforming problem with low complexity. The proposed TS algorithm is shown to work well, even in deep-fade AP–user link scenarios. The presented numerical results confirm the validity of our analysis and demonstrate the effectiveness of the proposed schemes over the benchmark schemes on different system setups and propagation environments.

Index Terms—Beamforming, complex sphere decoder, intelligent reflecting surfaces, MIMO, phase shift, sixth generation (6G), tabu search.

I. INTRODUCTION

It is expected that by 2026, more than half of the global data traffic will be carried by 5G networks [1]. This inevitable growth of mobile data traffic on wireless networks is not only accelerated by improvements in device capabilities, but also the emergence of and increase in data-intensive content, such as virtual reality (VR), augmented reality (AR), and video gaming. To accommodate this huge increase in data traffic, some revolutionary techniques, especially at the physical layer, must be deployed. The multiple-input multiple-output (MIMO) system is one of the key technologies, which

Manuscript received July 8, 2022 revised December 29, 2022; approved for publication by Xiang-Gen Xia, Division 2 Editor, January 29, 2023.

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2019R1A6A1A03032119 and in part by the NRF Grant funded by the Korean Government (MSIT) under Grant NRF-2022R1A2C1006566. (Corresponding author: Kyungchun Lee.)

S. Kimaryo is with the Department of Electrical and Information Engineering, Seoul National University of Science and Technology, Seoul 01811, Republic of Korea, email: seraphin@seoultech.ac.kr.

K. Lee is with the Department of Electrical and Information Engineering and the Research Center for Electrical and Information Technology, Seoul National University of Science and Technology, Seoul 01811, Republic of Korea, email: klee@seoultech.ac.kr.

K. Lee is the corresponding author.

Digital Object Identifier: 10.23919/JCN.2023.000007

has been extensively studied to cater for this increase in data traffic. The MIMO system has superbly managed to provide both higher spectral efficiencies and higher spatial diversity gains by exploiting multiple antennas at the transmitter and receiver. Nevertheless, it suffers from significantly high power consumption, complexity, and hardware cost, which reduce its suitability for the envisioned simple, low-complexity, and cost-effective wireless communication [2].

In contrast, intelligent reflecting surface (IRS), which is planar array consisting of a massive number of nearly passive reflecting elements, has recently gained much attention in both industry and academia for its ability to improve and control wireless environments. Each IRS element can be independently configured to modify the characteristics (amplitude and phase) of the impinging electromagnetic waves. It is therefore possible to leverage the independent configuration of each element to collaboratively beamform the impinging signals. Several use cases of this newly emerging technology have been reported. For example, IRS can be used to boost signal strength at the desired receiver by either reinforcing signals from other directions or suppressing the interfering signals from nearby users [3]. In addition, IRS can be used to cancel genuine signals at unintended users by sending a matching-magnitude opposite-phase signal toward an eavesdropper, thus enhancing communication secrecy [4]–[6]. By varying the reflection pattern of its elements, IRS can also be used to transfer information by a technique popularly called reflecting modulation [7] or IRS-based information transfer [8].

Unlike its closest competitor, half-duplex (HD) decode-and-forward (DF) relaying [9], which requires a separate power source to process and relay information from an access point (AP) to the user, IRS is full-duplex [9] and nearly passive. Following its passive nature, which is due to the absence of expensive and power-greedy radio frequency (RF) chains, IRS reflections are free of thermal noise. The absence of RF chains makes IRS not only passive, but also lightweight and compact in size [10]. This simple structure is another highly attractive feature of IRS, as it can be deployed flexibly in various places, including shopping malls, stadiums, and airports.

Therefore, IRS is considered one of the top candidate link layer technologies to enable forthcoming wireless communication standards, such as sixth-generation (6G) networks. However, there are several challenges remain to be solved to achieve such applications. For example, IRS is highly disadvantaged by its inability to estimate its associated channels. It is generally assumed that the entire channel-estimation task is performed by the AP. This can lead to prolonged channel-

Creative Commons Attribution-NonCommercial (CC BY-NC).

This is an Open Access article distributed under the terms of Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided that the original work is properly cited.

estimation time, which poses a threat to the application of IRS in fast-fading channels. Another predominant challenge, of which solution will unlock several applications, is the beamforming based on discrete phase shifts. Most of the envisioned IRS applications to date, regardless of whether they boost signal strength or mitigate interference, unanimously require proper IRS phase-shift configurations to achieve their desired goal. The study of discrete IRS phase shifts is therefore of vital importance. In this work, motivated by the above, we particularly focus on the optimization of discrete IRS phase shifts in both favorable and unfavorable wireless propagation environments.

A. Related Works

The paradigm for IRS phase-shift optimization can be described in two general ways. First, the optimization can be performed with continuous phases, where each element is allowed to take any phase in the range between 0 and 2π . Some early works on IRS adopted this strategy [9], [11]–[13]. For example, [9] and [11] both applied semidefinite relaxation (SDR) and alternating optimization (AO) to solve the joint active and passive beamforming problem to minimize transmit power at the AP and maximize signal strength at the receiver, respectively. In [12] and [13], the sum rate of the IRS-aided multiple-input single-output (MISO) network was maximized. Whereas the former used fractional programming to tackle the resulted nonconvex problem, the latter deployed alternating maximization with majorization-minimization. Owing to the hardware constraint at the IRS, the continuous-phase solutions provided by this strategy are not highly practicable, even though they can provide the foundation to tackle the more realistic problem, i.e., discrete IRS phase-shift optimization.

In [14], an IRS with a finite number of reflecting units, each with discrete phase shifts, was investigated to improve the energy efficiency (EE) of an outdoor cellular network by applying sequential fractional programming and conjugate gradient search. The work presented in [15] improved upon that in [9] by quantizing the continuous phases to their nearest discrete phases. This technique, however, leads to high performance loss, especially with a 1-bit IRS. A branch-and-bound (BnB) scheme was proposed in [15] and [16] to provide high-quality solutions; however, its high complexity, which scales with the size of the IRS, can significantly limit its application in large-scale IRS systems. The authors of [17] proposed cross-entropy (CE) optimization to overcome the performance loss due to phase discretization and gradient projection (GP) as an alternative to high-complexity schemes like BnB and manifold optimization (MO) [16], [18].

B. Contributions

IRS is envisioned to greatly assist the communication between the transmitter and receiver, especially when the link between them is in deep fade or completely blocked [19]. The most unique feature of IRS is its ability to achieve great performance while using a very small amount of power. Although it also provides performance gains when the direct link between the transmitter and receiver is strong, its performance becomes

comparable to or even lower than that of other technologies, such as DF, with the same settings [19]. Considering its inability to estimate the channel, plus deployment in a strong direct AP–user link, IRS is not very useful compared to DF. Most of the proposed algorithms yield good performance when there is a strong direct link between the AP and user. They can also work well in unfavorable wireless propagation conditions, albeit by increasing computational complexity. As such, the design of spectral- and complexity-efficient algorithms is still imperative to fully exploit the benefits of IRS. We address this by analyzing the IRS-assisted point-to-point MISO system with the following main contributions:

- We propose applying the sphere decoder (SD) and tabu search (TS) to optimize the discrete phase shifts of IRS. These two algorithms provide high performance with low complexity in extremely unfavorable propagation conditions. Besides, the proposed algorithms also work well even in good wireless conditions, hence making them a better choice for joint active (at the AP) and passive (at the IRS) beamforming.
- Because a large portion of the complexity of AO and CE [17] are accumulated from the computation of their objective values, we modify the QR-TS complexity-reduction technique proposed in [20] to work with non-square matrices. The modification allows the deployment of the technique in our proposed TS algorithm to efficiently compute the objective value of each neighbor, which in turn significantly reduces the complexity.
- Average computational complexities for the proposed algorithms and benchmark schemes are analyzed and their analytical expressions are derived. Furthermore, the performance is extensively evaluated to validate our analysis and demonstrate the effectiveness of the proposed algorithms in providing an improved trade-off between performance and complexity especially when the phase shifts of the IRS are coarsely quantized.

Notation: Throughout this paper, \mathbb{R} and \mathbb{C} denote real and complex domains, respectively. Scalars are denoted by lowercase italic letters. The bold-face lower-case (\mathbf{a}) and upper-case (\mathbf{A}) letters denote a vector and a matrix, respectively. For any general matrix \mathbf{M} , $M_{i,j}$ denotes the element at the i th row and j th column, whereas M_i and $M_{\bar{i}}$ denote the i th column and row of \mathbf{M} , respectively. M^T and M^H represent the transpose and conjugate transpose of \mathbf{M} , respectively. \mathbf{I}_N and $\mathbf{0}_{n,m}$, respectively, indicate an identity matrix with a dimension of $N \times N$ and an $n \times m$ matrix whose entries are all zeros. For a complex-valued vector \mathbf{x} , x_i is the i th element of \mathbf{x} ; $|\mathbf{x}|$ and $\|\mathbf{x}\|$ denote the absolute value and Euclidean norm of \mathbf{x} , respectively. $\mathbb{E}(\cdot)$ denotes the expectation operator. $\text{diag}(\mathbf{x})$ means a diagonal matrix with the elements of its main diagonal being the entries of \mathbf{x} . $\arg(\cdot)$ represents the phase extraction operator, which returns the phase of each element of its argument. $\text{card}(\cdot)$ is the cardinality operator, which returns the number of member elements of a set. $\text{mod}(\cdot, \cdot)$ is the modulus operator, which returns the remainder of the division of the operator's first argument by its second argument. A circularly symmetric complex Gaussian random variable with

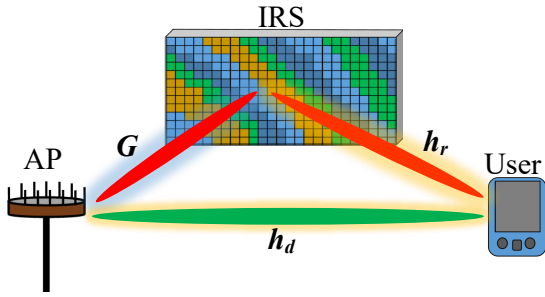


Fig. 1. IRS-assisted downlink MISO system.

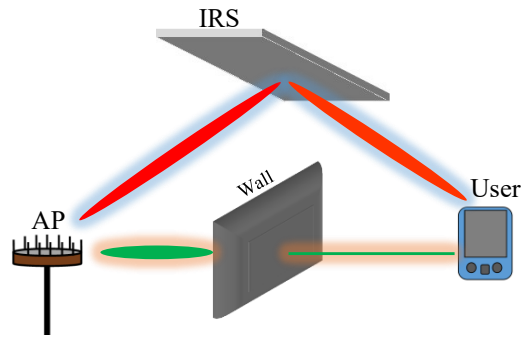


Fig. 2. IRS-assisted coverage extension.

mean μ and variance σ^2 is denoted by $\mathcal{CN}(\mu, \sigma^2)$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

We consider a downlink MISO communication system, which is illustrated in Fig. 1. The AP, equipped with M transmit antennas, communicates with a user, equipped with a single receive antenna. This communication is assisted by an IRS comprising N nearly passive reflecting elements, which reflect the incident electromagnetic waves from the AP in the direction of the user by changing the phase shift of each element such that the signals reinforce each other at the destination. Because IRS is a passive device, which is not able to estimate the channel by itself, in practice, it is attached with a smart controller, which communicates with the AP through a dedicated link to exchange the information [3]. For simplicity of analysis, we assume that the channel state information (CSI) of all the involved channels is available and known to both the AP and user. Furthermore, a quasistatic flat fading model is assumed for all the channels.

When the AP precodes the unit-power symbol s , i.e., $\mathbb{E}\{|s|^2\} = 1$, by a linear beamforming vector $\mathbf{f} \in \mathbb{C}^{M \times 1}$, the combined received signal y at the user via both the AP–user link and AP–IRS–user link is given by

$$y = \sqrt{P} \left(\mathbf{h}_d^T + \mathbf{h}_r^T \Phi \mathbf{G} \right) \mathbf{f} s + z, \quad (1)$$

where P is the maximum transmit power at the AP; $\mathbf{h}_d \in \mathbb{C}^{M \times 1}$, $\mathbf{h}_r \in \mathbb{C}^{N \times 1}$, and $\mathbf{G} \in \mathbb{C}^{N \times M}$ denote the baseband equivalent channels from the AP to the user, from the IRS to the user, and from the AP to the IRS, respectively; z is a zero-mean complex Gaussian noise signal with variance σ^2 ; and $\Phi \in \mathbb{C}^{N \times N} = \text{diag}(\eta_1 e^{j\theta_1}, \dots, \eta_N e^{j\theta_N})$ is the diagonal matrix containing the phase shifts $\theta_n \in (0, 2\pi]$, $\forall n = 1, \dots, N$ and reflection coefficients $\eta_n \in [0, 1]$, $\forall n = 1, \dots, N$ of all IRS elements. The reflection coefficients η_n represent the magnitude with which the impinging signal at each IRS element is reflected. In the IRS design perspective, it is desired to reflect signals from each element with the maximum amplitude possible [9], [21]–[23]. This guarantees stronger signals at the receiver, provided that phase shifts and interference are properly managed. Following this, we let the reflection coefficient of each element take the maximum possible value, i.e., $\eta_n = 1$, $\forall n = 1, \dots, N$.

Furthermore, note that it is theoretically feasible to tune the angles of the elements to any value in the range $(0, 2\pi]$. However, this approach is very costly in practice, especially with a large-scale IRS system. The practical solution is to design each element with finite quantization bits. Let $\mathcal{F}_b \triangleq \{0, 2\pi/2^b, \dots, 2\pi(2^b - 1)/2^b\}$ denote a finite set of phase shifts obtained by a uniform quantization of an interval $(0, 2\pi]$, where $b \geq 1$ is the number of quantization bits. Consequently, the notation of the diagonal matrix can be simplified to $\Phi = \text{diag}(e^{j\theta_1}, \dots, e^{j\theta_N})$, $\theta_n \in \mathcal{F}_b$. The signal-to-noise ratio (SNR) at the user is expressed as

$$\text{SNR} = \frac{P |(\mathbf{h}_d^T + \mathbf{h}_r^T \Phi \mathbf{G}) \mathbf{f}|^2}{\sigma^2}. \quad (2)$$

B. Problem Formulation

In this paper, we aim to maximize the signal strength at the user by jointly optimizing the active transmit precoder \mathbf{f} at the AP and phase shifts at the IRS under the constraint of maximum transmit power and discrete phase shift. This optimization is done under two practical IRS environments.

1) *Severely attenuated direct AP–user link*: It is a common scenario in wireless communication that users get blocked from the sight of an AP, or the signal reaching them is very weak due to severe attenuation along the path. To help improve this situation, IRS reflects the incident signal toward the user. In its most common usage, IRS is normally deployed within sight of an AP and in the vicinity of the user, which ensures reception of a stronger signal at the user [3], [9]. To demonstrate this environment, consider the deployment arrangement shown in Fig. 2, where the direct link suffers severe attenuation when penetrating the blocking wall, thus reaching a user with very weak strength. In this environment, IRS can be employed to provide an alternative path to reach the user. The maximization of the signal strength at the user can be formulated as

$$(P1): \quad \max_{\Phi, \mathbf{f}} \quad |(\mathbf{h}_d^T + \mathbf{h}_r^T \Phi \mathbf{G}) \mathbf{f}|^2, \quad (3)$$

$$\text{s.t.} \quad \|\mathbf{f}\|^2 \leq 1, \quad (4)$$

$$\theta_n \in \mathcal{F}_b. \quad (5)$$

The deep coupling of \mathbf{f} and Φ makes (3) intractable by normal heuristic methods [9]. However, by using the fact that for any

IRS phase shift matrix Φ , the optimal active precoding vector is given by [24]

$$\mathbf{f} = \frac{\left(\mathbf{h}_d^T + \mathbf{h}_r^T \Phi \mathbf{G}\right)^H}{\|\mathbf{h}_d^T + \mathbf{h}_r^T \Phi \mathbf{G}\|}, \quad (6)$$

and problem (3) can be decoupled to

$$\max_{\Phi} \|\mathbf{h}_d^T + \mathbf{h}_r^T \Phi \mathbf{G}\|^2, \quad (7)$$

$$\text{s.t. } \theta_n \in \mathcal{F}_b. \quad (8)$$

As shown in [9], by applying the change of variables $\mathbf{J}^T = \text{diag}(\mathbf{h}_r) \mathbf{G}$ and letting $\mathbf{v} = [e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_N}]^T$, the second inner term of (7) can be written as $\left(\mathbf{h}_r^T \Phi \mathbf{G}\right)^T = \mathbf{J}\mathbf{v}$. Accordingly, (7) can be rewritten as

$$\text{(P2): } \max_{\mathbf{v}} p(\mathbf{v}), \quad (9)$$

$$\text{s.t. } |\mathbf{v}_n| = 1, \forall n, \quad (10)$$

$$\arg(\mathbf{v}_n) \in \mathcal{F}_b, \forall n, \quad (11)$$

where $p(\mathbf{v}) = \|\mathbf{h}_d + \mathbf{J}\mathbf{v}\|^2$ is the objective function of \mathbf{v} . Due to severe signal attenuation over the direct link \mathbf{h}_d , it is obvious that $\|\mathbf{h}_d\| \ll \|\mathbf{J}\mathbf{v}\|$, which subsequently allows the transformation of (9) to

$$\text{(P3): } \max_{\mathbf{v}} \|\mathbf{J}\mathbf{v}\|^2, \quad (12)$$

$$\text{s.t. } |\mathbf{v}_n| = 1, \forall n, \quad (13)$$

$$\arg(\mathbf{v}_n) \in \mathcal{F}_b, \forall n. \quad (14)$$

The optimal method to solve (P3) is an exhaustive search; however, this is highly impractical owing to its exponentially-growing complexity in the dimension of the number of IRS elements. Fortunately, as described in Section III, near-optimal results can still be obtained with greatly reduced complexity by using the SD.

2) *Strong direct AP–user link*: By acknowledging that IRS is also expected to perform well in environments with a strong direct link, we extend the previous analysis to a more general setting. Note that most of the existing schemes have been developed to work with this kind of setup. The additional AP–IRS–user link not only boosts signal strength at the user, but also increases the spatial diversity of the overall composite channel matrix, which is crucial in enabling multistream transmission in MIMO [9]. To account for the effect of a strong direct link, we fall back to problem (P2). Solving this problem optimally would again require the use of exhaustive search. Due to the aforementioned limitations of this approach, several works have tried to avoid it by proposing alternative algorithms such as AO, GP, CE, etc. Although AO, GP, and CE have been excellent alternatives, it came to our attention that we can obtain higher performance by using TS especially in more practical settings like coarse (1-bit phase shifters) IRS. We therefore, in Section IV, propose and discuss the TS algorithm to solve (P2) in lesser complexity and with improved performance.

III. SPHERE DECODING

The SD is one of the powerful algorithms in solving different kinds of combinatorial problems. It provides a near-optimal solution with a great reduction of complexity. It has been extensively explored for the detection of wireless communication signals in MIMO systems [25], [26], and the problems of the like. In this section we demonstrate how to use SD to solve (P3).

A. SD Problem Formulation

Despite of being so powerful, SD is limited to solving minimization problems only. Because (12) is a maximization problem, the first step is to convert it to a minimization problem. From (12), it follows that if \mathbf{v} is constrained to have a unit Euclidean norm, i.e., $\|\mathbf{v}\| = 1$, the optimal solution denoted as $\tilde{\mathbf{v}}$ would simply be the singular vector of \mathbf{J} associated with the largest singular value δ_{\max} , i.e., $\tilde{\mathbf{v}} = \phi(\delta_{\max})$, where $\phi(\cdot)$ denotes a singular vector extraction operator for a given singular value. Note that $\tilde{\mathbf{v}}$ does not qualify to be the solution of problem (12) for two main reasons. First, it does not satisfy constraint (13), hence it cannot guarantee perfect reflection from each element. Second, it does not agree with constraint (14), which makes it a continuous solution rather than a practical discrete solution. To address these problems, we start by scaling $\tilde{\mathbf{v}}$ so that it has the same Euclidean norm as \mathbf{v} , i.e.,

$$\hat{\mathbf{v}} = \|\mathbf{v}\| \tilde{\mathbf{v}} = \sqrt{N} \tilde{\mathbf{v}}. \quad (15)$$

We note that $\hat{\mathbf{v}}$ is the optimal solution to (12) under the condition $\|\mathbf{v}\|^2 = N$. Although $\hat{\mathbf{v}}$ does not agree with both constraints, we are closer to achieving that goal. Essentially, this scaling allows the transformation of the maximization problem (12) to its minimization problem as follows:

$$\max_{\mathbf{v}} \|\mathbf{J}\mathbf{v}\|^2 = \min_{\mathbf{v}} (\|\mathbf{J}\hat{\mathbf{v}}\|^2 - \|\mathbf{J}\mathbf{v}\|^2), \quad (16)$$

$$= \min_{\mathbf{v}} \left(N \|\mathbf{J}\tilde{\mathbf{v}}\|^2 - \mathbf{v}^H \mathbf{J}^H \mathbf{J} \mathbf{v} \right), \quad (17)$$

$$= \min_{\mathbf{v}} \left(N \delta_{\max}^2 - \mathbf{v}^H \mathbf{J}^H \mathbf{J} \mathbf{v} \right), \quad (18)$$

$$\stackrel{(a)}{=} \min_{\mathbf{v}} \mathbf{v}^H \left(\delta_{\max}^2 \mathbf{I}_N - \mathbf{J}^H \mathbf{J} \right) \mathbf{v}, \quad (19)$$

where (a) is due to the fact that $\mathbf{v}^H \mathbf{v} = \|\mathbf{v}\|^2 = N$. Note that by adding a small number, e.g., $\tau = 10^{-25}$ to δ_{\max}^2 such that (19) remains approximately the same, the matrix $\mathbf{B} = (\delta_{\max}^2 + \tau) \mathbf{I}_N - \mathbf{J}^H \mathbf{J}$ becomes always positive definite, i.e., $\mathbf{B} \succ \mathbf{0}$. The positive definiteness of \mathbf{B} allows its Cholesky decomposition, which is necessary for further simplification of (19). Let $\mathbf{S} \in \mathbb{C}^{N \times N}$ be an upper triangular matrix obtained after applying Cholesky decomposition on \mathbf{B} , i.e.,

$$\mathbf{B} = \mathbf{S}^H \mathbf{S}. \quad (20)$$

By using (20), the problem (19) simplifies further as follows:

$$\min_{\mathbf{v}} \mathbf{v}^H \mathbf{S}^H \mathbf{S} \mathbf{v} = \min_{\mathbf{v}} \|\mathbf{S}\mathbf{v}\|^2. \quad (21)$$

Together with constraints (13) and (14), we can finally apply SD to solve problem (21).

Algorithm 1: SD algorithm**Input :** $\mathbf{J}, \mathcal{F}_b, N$

- 1 Compute the largest squared singular value δ_{\max}^2 of \mathbf{J}
- 2 Set $\mathbf{B} = (\delta_{\max}^2 + \tau) \mathbf{I}_N - \mathbf{J}^H \mathbf{J}$
- 3 Find \mathbf{S} according to (20)
- 4 Solve for \mathbf{v} in (21) by using SD

Output: \mathbf{v} *B. Application of SD*

In many applications, SD is used to solve real-valued problems. Hence, it starts by converting the complex-valued problems to their real-valued equivalent problems [25], [27]. In contrast, this work solves the original complex-valued problem as the conversion to the real-valued problem does not work in this context. What follows until the end of this section is the step-by-step description of the proposed SD algorithm summarized in Algorithm 1.

At the beginning, SD computes the square of the largest singular value of the matrix \mathbf{J} . This can be done by straightforwardly applying the computational-efficient power iteration method [28] on the Hermitian matrix $\mathbf{J}^H \mathbf{J}$, of which largest eigenvalue is the largest squared singular value of \mathbf{J} , i.e., $\delta_{\max}^2(\mathbf{J}) = \lambda_{\max}(\mathbf{J}^H \mathbf{J})$. In step 2, the algorithm stores the result of the computation $(\delta_{\max}^2 + \tau) \mathbf{I}_N - \mathbf{J}^H \mathbf{J}$ in the matrix \mathbf{B} . To make the problem SD-ready, the algorithm finalizes the preparatory stage by applying Cholesky decomposition on the positive definite matrix \mathbf{B} as in (20). The basic premise of solving for \mathbf{v} in step 4 lies in the observation that the matrix-vector product $\mathbf{S}\mathbf{v}$ forms a set of a complex lattice, where it is required to find the suitable combination of the elements of \mathbf{v} from the available set \mathcal{F}_b , which provides the smallest distance from the origin. To tackle this complex-valued combinatorial problem by using SD, we follow the description in [29].

Throughout this paper, we aim to optimize the performance–complexity tradeoff of our proposed algorithms. To achieve this goal in SD, we first apply Schnorr–Euchner (SE) enumeration [30]–[32] in each layer. Further, for a large number of close-to-optimal lattice points especially when b and N are large, the SD algorithm can demand high complexity to determine the final solution thus negatively impacting the performance–complexity tradeoff. To overcome this problem, we adopt an early-termination technique, where the algorithm is halted when it fails to reach the leaf node for a certain number of iterations $\varpi = \zeta N$, where ζ is the constant integer which is chosen to be greater than or equal to 2.

IV. TABU SEARCH

In this section, we demonstrate how TS can efficiently solve (P2) in any propagation environment (favorable or unfavorable) with low complexity. TS is an efficient searching algorithm for solving discrete optimization problems. TS iteratively searches through the domain of candidate solutions and their neighborhood, and avoids stagnation around local extrema by restricting itself from visiting recently visited

nodes. This cycling prevention is possible by the use of a short-term memory structure called tabu list, which stores the recent neighborhood moves. In doing so, it allows itself to have a wider search space, which eventually increases its chances of finding the global optimal point. For a full description of TS algorithm and its variants, refer to [33]–[35].

A. Main Concept

TS needs a starting point to begin its search. This point is generally constrained to be among the problem’s possible candidate solutions; otherwise, neighborhood definition fails. There are several techniques to determine the starting point, including using the solution of a less complex algorithm or a random choice from a set of candidate solutions. The latter technique is simpler; hence, we adopt it in this work and define the initial candidate solution as

$$\mathbf{c} \in \mathbb{C}^{N \times 1} = [e^{j\theta_1}, \dots, e^{j\theta_N}]^T, \theta_n \in \mathcal{F}_b, \forall n. \quad (22)$$

Denote a tabu list of length $L > 1$ by $\mathcal{T} \in \mathbb{C}^{N \times L}$ and set $\mathcal{T}_1 = \mathbf{c}$, where \mathcal{T}_i denotes the i th column of \mathcal{T} . Similarly, let \mathbf{x}^* denote the overall best solution found by the algorithm, and initialize it to \mathbf{c} , i.e., $\mathbf{x}^* = \mathbf{c}$. For each candidate solution, TS requires the definition of the set of its neighbors, $\mathcal{N}(\mathbf{c})$. In this paper, a non-tabu vector \mathbf{x} is a neighbor of the candidate solution \mathbf{c} if and only if \mathbf{x} differs with \mathbf{c} in only one position l such that $|\arg(c_l) - \arg(x_l)| = \Delta$, where $\Delta = 2\pi/2^b$ is the difference between two adjacent phase shifts in \mathcal{F}_b . Therefore, the collection of all neighbors of \mathbf{c} is given by

$$\mathcal{N}(\mathbf{c}) = \{\mathbf{x} \notin \mathcal{T}, |\arg(c_l) - \arg(x_l)| = \Delta, \forall l = 1, \dots, N\}. \quad (23)$$

The length of $\mathcal{N}(\mathbf{c})$ depends on the number of IRS elements, as well as the number of quantization bits. For $b > 1$, each element of \mathbf{c} provides, at most, two neighbors. For example, suppose $\arg(c_l)$ corresponds to $\varphi_k = \Delta k$ in \mathcal{F}_b , $k = 0, \dots, 2^b - 1$, then the two neighbors provided by c_l are

$$[\mathbf{x}]_1 = [c_1, \dots, c_{l-1}, e^{j\varphi_{\text{mod}(k-1, 2^b)}}, c_{l+1}, \dots, c_N]$$

and

$$[\mathbf{x}]_2 = [c_1, \dots, c_{l-1}, e^{j\varphi_{\text{mod}(k+1, 2^b)}}, c_{l+1}, \dots, c_N].$$

In contrast, when $b = 1$, \mathcal{F}_1 has only two phase shifts, i.e., $\{0, \pi\}$, with which only one neighbor can be provided by each element of the candidate solution \mathbf{c} . By using the same previous example, it is clearly seen that for $k = 1$, we have $\varphi_1 = \pi$, which implies that \mathcal{F}_1 is left with only one phase shift 0 to define the neighborhood for c_l . Note that to this point, we have only studied the maximum number of neighbors that a candidate may provide. However, because of the tabooing operation, we do not always have the maximum number of neighbors for each candidate solution at each iteration t . The precise expression for the number of neighbors at each iteration is complex compared to its relative importance in this work; hence, we adopt the simplest form. In particular, its approximate value for a given candidate vector \mathbf{c} and quantization bits b is given by

$$\text{card}(\mathcal{N}(\mathbf{c})) \leq 2N - 1\mathbb{1}_{(t>1)} - N\mathbb{1}_{(b=1)},$$

Algorithm 2: TS algorithm

Input : J, h_d, \mathcal{F}_b, N .

- 1 Initialize $c, \mathcal{T} = \mathbf{0}_{N \times L}$, and $noImprovement = 0$.
- 2 Set $x^* = c$.
- 3 **repeat**
- 4 Insert c into \mathcal{T} .
- 5 $noImprovement++$
- 6 Define neighbors $\mathcal{N}(c)$.
- 7 **if** $\mathcal{N}(c) = \emptyset$ **then**
- 8 Re-initialize c .
- 9 **if** $c \in \mathcal{T}$ **then**
- 10 go to 8
- 11 **end**
- 12 go to 4
- 13 **end**
- 14 Find the best neighbor \bar{x} .
- 15 **if** $p(\bar{x}) > p(x^*)$ **then**
- 16 $x^* = \bar{x}$
- 17 $noImprovement = 0$
- 18 **end**
- 19 $c = \bar{x}$
- 20 Pop out the oldest element in \mathcal{T} if \mathcal{T} is full.
- 21 **until** $noImprovement > \rho$

Output: x^*

where $\mathbb{1}_{(\cdot)}$ is an indicator function. The subtraction by 1 for $t > 1$ is to exclude at least one of the candidate's ancestors, which must be tabooed.

The best neighbor \bar{x} in $\mathcal{N}(c)$ is the one with the largest objective value, i.e.,

$$\bar{x} = \arg \max_{x \in \mathcal{N}(c)} p(x). \quad (24)$$

The best solution x^* is updated to \bar{x} if a larger objective value is found, i.e., $p(\bar{x}) > p(x^*)$. Next, the algorithm makes a move from c to \bar{x} , irrespective of the objective value of \bar{x} . We emphasize that, to avoid cycling around the local maxima, this is a mandatory move, even when \bar{x} provides a smaller objective value than c , i.e., $p(\bar{x}) < p(c)$. Finally, \mathcal{T} is updated by pushing in the new candidate solution \bar{x} . If \mathcal{T} is full, the oldest element in the list is popped out. The release of vectors from a tabu list legitimize their consideration of being neighbors of future candidates. The steps above are then iteratively repeated until a termination condition is met. There are several conditions to terminate the algorithm, including termination after a certain number of predefined iterations, termination after a certain number of iterations without an improvement of the best solution, or any combination of conditions. While both can work fine, we opt to use the second criterion in this paper. Note that when using this criterion, to avoid premature termination of the algorithm, a $noImprovement$ counter is reset whenever a new best solution is found; otherwise, it is incremented. The last x^* is returned as the best solution of the algorithm. These steps are summarized in Algorithm 2.

In the first step, TS randomly initializes the candidate solution and defines the control parameters, such as \mathcal{T} , which

is set to an all-zero matrix of appropriate dimensions, and the termination criterion, which is set to zero. Next, the algorithm assumes the randomly generated candidate to be the best solution before starting its optimization. In each iteration, TS starts by tabooing the current candidate solution so that it is not revisited in the next L iterations. The $noImprovement$ counter is then incremented before finding the candidate's neighbors. Note that the quality of the TS solution depends greatly on the extent of exploration over the domain of candidate solutions. Automatic parameter tuning [33], frequency listing [33], and random restart [34] are some techniques that enhance exploration of TS. To achieve the same goal, a precautionary procedure is provided in steps 7–13 of Algorithm 2, where we partially (without resetting \mathcal{T} and $noImprovement$ counter) restart the algorithm whenever the candidate solution lacks neighbors to explore. We specifically choose this strategy because it is less complex than a complete restart. The situation of $\mathcal{N}(c) = \emptyset$ occurs frequently when N is small and L is large; thus, it is important to restart the algorithm with a random candidate to expand its search space, which in turn ensures that a near-optimal solution can be found.

If $\mathcal{N}(c) \neq \emptyset$, TS evaluates each member against the objective function to determine the best neighbor. Whenever a neighbor provides a larger objective value than the current best solution, the algorithm updates the best solution and resets the $noImprovement$ counter, as shown in steps 14–18. In step 19, to avoid cycling, TS moves to the best neighbor and removes the oldest element from its tabu list to accommodate a new vector. Step 21 indicates that at the end of each iteration, the $noImprovement$ control parameter is examined to determine whether it meets the termination condition, upon which the algorithm is terminated; otherwise, a new iteration is started.

B. Complexity Reduction

Note that a large portion of the complexity of the TS algorithm comes from the computation of the objective value of each neighbor. However, a close scrutiny of the working principle of the TS algorithm, that each candidate must differ from all of its neighbors in only one position, reveals that $N - 1$ of each neighbor's metric computations are exactly the same as those of the current candidate. This suggests that the computational complexity of each neighbor can be reduced significantly by employing another memory structure, in which the cumulative computation load of each candidate is temporarily stored so that it can be shared by all of its neighbors. A novel technique called QR-TS has been proposed in [20], which serves this same purpose. Because this work follows this technique, and for the sake of convenience, we briefly describe it while modifying some steps to make it compatible with our problem. By applying QR-decomposition on the matrix J such that $QR = J$, where $Q \in \mathbb{C}^{M \times M}$ and $R \in \mathbb{C}^{M \times N}$ are the unitary matrix and an upper-triangular matrix, respectively, the expression for the objective function can be written as $p(v) = \|Q^H h_d + Rv\|^2$. The objective

value of neighbor \mathbf{x} of candidate \mathbf{c} can then be given by

$$p(\mathbf{x}) = \|\mathbf{Q}^H \mathbf{h}_d + \mathbf{R}\mathbf{x}\|^2 = \|\mathbf{w} - \mathbf{R}_l \xi_l\|^2 \quad (25)$$

$$= \sum_{i=1}^n |\mathbf{w}_i - \mathbf{R}_{i,l} \xi_l|^2 + \sum_{i=n+1}^M |\mathbf{w}_i|^2, \quad (26)$$

where $\mathbf{w} \in \mathbb{C}^{M \times 1} = \mathbf{Q}^H \mathbf{h}_d + \mathbf{R}\mathbf{c}$; ξ_l is the l th and only nonzero element of $\mathbf{c} - \mathbf{x} = [0, \dots, 0, e^{j\arg(\mathbf{c}_l)} - e^{j\arg(\mathbf{x}_l)}, 0, \dots, 0]^T$, and $n = \min(M, l)$ is the number of nonzero elements of \mathbf{R}_l . It is important to note that the vector \mathbf{w} is constant for all neighbors of the same candidate. This insight enables the computation of its cumulative amount, which is then saved in a short-lived memory \mathbf{e} as follows [20]:

$$\mathbf{e} = \left[\sum_{i=2}^M |\mathbf{w}_i|^2, \sum_{i=3}^M |\mathbf{w}_i|^2, \dots, |\mathbf{w}_M|^2, 0 \right]. \quad (27)$$

The short-lived memory is thereafter used by each neighbor to efficiently obtain its metrics by

$$p(\mathbf{x}) = \sum_{i=1}^n |\mathbf{w}_i - \mathbf{R}_{i,l} \xi_l|^2 + \mathbf{e}_n. \quad (28)$$

Up to this point, (28) is already appealing, as it significantly reduces the computational load by performing some of the repetitive operations only once and allows their results to be shared across different neighbors. However, to finalize this technique, we simplify one more operation. It is worth pointing out that in two successive iterations, t and $t+1$, $\mathbf{w}^{(t)}$ and $\mathbf{w}^{(t+1)}$ must differ in only one position \tilde{l} , similar to the candidate–neighbor relationship. From this observation, \mathbf{w} is computed as

$$\mathbf{w}^{(t)} = \begin{cases} \mathbf{Q}^H \mathbf{h}_d + \mathbf{R}\mathbf{c}^{(t)}, & t = 1, \\ \mathbf{w}^{(t-1)} + \mathbf{R}_{\tilde{l}} \xi_{\tilde{l}}, & t > 1, \end{cases} \quad (29)$$

where $\xi_{\tilde{l}}$ is the \tilde{l} th element of $\mathbf{c}^{(t)} - \mathbf{c}^{(t-1)}$. In summary, whenever we compute the objective value of each neighbor as in (24), we start by finding the vector \mathbf{w} , as shown in (29), followed by evaluating the short-lived memory \mathbf{e} . Finally, we find the objective value by using (28).

V. SIMULATION RESULTS

A. Simulation Setup

In this section, we present numerical results to demonstrate the effectiveness of our proposed schemes as well as to validate our analysis. For performance comparison, we adopt state-of-the-art algorithms AO [9], GP [17], and CE [17] as the benchmark schemes, which, to the best of our knowledge, deliver the highest performance–complexity tradeoff among the existing schemes for discrete IRS phase-shift optimization in downlink MISO systems. We consider an indoor environment, where an AP with $M = 6$ antennas located at $(0, 0)$ serves a single-antenna user located at (D, d_y) . The IRS located at $(D, 0)$ is used to help this communication by reflecting phase-adjusted signals in the direction of the user. Unless otherwise stated, the transmit power P at the AP is set to 23 dBm,

whereas the noise power σ^2 at the user is set to -80 dBm. The distance-dependent path loss β for all channel components is modeled as [9]

$$\beta(d, \alpha)[\text{dB}] = -30 - 10\alpha \log_{10}(d), \quad (30)$$

where d is the distance in metres of the individual link, and α is a path loss exponent.

In what follows, any parameter with a subscript AI , IU , and AU denote the belonging of the parameter to the links \mathbf{G} , \mathbf{h}_r , and \mathbf{h}_d , respectively. To account for small-scale fading, we further assume that AP–IRS and IRS–user links follow the Rician fading channel model, particularly due to their line-of-sight (LoS) nature, whereas the direct AP–user link follows the Rayleigh fading channel model. The antenna elements at the AP and IRS are assumed to form half-wavelength uniform linear arrays (ULAs) [12]. The channels \mathbf{G} , \mathbf{h}_r , and \mathbf{h}_d can be given by

$$\mathbf{G} = \sqrt{\frac{\kappa \bar{\beta}_{AI}}{\kappa + 1}} \hat{\mathbf{G}} + \sqrt{\frac{\bar{\beta}_{AI}}{\kappa + 1}} \bar{\mathbf{G}}, \quad (31)$$

$$\mathbf{h}_r = \sqrt{\frac{\kappa \bar{\beta}_{IU}}{\kappa + 1}} \hat{\mathbf{h}}_r + \sqrt{\frac{\bar{\beta}_{IU}}{\kappa + 1}} \bar{\mathbf{h}}_r, \quad (32)$$

$$\mathbf{h}_d = \sqrt{\bar{\beta}_{AU}} \bar{\mathbf{h}}_d, \quad (33)$$

where κ is the Rician factor; $\bar{\beta}$ indicates a linear scale of the path loss β ; $\hat{\mathbf{G}} \in \mathbb{C}^{N \times M}$ and $\hat{\mathbf{h}}_r \in \mathbb{C}^{N \times 1}$ denote LoS components obtained by the product of two array response vectors at the two sides of the link; $\bar{\mathbf{G}} \in \mathbb{C}^{N \times M}$ and $\bar{\mathbf{h}}_r \in \mathbb{C}^{N \times 1}$ denote the non-LoS (NLoS) components, whose entries are distributed as $\mathcal{CN}(0, 1)$. Moreover, $\bar{\mathbf{h}}_d \in \mathbb{C}^{N \times 1}$ denotes the Rayleigh distributed channel vector, whose elements are assumed to be independent and identically distributed zero-mean complex Gaussian random variables with a variance of $1/2$ per dimension. Without loss of generality, algorithm-specific parameters are set as follows: the length of tabu list is set to $L = \min(2^{N_b}, 50)$. TS is halted after $\varrho = 100$ iterations without improvement of the best solution, whereas SD is terminated early after $\varpi = \max(5N, 100)$ iterations without reaching a leaf node.

B. SNR Improvement on a Deep-Fade Direct Link

We set the distances $D = 150$ m and $d_y = 3$ m. The path loss parameters are set to $\alpha_{AI} = 2.2$, $\alpha_{IU} = 2.8$, and $\alpha_{AU} = 3.5$. Rician factors for different links are set to $\kappa_{AI} = \infty$ and $\kappa_{IU} = 10$ [12]. Because we assume a coarse IRS, we set $b = 1$. To model a deep fading scenario, we adopt a penetration loss factor ε , which introduces the necessary attenuation to the commonly suffering link. The penetration loss factor is added to (30). As shown in Fig. 2, we consider the link \mathbf{h}_d to be in a deep fade as the AP and user are blocked by a 12-inch thick concrete wall, hence on the basis of [36], we set $\varepsilon_{AU} = -70.4$ dB. To demonstrate the optimality of the proposed algorithms, in Fig. 3 we compare their performance to that of the exhaustive search. However, due to the excessive computational load of the exhaustive search, we only consider it for a small number of IRS elements, i.e.,

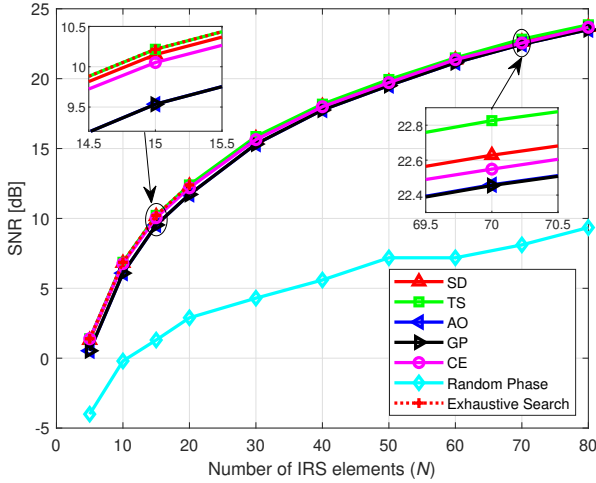


Fig. 3. SNR versus the number of IRS elements for an IRS-assisted downlink MISO system with $M = 6$, $b = 1$, and $\varepsilon_{AU} = -70.4$ dB.

$N = \{5, 10, 15, 20\}$. Apart from benchmark schemes, we also include the random-phase scheme, where the phase shift of each element is determined by a random selection from \mathcal{F}_b . In Fig. 3, it is seen that when the phase shifts of the IRS elements are chosen randomly, the performance becomes very poor, as expected. This simply demonstrates the absolute importance of proper phase shift optimization to fully exploit the potential of the IRS. Second, it is clearly observed that for a small number of IRS elements, our proposed algorithms, i.e., SD and TS, nearly attain the optimal performance of the exhaustive search. However, when the size of IRS increases, the performance of CE is observed to drop marginally with respect to that of TS. Furthermore, we note that CE requires significantly higher complexity compared to our proposed schemes, as shown in Section V.G. The performance gains of both the proposed and benchmark schemes over that of random phase scheme increase to 15 dB when $N = 80$, which demonstrates that the importance of IRS phase shift optimization is more evident with large IRS size. Another interesting observation is that, despite of ignoring the direct link, the SD's performance loss with respect to TS is almost negligible. This further validates our assumption that when the strength of the deep-fade direct AP–user link is very small compared to the composite AP–IRS–user link, it can be safely ignored.

C. Convergence Behavior of TS

We provide the convergence behavior of our proposed TS algorithm in Fig. 4. For this purpose, we set $M = 6$, $N = 50$, $b = 2$, and $\varepsilon_{AU} = 0$ dB. Fig. 4 shows that the convergence of the TS algorithm is monotonic and converges at 40 iterations. Notably, the TS algorithm is designed to store the best-found solution at each iteration. This attribute guarantees the monotonic convergence behavior, as seen in Fig. 4.

D. Performance on a Normal Direct AP–User Link

Next, we examine the performance of our proposed algorithms in normal wireless propagation environments. Except

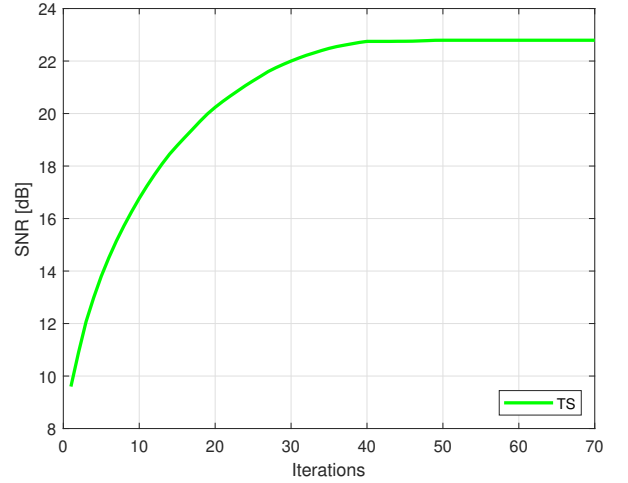


Fig. 4. TS convergence for a downlink MISO system with $M = 6$, $N = 50$, $b = 2$, and $\varepsilon_{AU} = 0$ dB.

for the attenuation factor ε_{AU} , which is ignored in this subsection, the rest of parameters remain the same as those used in Fig. 3. Fig. 5 plots the SNRs of all the algorithms versus the number of IRS elements for quantization bits $b = \{1, 2\}$. We first observe that for $b = 1$, TS attains the highest performance amongst all the schemes. Compared to the case when the direct link is severely attenuated in Fig. 3, there is almost negligible improvement in terms of performance. This implies that in wireless communications, the LoS channels, especially when properly directed to the destination as in this case, deliver stronger signals than the NLoS channels. This is also the reason that SD, which ignores the NLoS direct link, still maintains its good performance over the benchmark schemes.

For $b = 2$, the performance of the proposed algorithms is almost indistinguishable from that of AO and GP. Nevertheless, the performance of CE seems to degrade significantly, especially for a large IRS. In our analysis, we note that this loss is highly contributed by the difficulty on the proper selection of algorithm parameters like the smoothing strategy that prevents premature convergence, the update rule of the tilting parameter, the choice of the numbers of candidate and elite samples, and the maximum number of iterations (t_{\max}). The dependency of CE on these many parameters not only degrades its performance upon improper parameter settings, but also makes the algorithm selective of the propagation environment and system setups.

Another important observation from Fig. 5 is that when the number of quantization bits increases from 1 to 2, a performance gain of approximately 2.5 dB is obtained for the entire range of the number of IRS elements. However, it comes at the expense of increased power consumption and complexity on the design of IRS elements as well as the optimization of their phase shifts. Therefore, in the design of IRS-assisted MIMO systems, it is necessary to properly balance the desired performance versus the complexity and power consumption of the system.

TABLE I
Computational complexity for SD, TS, AO, GP, and CE algorithms.

| Algorithm | Operation | FLOPS | Frequency |
|-----------|--|--|-----------|
| SD | SD problem formulation | $7N^3/3 + 0.5N^2 - M^2 + NM(2N + 2M + 1) + 13N/6$ | 1 |
| | SE-ordering | $2N + 4 \text{ card}(\mathcal{F}_b)$ | t |
| | Solution saving and radius update | $N^2 + 1.5N - 1$ | t |
| TS | $\mathbf{J} = \text{diag}(\mathbf{h}_r) \mathbf{G}^T$ | NM | 1 |
| | QR-decomposition of \mathbf{J} | $2N^2(M - N/3)$ | 1 |
| | $\mathbf{w} = \mathbf{Q}^H \mathbf{h}_d + \mathbf{R}\mathbf{c}$ | $2M^2 + 2MN - M$ | 1 |
| | Neighbor's objective value computation | $3M + 5 \text{ card}(\mathcal{N}(\mathbf{e})) - 3$ | t |
| AO | $\bar{\mathbf{w}} = \mathbf{h}_d / \ \mathbf{h}_d\ $ | $3M - 1$ | 1 |
| | $\gamma_0 = \arg(\mathbf{h}_d^H \bar{\mathbf{w}})$ | $2M$ | t |
| | $\mathbf{a} = \text{diag}(\mathbf{h}_r^H) \mathbf{G} \bar{\mathbf{w}}$ | $3NM - N$ | t |
| | $\mathbf{v}^* = e^{j(\gamma_0 - \arg(\mathbf{a}))}$ | $3N$ | t |
| | $\bar{\mathbf{w}} = \frac{(\mathbf{h}_d^H + \mathbf{h}_r^H \Phi \mathbf{G})^H}{\ \mathbf{h}_d^H + \mathbf{h}_r^H \Phi \mathbf{G}\ }$ | $2MN + 3M + N - 1$ | t |
| | Quantization of \mathbf{v}^* to the nearest phase shift in \mathcal{F}_b | $N^2 + N$ | 1 |
| GP | $\mathbf{A} = \text{diag}(\mathbf{h}_r^H) \mathbf{G} \mathbf{G}^H \text{diag}(\mathbf{h}_r^H)$ | $N^2(2M - 1) + NM$ | 1 |
| | $\mathbf{b} = \text{diag}(\mathbf{h}_r^H) \mathbf{G} \mathbf{h}_d$ | $3NM - N$ | 1 |
| | Eigenvalue decomposition to obtain λ_{\max} | N^3 | 1 |
| | $\bar{\mathbf{A}} = 2\mu\mathbf{A}$ and $\bar{\mathbf{b}} = 2\mu\mathbf{b}$ | $N^2 + N$ | 1 |
| | Update procedure | $2N^2 + N$ | t |
| | Objective value computation | $2N^2 + 7N - 3$ | t |
| CE | Computation of the initial probability matrix $\mathbf{P}^{(0)}$ | N | 1 |
| | Objective value computation of each sample | $2N^2 + 7N - 3$ | tS |
| | Update of probability matrix \mathbf{P} | N | tS |

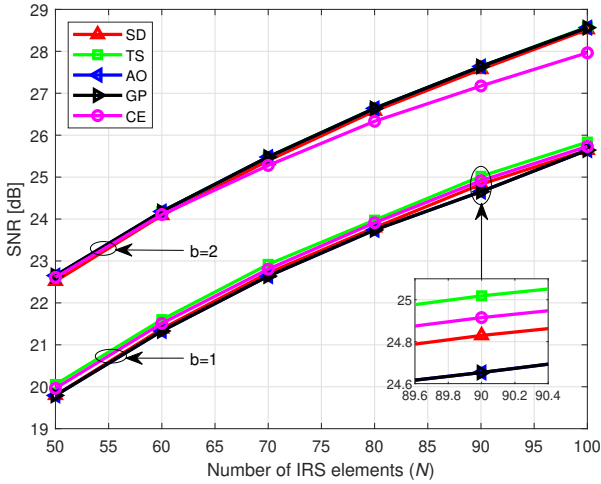


Fig. 5. SNR versus the number of IRS elements for a downlink MISO system with $M = 6$, $b = \{1, 2\}$, and $\varepsilon_{AU} = 0$ dB.

E. SNR versus Quantization Bits

To cement the discussion ignited in the previous subsection about the effect of quantization bits on the performance of IRS-assisted MIMO systems, we analyze it in different system setups. We use same system parameters as those used in the previous subsection although to capture different system setups, the set of quantization bits $b = \{1, 2, 3, 4\}$ is analyzed for $N = \{30, 50\}$. In Fig. 6, where we plot SNR versus the number of quantization bits for different algorithms, we also include the performance of the continuous phase, which is the unquantized solution of AO, to act as an upper bound. From Fig. 6, we observe that for both $N = 30$ and $N = 50$, our

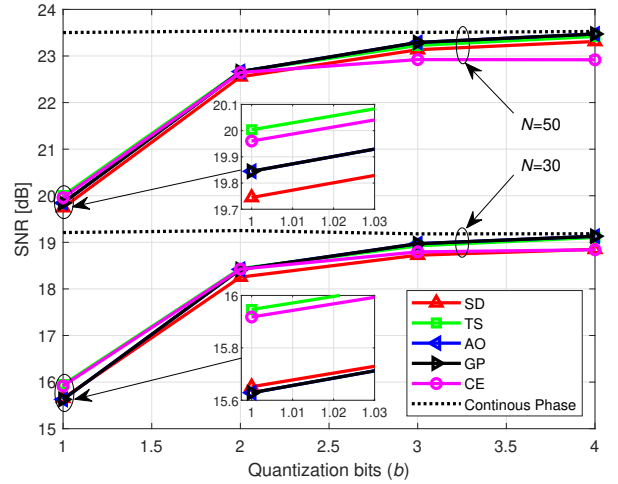


Fig. 6. SNR versus the number of quantization bits for a downlink MISO system with $M = 6$ and $N = \{30, 50\}$.

proposed algorithms attain better performance than benchmark schemes for low-resolution IRS, whereas with high-resolution IRS, their performance indistinguishably drops below that of AO and GP.

F. Effect of IRS and User Locations

In this subsection, we evaluate the effect of the IRS and user locations on the performance of IRS-assisted MISO systems. Specifically, we start by moving the user on the range $\vartheta = [100, 200]$ m, whereas IRS is fixed at its original position $(D, 0)$. In the second phase, IRS is moved over the same range ϑ , whereas the user is reset to its initial location

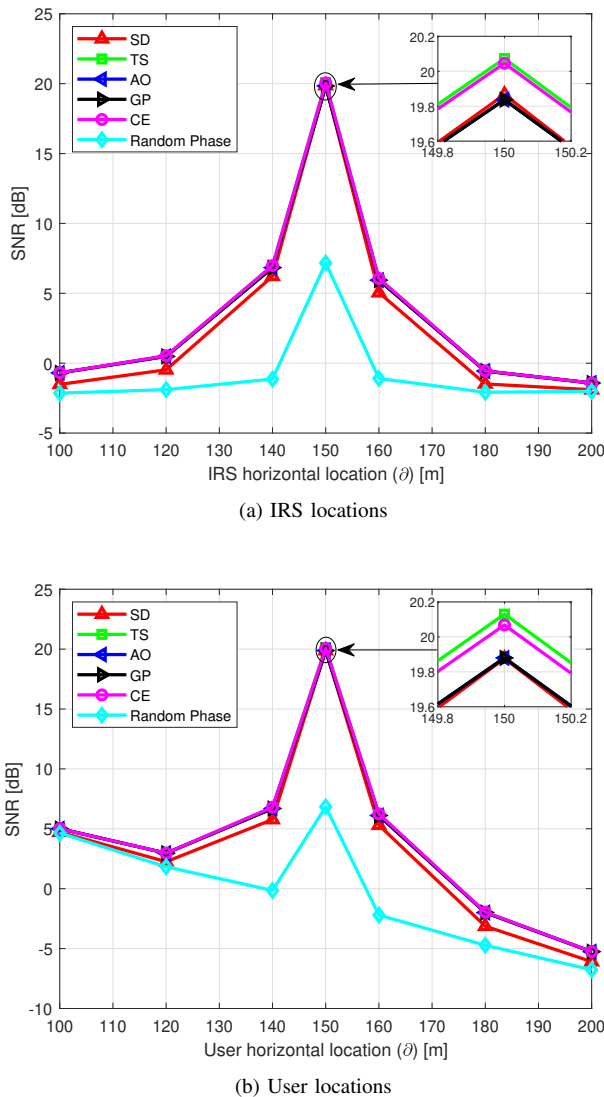


Fig. 7. SNR versus location of IRS and user for an IRS-assisted downlink MISO system with $M = 6$, $N = 50$ and $b = 1$.

(D, d_y). We assume $N = 50$, $b = 1$, and no penetration loss. The rest of parameters remain the same as in Section V.B. In Fig. 7, the SNRs of the proposed and benchmark schemes are plotted against various IRS/user horizontal locations. First, we observe that when both the IRS and user have the shortest separation between them, all the schemes provide the maximum performance. This is expected because of the small path loss on the IRS–user link. We also note that a higher SNR is attained when the user is close to the AP than when IRS is close to the AP; this is because when a user is close to the AP, the gain on the strength of the direct link (due to reduced path loss on the AP–user link) is relatively large compared to the loss of strength (due to the increased path loss resulted from the increased distance) on the IRS–user link. By contrast, when IRS is close to the AP, the gain of the strength on the AP–IRS link is comparable to the loss on the IRS–user link, hence neutralizes the effect of IRS being close to the AP. Moreover, we can clearly see from both Figs. 7(a) and

7(b) that, irrespective of IRS/user locations, our proposed TS algorithm provides higher SNRs than the benchmark schemes.

G. Computational Complexity

Furthermore, we investigate the complexity of the proposed algorithms and benchmark schemes in terms of the number of floating-point operations (FLOPS) [28], [37]. In Table I, where we summarize the number of FLOPS needed by each main operation of each algorithm, t denotes the number of times each operation is performed. Note that for fair comparisons, we have also included the GP’s computation load N^3 for finding λ_{\max} , which was not included in its original work [17]. For the CE algorithm, the notation S denotes the number of samples of candidate solutions. We further assume that the QR Householder method [28] is used to perform QR decomposition for the TS algorithm. In Fig. 8, we compare the complexity of different schemes under different system configurations by using the same parameters as those used in Fig. 3, with the exception of ε_{AU} , which is ignored.

Specifically, Figs. 8(a)–8(c) plot the FLOPS against the number of IRS elements for $b = 1$, IRS horizontal distance ϑ for $b = 1$ and $N = 50$, and number of quantization bits for $N = 30$, respectively. We observe from Figs. 8(a)–8(c) that the complexities of the proposed algorithms are lower than those of CE and GP and higher than that of AO. AO has the smallest complexity because it does not directly optimize the discrete phases. Instead, it quantizes its continuous solution to the nearest discrete phase shift in \mathcal{F}_b . Furthermore, unlike GP, which accumulates most of its complexities from the computation of the objective value and its two update procedures [17, eq. (8a) and (8b)], AO computes the objective value only in each iteration. In contrast, the complexity reduction of the proposed algorithms over GP and CE is remarkable. For instance, in Fig. 8(a), we observe that for $N < 50$, the complexity of SD is smaller than those of TS, CE, and GP. However, as N increases beyond 50, the SD’s complexity surpasses that of TS although it still remains below those of GP and CE. Quantitatively, we note that when $N = 50$, the gains in complexity reduction of TS and SD over GP are 99.28% and 99.03%, respectively, whereas those over CE are 99.32% and 99.08%, respectively.

In Fig. 8(b), we observe that when IRS is moved along different positions, the complexity of GP also changes significantly, with the highest complexity obtained for the IRS and user being closest to each other. This causes the gain in complexity reduction of the proposed algorithms over GP to vary significantly. For example, when IRS is at $\vartheta = 100$ m, the complexity-reduction ratios of TS and SD over GP are 78.74% and 70.58%, respectively, whereas at $\vartheta = 150$ m they increase to 99.21% and 98.91%, respectively.

Finally, from Fig. 8(c), it is clearly seen that the complexities of all the schemes are less sensitive to the number of quantization bits. This is due to the design of these algorithms. For example, TS and SD use termination criteria, which are independent of b ; likewise, CE is also capped at a predefined number of maximum iterations, whereas AO and GP only quantize their continuous solutions to the set of discrete phase

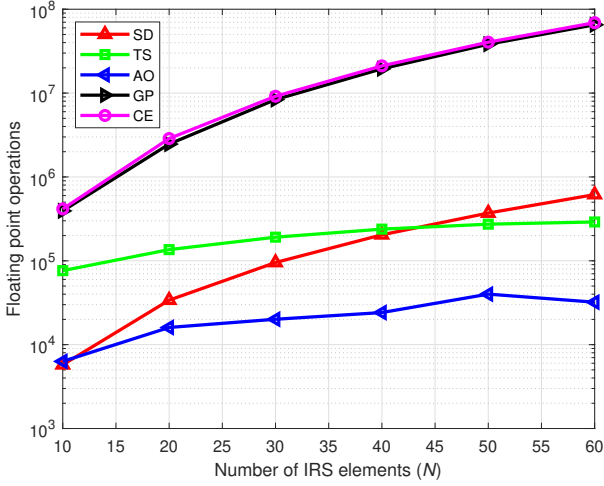
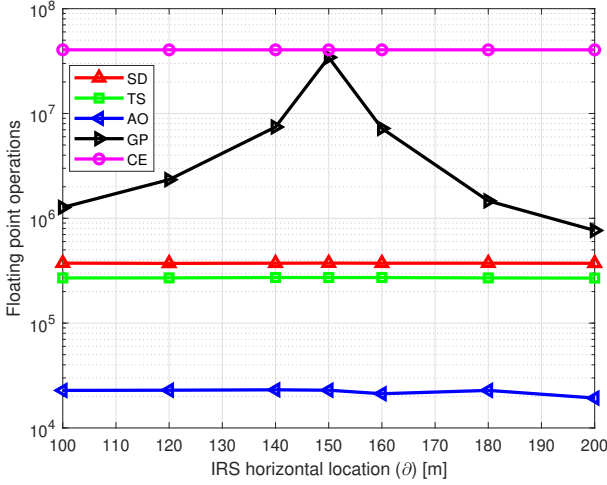
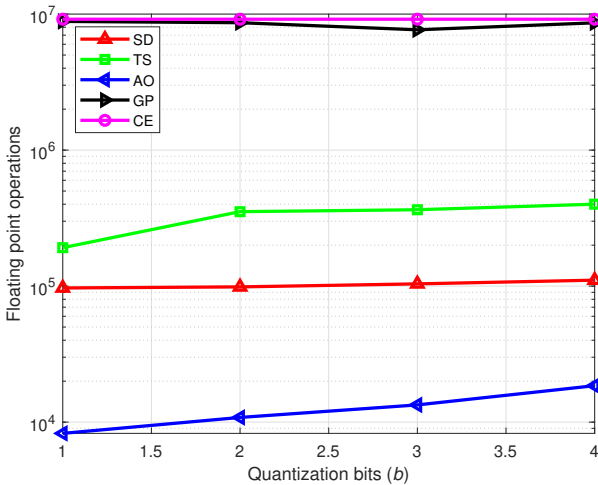
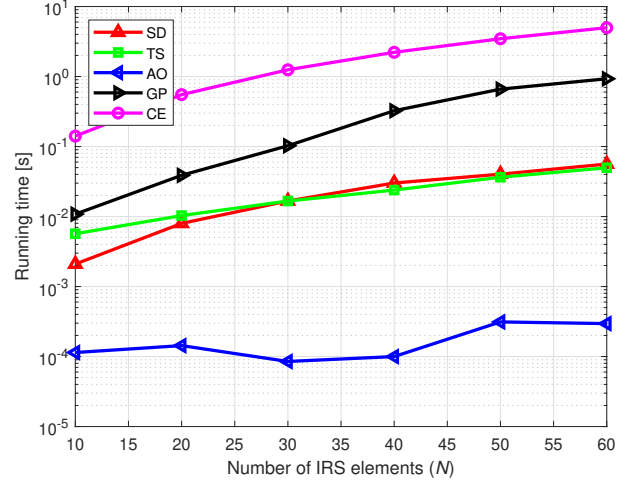

 (a) $b = 1$

 (b) $b = 1, N = 50$

 (c) $N = 30$

 Fig. 8. Comparison of computational complexity of proposed algorithms with benchmark schemes for an IRS-assisted downlink MISO system with $M = 6$ and $\varepsilon_{AU} = 0$ dB.

 Fig. 9. Running time versus the number of IRS elements for a downlink MISO system with $M = 6$, $b = 1$, and $\varepsilon_{AU} = 0$ dB.

shifts. Despite of that, the proposed algorithms still attain substantial complexity reductions up to 97.83% and 97.91% for TS over GP and CE, respectively, whereas those of SD over GP and CE are 98.90% and 98.94%, respectively.

H. Running Time

Finally, we compare the running time of the proposed algorithms with the benchmark schemes for various sizes of the IRS. The results in Fig. 9 are obtained by using the same parameters as those used in Fig. 5 for $b = 1$. The simulation was performed by MATLAB R2021a running on a Windows 11 computer having an Intel(R) Core(TM) i7-10700 CPU @ 2.90 GHz processor and 16 GB of RAM. From Fig. 9, we observe that the CE algorithm has the longest running time of all the tested schemes, whereas AO has the shortest. Furthermore, the average running times of the proposed algorithms are still shorter than those of GP and CE.

VI. CONCLUSION

In this study, we have investigated an IRS-assisted downlink MISO system with the aim of enhancing the signal strength at the user. To achieve this, the IRS reconfiguration problem has been addressed in two stages. First, a narrowed analysis, which ignores a severely attenuated NLoS direct AP–user link, has been studied. We have then generalized the analysis to the case of a normal AP–user link. In the former scenario, the IRS phase-shift problem is reformulated so that it can be efficiently solved by the proposed SD algorithm, whereas in the latter scenario, we have resorted to the original joint active and passive beamforming problem and proposed the TS algorithm. Both proposed algorithms have been carefully designed to deliver high performance with significant reduction in complexity compared to the conventional GP and CE schemes. Specifically, the complexity of SD is regulated by an early-termination criterion, which halts the algorithm whenever slow

convergence is detected, especially when there is a large number of close-to-optimal lattice points. Furthermore, TS uses a novel QR-TS technique to reduce its complexity by reusing the computational results of the candidate solution in computing the neighbor's metrics. Extensive simulation results verify the validity of our analysis as well as prove the effectiveness of our proposed algorithms over the benchmark schemes. In particular, TS provides a 1-dB SNR gain over AO and GP when the direct AP–user link is in deep fade. The proposed schemes not only deliver higher performance than the benchmark schemes, but also attain near-optimal performance of an exhaustive search for small-size IRS. More importantly, such high performance in different system setups is obtained with significant complexity reduction. For example, complexity-reduction ratios of 99.28% and 99.03% have been obtained by TS and SD, respectively, over GP when the phase shifts of the 50-element IRS are optimized in a normal wireless propagation environment. Furthermore, the phase-shift optimization by SD for a 30-element IRS with quantization bits $b = \{1, 2, 3, 4\}$ only requires 1.1% and 1.06% of the computational loads of GP and CE, respectively.

Despite performance gains provided by our proposed algorithms, their design relies on the assumption of perfect CSI knowledge at the AP. Also, the algorithms have been designed and analyzed for single-user systems only. However, in practical systems, getting perfect CSI can be challenging, especially when there are multiple users in the network. Therefore, the extension of the proposed algorithms to the IRS-assisted multi-user wireless communication system with imperfect CSI can be an interesting direction for future research.

REFERENCES

- [1] Ericsson, *Mobile data traffic forecast – Mobility Report*. (2021), [Online]. Available: <https://www.ericsson.com/en/mobility-report/dataforecasts/mobile-traffic-forecast> (visited on 06/01/2021).
- [2] S. Zhang, Q. Wu, S. Xu, and G. Y. Li, “Fundamental green tradeoffs: Progresses, challenges, and impacts on 5G networks,” *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 33–56, 2017.
- [3] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2020.
- [4] M. Cui, G. Zhang, and R. Zhang, “Secure wireless communication via intelligent reflecting surface,” *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, 2019.
- [5] J. Chen, Y.-C. Liang, Y. Pei, and H. Guo, “Intelligent reflecting surface: A programmable wireless environment for physical layer security,” *IEEE Access*, vol. 7, pp. 82 599–82 612, 2019.
- [6] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, “Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications,” *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, 2019.
- [7] S. Guo, S. Lv, H. Zhang, J. Ye, and P. Zhang, “Reflecting modulation,” Dec. 2018. arXiv: 1912.08428v2. [Online]. Available: <http://arxiv.org/abs/1912.08428v2>.
- [8] W. Yan, X. Yuan, and X. Kuai, “Passive beamforming and information transfer via large intelligent surface,” *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 533–537, 2020.
- [9] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, 2019.
- [10] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, “A new wireless communication paradigm through software-controlled metasurfaces,” *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, 2018.
- [11] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design,” in *Proc. IEEE GLOBECOM*, 2018.
- [12] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, “Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, 2020.
- [13] C. Huang, A. Zappone, M. Debbah, and C. Yuen, “Achievable rate maximization by passive intelligent mirrors,” in *Proc. IEEE ICASSP*, 2018.
- [14] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, 2019.
- [15] Q. Wu and R. Zhang, “Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts,” *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, 2020.
- [16] X. Yu, D. Xu, and R. Schober, “MISO wireless communication systems via intelligent reflecting surfaces : (invited paper),” in *Proc. IEEE/CIC ICC*, 2019.
- [17] J.-C. Chen, “Beamforming optimization for intelligent reflecting surface-aided MISO communication systems,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 504–513, 2021.
- [18] X. Yu, D. Xu, and R. Schober, “Optimal beamforming for MISO communications via intelligent reflecting surfaces,” Jan. 2020. arXiv: 2001.11429v1. [Online]. Available: <http://arxiv.org/abs/2001.11429v1>.
- [19] E. Björnson, Ö. Özdogan, and E. G. Larsson, “Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?,” *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, 2020.
- [20] N. T. Nguyen, K. Lee, and H. Dai, “QR-decomposition-aided tabu search detection for large MIMO systems,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4857–4870, 2019.
- [21] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, “Coding metamaterials, digital metamaterials and programmable metamaterials,” *Light Sci. Appl.*, vol. 3, no. 10, e218–e218, Oct. 2014.
- [22] L. Zhang *et al.*, “Space-time-coding digital metasurfaces,” *Nat. Commun.*, vol. 9, no. 1, p. 4334, Oct. 2018.
- [23] P. Nayeri, F. Yang, and A. Z. Elsherbeni, *Reflectarray Antennas: Theory, Designs, and Applications*. John Wiley & Sons, Ltd, 2018.
- [24] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [25] O. Damen, A. Chkeif, and J.-C. Belfiore, “Lattice code decoder for space-time codes,” *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161–163, 2000.
- [26] B. Hassibi and H. Vikalo, “On the sphere-decoding algorithm I. Expected complexity,” *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2806–2818, 2005.
- [27] B. Shim and I. Kang, “Sphere decoding with a probabilistic tree pruning,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4867–4878, 2008.
- [28] G. H. Golub and C. F. Van Loan, *Matrix computations*, in Baltimore, MD, USA: Johns Hopkins University Press, 2012.
- [29] L. G. Barbero and J. S. Thompson, “A fixed-complexity MIMO detector based on the complex sphere decoder,” in *Proc. IEEE ICASSP*, 2006.
- [30] C. P. Schnorr and M. P. Euchner, *Lattice basis reduction: Improved practical algorithms and solving subset sum problems*. Math. Program., 1994, vol. 66.

- [31] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, 2002.
- [32] M. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, 2003.
- [33] C. Lagos, B. Crawford, R. Soto, *et al.*, "Improving tabu search performance by means of automatic parameter tuning," *Can. J. Electr. Comput. Eng.*, vol. 39, no. 1, pp. 51–58, 2016.
- [34] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Random-restart reactive tabu search algorithm for detection in large-MIMO systems," *IEEE Commun. Lett.*, vol. 14, no. 12, pp. 1107–1109, 2010.
- [35] N. Srinidhi, T. Datta, A. Chockalingam, and B. S. Rajan, "Layered tabu search algorithm for large-MIMO detection and a lower bound on ML performance," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2955–2963, 2011.
- [36] W. C. Stone, "Electromagnetic signal attenuation in construction materials," Tech. Rep., Oct. 1997.
- [37] R. Hunger, *Floating point operations in matrix-vector calculus*. Munich University of Technology, Inst. for Circuit Theory and Signal Processing, 2005.



Seraphin F. Kimaryo received the B.Sc. degree (Hons.) in Telecommunications Engineering from the College of Information and Communication Technologies, University of Dar es Salaam, Tanzania, in 2016, and the M.S. degree in Electrical and Information Engineering from the Seoul National University of Science and Technology, South Korea, in 2022, where he is currently pursuing the Ph.D. degree in Electrical and Information Engineering. His research interests include signal processing, optimization, and applied machine learning for wire-

less communication with a focus on massive multiple-input multiple-output systems.



Kyungchun Lee (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, in 2000, 2002, and 2007, respectively. From April 2007 until June 2008, he was a Post-Doctoral Researcher with the University of Southampton, U.K. From July 2008 to August 2010, he was with Samsung Electronics, Suwon, South Korea. Since September 2010, he has been with the Seoul National University of Science and Technology, South Korea. In 2017,

he was a Visiting Assistant Professor with North Carolina State University, Raleigh, NC, USA. His research interests include wireless communications and applied machine learning. He received the Best Paper Awards at the IEEE International Conference on Communications (ICC) and IEEE Wireless Communications and Networking Conference (WCNC) in 2009 and 2020, respectively.