Reconfigurable Intelligent Surfaces Based on Single, Group, and Fully Connected Discrete-Value Impedance Networks

Matteo Nerini, Bruno Clerckx, Fellow, IEEE

Abstract

Reconfigurable intelligent surfaces (RISs) allow to control the propagation environment in wireless networks by properly tuning multiple reflecting elements. Traditionally, RISs have been realized through a single connected architecture, where each RIS element is controlled by an impedance connected to ground. In a recent work, this architecture has been generalized by realizing RISs through group and fully connected impedance networks. However, impedance networks reconfigurable with arbitrary precision are hard to realize in practice. In addition, it is still unexplored how to group together the RIS elements in group connected architectures. These two problems are addressed in this paper. Firstly, we propose a RIS design strategy based on reconfigurable impedance networks with discrete values. Secondly, we present three approaches to design the grouping strategy in group connected RISs. Numerical results show that fewer resolution bits are necessary to achieve the performance upper bound as the group size increases. While four resolution bits are needed in single connected architectures, only a single resolution bit is sufficient in fully connected ones. In addition, we show that by dynamically optimizing the grouping strategy, RISs with group size 4 nearly achieve the same performance as fully connected RISs, with reduced hardware complexity.

Index Terms

Reconfigurable intelligent surface, Discrete-value impedance network, Pattern search, $k$-means clustering, Alternating optimization.

M. Nerini and B. Clerckx are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: m.nerini20, b.clerckx@imperial.ac.uk).
I. INTRODUCTION

Reconfigurable intelligent surfaces (RISs), or intelligent reflecting surfaces, are an emerging technology that will enhance the performance of future wireless communications [1], [2], [3]. This technology relies on large planar surfaces comprising multiple reflecting elements, each of them able to induce a certain amplitude and phase change to the incident electromagnetic wave. Thus, an RIS can steer the reflected signal towards the intended direction by smartly coordinating the reflection coefficients of its elements. RIS-aided communication systems benefit from three main advantages. Firstly, the received signal power is increased by means of the reconfigurable propagation environment created by the RIS. Secondly, less inter-user interference is experienced in multi-user systems and, consequently, a larger capacity region is achieved. Finally, the outage probability is decreased since RISs can better serve users in dead zones and at the cell edge. Beyond these conceptual benefits, RISs are a cost-effective solution, since they are composed solely of passive reflective elements, with no expensive radio frequency (RF) chains.

To avoid difficult optimization problems, many studies on RISs do not pose limitations on the allowed reflection coefficient values. However, in practical implementation, they are selected from a finite number of discrete values. Indeed, reflection coefficients tunable with finer resolution require a more complex hardware design, which can be prohibitive when the number of RIS elements is high [4]. Considering a single-user RIS-aided single-input single-output (SISO) system, the effects of discrete phase shifts have been investigated on the diversity order [5], [6], on the achievable rate [7], and on the ergodic capacity [8]. Furthermore, in [9], [10], [11], channel estimation strategies have been proposed for discrete phase shift RISs.

Several methods to design RISs with discrete reflection coefficients have been recently presented with different objectives. Discrete reflection coefficients have been optimized to enhance the performance of multiple-input multiple-output (MIMO) communications by solving rate maximization problems in single-user systems [12], [13], [14], [15] and sum-rate maximization problems in multi-user systems [16], [17], [18], [19], [20], [21]. In [14], besides the discrete reflection coefficients assumption, a practical model capturing the phase-dependent amplitude variation in the reflection coefficients is considered. RISs with discrete reflection coefficients have been applied also with the objective of enhancing resource utilization. In [22], [23], the transmit power is minimized by jointly optimizing the continuous transmit precoding and the discrete phase shifts at the RIS. For RIS-aided downlink communications, works have been conducted
to maximize energy efficiency [24], [25], and the spectral efficiency in the presence of statistical channel state information (CSI) [26]. In RIS-aided communications, orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA) schemes have been compared in terms of minimum transmit power required with given user rates in [27]. With a focus on RIS-aided NOMA systems, the minimum user rate maximization problem has been considered in [28], and the outage probability minimization problem has been studied in [29]. In [30], the joint usage of a decode-forward relay together with a discrete phase shift RIS has been studied to significantly improve the achievable rate. In addition, RISs with discrete reflection coefficients have been investigated for Terahertz (THz) communications [31], [32], [33], [34], wideband communications [35], and simultaneous wireless information and power transfer (SWIPT) [36], [37], [38]. Finally, prototypes of discrete phase shift RISs have been designed in [39], [40].

In the aforementioned literature, it is always assumed that each RIS element is independently controlled by a tunable impedance connected to ground [41]. Recently, this widely adopted single connected architecture has been generalized by connecting all or a subset of RIS elements through a reconfigurable impedance network, resulting in the fully and group connected architecture, respectively [42]. These architectures lead to non-diagonal scattering matrices, in contrast with the conventional diagonal scattering matrix of the single connected architecture. Because of the additional degrees of freedom provided by the more complex architecture, the fully connected architecture enables the best performance gain with respect to all other RIS models proposed to date [42]. However, the group and fully connected architectures have been studied with no practical constraints on the reflecting elements, and it is not clear how to design these architecture with discrete reflection coefficients. On the one hand, continuous reflection coefficients can be optimized with methods based on the partial derivatives of the objective function, or their approximations. On the other hand, the optimization of discrete reflection coefficients poses two subproblems: Firstly, the construction of a discrete codebook; secondly, the search of the optimal discrete values within the codebook entries. In addition, in the case of group connected architectures, how to group together the RIS elements is still an open design challenge.

Driven by the success of these novel RIS architectures, in this paper, we propose a design strategy for group and fully connected RISs with discrete reflection coefficients. We generalize the well-established approach based on discrete phase shifts valid for the single connected architecture to the group and fully connected ones. Furthermore, in the case of group connected architectures, we investigate how to group together the RIS elements to maximize the performance with...
TABLE I

|                  | Single connected RISs | Group connected RISs | Fully connected RISs |
|------------------|-----------------------|----------------------|----------------------|
| Continuous       | Traditional literature: | [42].                | [42].                |
| reflection       | The RIS has continuous | The RIS has continuous | The RIS has continuous |
| phase coefficients | phase shifts in [0, 2\pi). | reactance matrix entries in \mathbb{R}. | reactance matrix entries in \mathbb{R}. |
| Discrete         | Traditional literature: | This work:           | This work:           |
| reflection       | The RIS has discrete | The RIS has discrete | The RIS has discrete |
| phase coefficients | phase shifts in \{0, \frac{2\pi}{2^M}, \ldots, (2^M - 1) \frac{2\pi}{2^M}\}. | reactance matrix entries. | reactance matrix entries. |

reduced hardware complexity. The contributions of this paper are summarized as follows.

First, we propose a RIS design strategy based on discrete-value reconfigurable impedance networks. In Tab. I, we provide a taxonomy of studies on RIS to better contextualize our contribution. The traditional literature on the single connected architecture with discrete phase shifts is extended in this work to the group and fully connected cases. More precisely, firstly we propose an RIS design based on scalar quantization, assigning a finite number of resolution bits to each reconfigurable impedance. Then, a second discretization strategy based on vector quantization is derived, in which groups of discrete-value reconfigurable impedances are jointly optimized. RISs designed with these two strategies are denoted as scalar-discrete RISs and vector-discrete RISs, respectively. Vector-discrete RISs achieve higher performance than scalar-discrete RISs at the cost of increased optimization computational complexity.

Second, we address the problem of grouping the RIS elements in the group connected architecture, namely, the grouping strategy. We describe two approaches to design the grouping strategy. The former is based on the channel statistics, i.e., the grouping strategy is fixed in time. The latter is based on the channel realization, i.e., the grouping strategy is dynamically optimized. To dynamically optimize the grouping strategy at run-time, we propose a practical iterative algorithm since the exhaustive search has prohibitive complexity. When this algorithm is used to design the grouping strategy, group connected RISs with group size 4 nearly achieve the same performance as fully connected RISs, with reduced hardware complexity.

Third, we assess the performance in terms of received signal power of single, group, and fully connected RISs employing different numbers of resolution bits. The performance of these practical RISs is compared with the performance of ideal RISs and with a novel received signal
power upper bound valid for RIS-aided MIMO systems. We verify that the more connections are present in the reconfigurable impedance network, the less resolution is needed to achieve the optimality. Because of this property, in the fully connected architecture, a single resolution bit allocated to each reconfigurable impedance is sufficient to achieve the performance of ideal RISs. In a practical development, this simplifies significantly the hardware complexity of the fully connected RIS. Furthermore, the received signal power obtained with fully connected RISs can nearly reach the proposed performance upper bound.

Fourth, we compare the single connected, group connected, and fully connected architectures when the number of total resolution bits is limited. On the one hand, the single connected is the simplest architecture having the lowest number of impedances to be optimized. On the other hand, the fully connected architecture includes a larger reconfigurable impedance network, but it allows, at the same time, better performance with lower resolution per reconfigurable impedance. Between these two extremes, a trade-off is given by the group connected architecture. We show how single, group, and fully connected architectures should be used to maximize the received signal power under a constrained number of total resolution bits.

Organization: In Section II, we introduce the system model and the properties of the three RIS architectures. In Section III, we present our novel RIS design strategy based on discrete-value impedance networks. In Section IV, we address the problem of grouping the RIS elements in the case of group connected architecture. In Section V, we assess the performance of RISs based on discrete-value impedance networks in terms of received signal power. Finally, Section VI contains the concluding remarks.

Notation: Vectors and matrices are denoted with bold lower and bold upper letters, respectively. Scalars are represented with letters not in bold font. $|a|$, and $\arg(a)$ refer to the modulus and phase of a complex scalar $a$, respectively. $[a]_i$ and $\|a\|$ refer to the $i$-th element and $l_2$-norm of vector $a$, respectively. $A^T$, $A^H$, $[A]_{i,j}$, and $\|A\|$ refer to the transpose, conjugate transpose, $(i,j)$-th element, and spectral norm of matrix $A$, respectively. $\mathbb{R}$ and $\mathbb{C}$ denote the real and complex number sets, respectively. $j = \sqrt{-1}$ denotes the imaginary unit. $0$ and $I$ denote an all-zero matrix and an identity matrix, respectively, with appropriate dimensions. $\mathcal{CN}(0,\mathbb{I})$ denotes the distribution of a circularly symmetric complex Gaussian random vector with mean vector $0$ and covariance matrix $\mathbb{I}$ and $\sim$ stands for “distributed as”. $\text{diag}(a_1,\ldots,a_N)$ refers to a diagonal matrix with diagonal elements being $a_1,\ldots,a_N$. $\text{diag}(A_1,\ldots,A_N)$ refers to a block diagonal matrix with blocks being $A_1,\ldots,A_N$. 

September 20, 2022 DRAFT
II. SYSTEM MODEL

Let us consider a single-user RIS-aided MIMO system, as represented in Fig. 1. We denote as $N_T$ the number of antennas at the transmitter, $N_R$ the number of antennas at the receiver, and $N_I$ the number of antennas at the RIS. The $N_I$ antennas of the RIS are connected to a $N_I$-port reconfigurable impedance network, with scattering matrix $\Theta \in \mathbb{C}^{N_I \times N_I}$. We assume single-stream transmission, to exploit the diversity gain offered by the multiple antennas at the transmitter and receiver, and by the reflecting elements at the RIS. We denote the transmit signal as $x = wc \in \mathbb{C}^{N_T \times 1}$, where $w \in \mathbb{C}^{N_T \times 1}$ is the precoding vector subject to the constraint $\|w\| = 1$, and $c$ is the transmitted symbol with average power $P_T = \mathbb{E}[|c|^2]$. Denoting the receive signal as $y \in \mathbb{C}^{N_R \times 1}$, we have $y = Hx + n$, where $H \in \mathbb{C}^{N_R \times N_T}$ is the channel matrix seen by the receiver and $n \sim \mathcal{CN}(0, \sigma_n^2 I)$ is the additive white Gaussian noise (AWGN) at the receiver. Thus, the signal used for detection is given by $z = gy$, where $g \in \mathbb{C}^{1 \times N_R}$ is the combining vector subject to the constraint $\|g\| = 1$. Eventually, we can express the signal $z$ as

$$z = gHwc + \tilde{n},$$

where $\tilde{n} = gn$ is the AWGN with variance $\sigma_{\tilde{n}}^2$.

To characterize the channel matrix $H$ as a function of $\Theta$, we make the following two assumptions. Firstly, we assume that there is no impedance mismatching at the transmitter, receiver, and RIS by considering all antennas perfectly matched to the characteristic impedance, set as $Z_0 = 50 \, \Omega$. Secondly, we assume no mutual coupling between the antennas, which can be achieved in practice by setting the antenna spacing larger than half-wavelength. With these
two assumptions, the channel matrix $H$ can be written as

$$H = H_{RT} + H_{RI} \Theta H_{IT},$$

(2)

where $H_{RT} \in \mathbb{C}^{N_R \times N_T}$, $H_{RI} \in \mathbb{C}^{N_R \times N_I}$, and $H_{IT} \in \mathbb{C}^{N_I \times N_T}$ refer to the channel matrices from transmitter to receiver, RIS to receiver, and transmitter to RIS, respectively \[42\].

We denote the impedance matrix of the $N_I$-port reconfigurable impedance network as $Z_I \in \mathbb{C}^{N_I \times N_I}$. The $N_I$-port reconfigurable impedance network is constructed with passive elements which can be adapted to the channel to properly reflect the incident signal. To maximize the power reflected by the RIS, we consider $Z_I$ purely reactive and we can write $Z_I = jX_I$, where $X_I \in \mathbb{R}^{N_I \times N_I}$ denotes the reactance matrix of the $N_I$-port reconfigurable impedance network. Hence, according to network theory \[43\], $\Theta$ is given by

$$\Theta = \left( jX_I + Z_0 I \right)^{-1} \left( jX_I - Z_0 I \right).$$

(3)

Furthermore, the reconfigurable impedance network is also reciprocal so that we have $X_I = X_I^T$ and $\Theta = \Theta^T$. Depending on the connections between the RIS elements, three different RIS architectures have been identified by Shen et al., which are described in the following \[42\].

A. Single Connected Reconfigurable Impedance Network

This is the conventional architecture traditionally employed in the literature. In this architecture, each port of the RIS is connected to ground with a reconfigurable impedance and is not connected to the other ports. We denote with $Z_{ni} = jX_{ni}$ the impedance connecting the $n_I$-th port to ground, for $n_I = 1, \ldots, N_I$. Thus, the reactance matrix $X_I$ is a diagonal matrix given by

$$X_I = \text{diag} \left( X_1, X_2, \ldots, X_{N_I} \right).$$

(4)

According to \[3\], $\Theta$ is also a diagonal matrix whose $(n_I, n_I)$-th element is the reflection coefficient of the reconfigurable impedance $Z_{ni}$, and is given by $\frac{jX_{ni} - Z_0}{jX_{ni} + Z_0} = e^{j\theta_{ni}}$. Hence, the scattering matrix $\Theta$ in the single connected case can be written as

$$\Theta = \text{diag} \left( e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_{N_I}} \right).$$

(5)

B. Fully Connected Reconfigurable Impedance Network

In a more general perspective, the fully connected reconfigurable network is obtained by connecting every port of the network to all other ports. Therefore, the reactance matrix $X_I$
can be an arbitrary full matrix depending on the impedances chosen to build the reconfigurable impedance network, assumed purely reactive \cite{42}. For the reciprocity of the $N_I$-port, the only constraint on $X_I$ in the fully connected architecture is

$$X_I = X_I^T.$$  \hfill (6)

According to network theory \cite{43}, and (3), $\Theta$ is a complex symmetric unitary matrix

$$\Theta = \Theta^T, \Theta^H \Theta = I.$$  \hfill (7)

C. Group Connected Reconfigurable Impedance Network

The group connected architecture has been proposed to reach a trade-off between performance and complexity. On the one hand, the single connected architecture is the simplest one, but also the most limiting since $\Theta$ is a diagonal matrix. On the other hand, the fully connected architecture offers the highest degree of freedom but at the cost of a complex circuit topology. In the group connected architecture, the $N_I$ elements are divided into $G$ groups, each having $N_G = \frac{N_I}{G}$ elements. Each element of the $N_I$-port is connected to all other elements in its group, while there is no connection inter-group. In other words, for the $g$-th group, a $N_G$-port fully connected reconfigurable impedance network with reactance matrix of $X_{I,g} \in \mathbb{R}^{N_G \times N_G}$ is used. Thus, $X_I$ is a block diagonal matrix given by

$$X_I = \text{diag} (X_{I,1}, X_{I,2}, \ldots, X_{I,G}), \quad X_{I,g} = X_{I,g}^T, \forall g.$$  \hfill (8)

According to (3), the following constraints can be found for the scattering matrix in the group connected architecture

$$\Theta = \text{diag} (\Theta_1, \Theta_2, \ldots, \Theta_G), \quad \Theta_g = \Theta_g^T, \quad \Theta_g^H \Theta_g = I, \forall g,$$  \hfill (9)

which show that $\Theta$ is a block diagonal matrix with each block $\Theta_g$ being a complex symmetric unitary matrix, for $g = 1, \ldots, G$.

III. Discrete RIS Design

In this section, we formulate the received signal power maximization problem in a single-user RIS-aided MIMO system and we optimize group and fully connected RISs based on discrete-value impedance networks. Our goal is to design the matrix $\Theta$ and the vectors $g$ and $w$ to maximize the received signal power, which is given by $P_R = P_T |g (H_{RT} + H_{RI} \Theta H_{IT}) w|^2$. In
the case of single-stream transmission, the optimal precoding and combining vectors are given by the dominant eigenmode transmission [44]. Thus, maximizing the received signal power is equivalent to maximize \( \| H_{RT} + H_{RI} \Theta H_{IT} \| ^2 \). Depending on the considered RIS architecture, the scattering matrix \( \Theta \) is subject to different constraints, and different optimization problems are formulated.

### A. Discrete Single Connected RISs

We firstly review how the received signal power maximization problem has been traditionally solved in the well-known single connected case, where \( \Theta \) is characterized as in (5). In this case, \( \Theta \) is completely described by the phase shifts \( \theta_{n_I} \in [0, 2\pi) \), \( \forall n_I = 1, \ldots, N_I \). Assuming no restrictions on the phase shifts, the power maximization problem can be formulated as follows

\[
\max_{\theta_{n_I}} \| H_{RT} + H_{RI} \Theta H_{IT} \| ^2 \\
\text{s.t.} \quad \Theta = \text{diag} \left( e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_{N_I}} \right) , \\
\theta_{n_I} \in [0, 2\pi) , \forall n_I .
\]

(10) (11) (12)

However, continuous phase shifts are difficult to realize in practice since complex hardware is needed to tune the reconfigurable impedances with high precision.

In the traditional literature on RISs, the phase shifts have been typically discretized uniformly in the interval \([0, 2\pi)\), by assigning \( B \) resolution bits per phase shift. In other words, the phase shift of each diagonal element of \( \Theta \) is chosen in the codebook

\[
\Omega = \{ 0, \delta, \ldots, (2^B - 1) \delta \} ,
\]

(13)

where \( \delta = \frac{2\pi}{2^B} \) is the quantization step. Considering this constraint, we can reformulate the received signal power maximization problem (10)-(12) in the discrete case as

\[
\max_{\theta_{n_I}} \| H_{RT} + H_{RI} \Theta H_{IT} \| ^2 \\
\text{s.t.} \quad \Theta = \text{diag} \left( e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_{N_I}} \right) , \\
\theta_{n_I} \in \{ 0, \delta, \ldots, (2^B - 1) \delta \} , \forall n_I .
\]

(14) (15) (16)

To solve this problem, an exhaustive search could be employed since each phase shift can assume a finite number of values. However, this is prohibitive in practice because the RIS system would be required to search among \( 2^{N_I B} \) possible \( \Theta \) matrices, a number exponentially increasing with
To reduce the search space, we solve this problem through alternating optimization, also known as successive refinement, or local search, widely applied in the literature to optimize RIS with discrete phase shifts [13], [14], [22]. In this sub-optimal method, each of the $N_I$ phase shifts is optimized individually by searching among all the possible $2^B$ values, while fixing the other $N_I - 1$ phase shifts. This procedure is repeated for all $N_I$ phase shifts, iterating multiple times until the convergence is reached, i.e., until the fractional decrease of the objective value is below a certain parameter $\epsilon$. With alternating optimization, the total number of search times in a complete iteration is decreased to $N_I 2^B$.

In the following, we propose two novel design strategies for group and fully connected RISs based on discrete-value impedance networks. In the first, scalar-discrete RISs, $B$ resolution bits are allocated to each reactance matrix entry. In the second, vector-discrete RISs, $B_V$ resolution bits are allocated to each reactance matrix block $X_{I,g}$.

B. Scalar-Discrete Group and Fully Connected RISs

Let us now consider the group connected architecture. In the following, the fully connected architecture is not explicitly mentioned since it is considered a special case of group connected architecture, with $G = 1$ and $N_G = N_I$. The constraints (9) indicate that the scattering matrix $\Theta$ is a block diagonal matrix with each block being a complex symmetric unitary matrix, which complicates the optimization. Thus, the relationship between $\Theta$ and the reactance matrix $X_I$ given by (3) is exploited to write the maximization problem as

$$\max_{X_{I,g}} \|H_{RT} + H_{RI}\Theta H_{IT}\|^2$$

s.t. $\Theta = \text{diag}(\Theta_1, \Theta_2, \ldots, \Theta_G)$,

$$\Theta_g = (jX_{I,g} + Z_0I)^{-1}(jX_{I,g} - Z_0I), \ \forall g,$$  

$$X_{I,g} = X_{I,g}^T, \ \forall g,$$

which can be transformed into an unconstrained problem. More precisely, exploiting the constraints $\Theta_g$ and $\Theta_g$, the objective $\|H_{RT} + H_{RI}\Theta H_{IT}\|^2$ can be expressed as a function of $X_{I,g}, \forall g$. Since $X_{I,g}$ is an arbitrary $N_G \times N_G$ real symmetric matrix, $X_{I,g}$ is an unconstrained function of the $\frac{N_G(N_G+1)}{2}$ entries in its upper triangular part. Thus, the obtained problem is an unconstrained optimization problem in the variables $[X_{I,g}]_{i,j}$, with $i \leq j$ and $\forall g$. This problem has been solved for RIS-aided SISO systems in [42] by using the quasi-Newton method to find
the optimal upper triangular part of each $X_{I,g}$ without any constraints. Although this approach has been proved to achieve the theoretical upper bound in SISO systems, the reactance matrix entries are allowed to assume any real value in $[42]$, which is hard to realize in practice.

Since the entries of $X_{I,g}$ are not distributed in a finite interval, to define the best discrete codebook is not as straightforward as in the single connected case. For this reason, when considering the optimization of discrete group connected RISs, a two-fold problem arises: $i)$ the definition of a suitable discrete reactance codebook for the elements of $X_{I,g}$, and $ii)$ the optimization of $X_{I,g}$ over the possible values given by the codebook. To solve these two problems, we propose a RIS design based on two stages, namely, offline learning and online deployment. Firstly, the offline learning stage has the objective of building a suitable discrete reactance codebook. Secondly, during the online deployment, the discrete values of $X_I$ are optimized to maximize the received signal power.

During offline learning, the RIS uses previously acquired channel realizations to design the discrete reactance codebook. In our solution, we consider a codebook $\Psi$ symmetric around zero, which can be written as

$$\Psi = \{\pm \psi_1, \pm \psi_2, \ldots, \pm \psi_{2^{B-1}}\},$$

(21)

where $\psi_b \in \mathbb{R}, \forall b = 1, \ldots, 2^{B-1}$. To construct the codebook $\Psi$, we treat differently the two cases in which $B = 1$ and $B > 1$.

1) Case $B = 1$: When the number of resolution bits per reconfigurable reactance is $B = 1$, the entries of the reactance matrix $X_I$ can assume only the two values in $\Psi = \{-\psi_1, +\psi_1\}$, where $\psi_1$ is a positive real number. Thus, the following optimization problem can be considered

$$\max_{X_{I,g}} \|H_{RT} + H_{RI}\Theta H_{IT}\|^2$$

(22)

s.t. $\Theta = \text{diag}(\Theta_1, \Theta_2, \ldots, \Theta_G),$

(23)

$$\Theta_g = (jX_{I,g} + Z_0I)^{-1} (jX_{I,g} - Z_0I), \forall g,$$

(24)

$$X_{I,g} = X_{I,g}^T, \forall g,$$

(25)

$$[X_{I,g}]_{i,j} \in \{-\psi_1, +\psi_1\}, \forall g.$$

(26)

Similarly to the problem $[14]$, $[16]$, we can apply alternating optimization to find the reactance matrix $X_I$ which maximize the objective. However, the optimal $X_I$ is function of $\psi_1$, which
also needs to be optimized. To this end, we introduce a second optimization problem as follows

$$\max_{\psi_1} \| \mathbf{H}_{RT} + \mathbf{H}_{RI} \mathbf{\Theta}_{HT} \|^2$$

s.t. $\mathbf{\Theta} = \text{diag} (\mathbf{\Theta}_1, \mathbf{\Theta}_2, \ldots, \mathbf{\Theta}_G),$ 

$$\mathbf{\Theta}_g = (j \mathbf{X}_{I,g} + \mathbf{Z}_0 \mathbf{I})^{-1} (j \mathbf{X}_{I,g} - \mathbf{Z}_0 \mathbf{I}), \ \forall g,$$

$$\mathbf{X}_{I,g} \text{ solves (22)-(26), } \forall g.$$ 

These two nested optimization problems form a bilevel programming problem, in which (27)-(30) is the upper-level problem, and (22)-(26) is the lower-level problem. In general, the upper-level problem (27)-(30) is difficult to solve since we do not have an expression of the constraint (30), and we cannot write explicitly $\mathbf{X}_{I,g}$ as function of $\psi_1$. However, we are able to evaluate $\mathbf{X}_{I,g}$ as function of $\psi_1$ by solving (22)-(26). Thus, the upper-level problem is solvable with pattern search, an optimization algorithm that only requires the evaluation of the objective function in a series of points until the convergence is reached, without using derivatives. In a practical development, we aim to adjust the reconfigurable impedance network by choosing among the values of a fixed codebook. For this reason, we propose to optimize $\psi_1$ offline, on the basis of a training set of channel realizations, and to solve the lower-level problem (22)-(26) online, depending on the actual channels $\mathbf{H}_{RT}$, $\mathbf{H}_{RI}$, and $\mathbf{H}_{IT}$.

2) Case $B > 1$: Pattern search can be used to solve the upper-level problem (27)-(30) only when the optimization involves 1-dimensional search, that is only when $B = 1$. When more resolution bits are considered, pattern search becomes highly sub-optimal since it is likely to converge to a local maximum in a multi-dimensional search space. For this reason, we rely on another method to define the codebook $\mathbf{\Psi}$ when $B > 1$. The main idea is to learn the discrete codebook from the distribution of the optimal continuous elements of $\mathbf{X}_I$ which solve (17)-(20). This can be realized with $k$-means clustering, an unsupervised machine learning technique able to group the elements of a distribution into $k$ clusters. More precisely, since the codebook is assumed symmetric around zero, we need to obtain the $2^{B-1}$ values of the positive half of the codebook. To this end, a training set is formed with the modulus of the optimal continuous $\mathbf{X}_I$ elements. Thus, $\psi_1, \psi_2, \ldots, \psi_{2^{B-1}}$ are given by the $k$ cluster centers obtained by clustering the training set, with $k = 2^{B-1}$.

Once the codebook has been learned offline, during the online deployment stage, the discrete $\mathbf{X}_I$ entries are optimized by maximizing the received signal power. This is realized by applying
alternating optimization to solve

\[
\max_{X_{I,g}} \| \mathbf{H}_{RT} + \mathbf{H}_{RI} \Theta \mathbf{H}_{IT} \|^2
\]

\[
\text{s.t. } \Theta = \text{diag} (\Theta_1, \Theta_2, \ldots, \Theta_G),
\]

\[
\Theta_g = (jX_{I,g} + Z_0 I)^{-1} (jX_{I,g} - Z_0 I), \forall g,
\]

\[
X_{I,g} = X_{I,g}^T, \forall g,
\]

\[
[X_{I,g}]_{i,j} \in \{ \pm \psi_1, \pm \psi_2, \ldots, \pm \psi_{2^{B-1}} \}, \forall g.
\]

More precisely, each reactance matrix entry is optimized individually by searching among all possible \(2^B\) values, while fixing the other entries. This procedure is repeated for all reactance matrix entries, iterating multiple times until the convergence is reached, i.e., until the fractional decrease of the objective value is below a certain parameter \(\epsilon\). In group connected architectures, the total number of search times in a complete iteration of the alternating optimization algorithm is \(G \frac{N_G(N_G+1)}{2} 2^B\), which boils down to \(N_1(N_1+1) 2^B\) in the fully connected case. Note that the channel statistics should be assumed to be unchanged during the two stages. Thus, although the offline learning stage is conducted offline, it has to be repeated when the channel statistics changes. Since the only information needed during offline learning consists of channel samples, the training set can be obtained with no additional overhead by storing at the RIS the CSI estimated at every channel realization.

C. Vector-Discrete Group and Fully Connected RISs

In the previous subsection, the scalar codebook \(\Psi\) has been introduced to define the possible values that each element in \(X_I\) can assume. This concept can be extended by introducing a second discretization strategy based on vector quantization. To realize vector-discrete RISs, we assign \(B_V\) resolution bits to each reactance matrix block \(X_{I,g}\). Thus, we can introduce a vector codebook \(\Psi_V\) as a set of vectors

\[
\Psi_V = \{ \psi_1, \psi_2, \ldots, \psi_{2^{B_V}} \}.
\]

where \(\psi_b \in \mathbb{R}^{\frac{N_G(N_G+1)}{2}}, \forall b = 1, \ldots, 2^{B_V}\). Here, \(\psi_b\) is a possible value that each block \(X_{I,g}\) can assume. Similarly to the scalar discretization case, \(\Psi_V\) can be obtained with \(k\)-means clustering through an offline learning phase, with the difference that a multi-dimensional feature space with
dimension $\frac{N_G(N_G+1)}{2}$ is now considered, and $k = 2^{B_V}$. Thus, the online optimization problem is formulated as

$$\max_{\mathbf{X}_{I,g}} \| \mathbf{H}_{RT} + \mathbf{H}_{RI} \Theta \mathbf{H}_{IT} \|^2$$  \hspace{1cm} (37)

s.t. $\Theta = \text{diag} (\Theta_1, \Theta_2, \ldots, \Theta_G)$,

$$\Theta_g = (j \mathbf{X}_{I,g} + Z_0 \mathbf{I})^{-1} (j \mathbf{X}_{I,g} - Z_0 \mathbf{I}), \ \forall g,$$  \hspace{1cm} (38)

$$\mathbf{X}_{I,g} = \mathbf{X}_{I,g}^T, \ \forall g,$$  \hspace{1cm} (39)

$$\mathbf{X}_{I,g} \in \{ \psi_1, \psi_2, \ldots, \psi_{2^{B_V}} \}, \ \forall g.$$  \hspace{1cm} (40)

Alternating optimization is the selected strategy to solve this problem, in which the $G$ blocks $\mathbf{X}_{I,g}$ are iteratively optimized by searching among their possible $2^{B_V}$ values contained in $\Psi_V$.

The cost of this second strategy relies on its complexity. For example, if $B$ bits are allocated to each reactance element, the number of clusters needed to quantize a vector of $\frac{N_G(N_G+1)}{2}$ elements is $k = 2^{B \cdot \frac{N_G(N_G+1)}{2}}$, growing exponentially with the square of $N_G$. Consequently, the number of search times in a complete iteration of the alternating optimization algorithm becomes $G 2^{B \cdot \frac{N_G(N_G+1)}{2}}$. However, this discretization strategy based on vector quantization brings two benefits. First, for the same number of total resolution bits, vector quantization is inherently more efficient than scalar quantization. Second, the number of total resolution bits can be chosen with more degrees of freedom, since it is no longer limited to $B$ bits for each reactance element.

IV. GROUPING STRATEGY

In this section, we address the problem of grouping the RIS elements in the case of a group connected architecture. We refer to a particular partition of the $N_I$ RIS elements into $G$ groups, each containing $N_G = \frac{N_I}{G}$ elements, as a grouping strategy. The role played by the grouping strategy is still unexplored in the literature. To shed light on this, we firstly derive an upper bound on the received signal power in RIS-aided MIMO systems. Secondly, we define the characteristics of the grouping strategy which maximizes this upper bound when the direct link $\mathbf{H}_{RT}$ is negligible. Finally, we present two approaches to design the grouping strategy: One based on the channel statistics, and the other based on the channel realization.

The received signal power in an RIS-aided MIMO system using dominant eigenmode transmission is given by $P_R = P_T \| \mathbf{H}_{RT} + \mathbf{H}_{RI} \Theta \mathbf{H}_{IT} \|^2$. Considering unitary transmit power for
simplicity, we can find an upper bound for the maximum received signal power as

\[ P_R = \|H_{RT} + H_{RI}\Theta H_{IT}\|^2 \]  \hfill (42)

\[ \leq (\|H_{RT}\| + \|H_{RI}\Theta H_{IT}\|)^2 \]  \hfill (43)

\[ = \|H_{RT}\|^2 + \|H_{RI}\Theta H_{IT}\|^2 + 2\|H_{RT}\|\|H_{RI}\Theta H_{IT}\| \]  \hfill (44)

\[ \leq \|H_{RT}\|^2 + \|H_{RI}\|^2\|H_{IT}\|^2 + 2\|H_{RT}\|\|H_{RI}\|\|H_{IT}\| = \bar{P}_R, \]  \hfill (45)

where (43) follows from the triangle inequality, and (45) follows from the sub-multiplicativity of the spectral norm. To find the requirements that \( \Theta \) should satisfy in order to achieve this upper bound, in the following discussion we neglect the direct link \( H_{RT} \). This assumption reflects real scenarios where the direct channel is highly obstructed and significantly weaker than the RIS-aided link. In this case, the requirement for \( \Theta \) to achieve the upper bound is given in the following proposition.

**Proposition 1.** In the absence of the direct link \( H_{RT} \), the received signal power upper bound is achieved if and only if \( \Theta \) is a complex symmetric unitary matrix satisfying

\[ u_{RI} = \Theta u_{IT}, \]  \hfill (46)

where \( u_{RI} \) and \( u_{IT} \) are the dominant left singular vectors of \( H_{RI}^H \) and \( H_{IT} \), respectively.

**Proof.** Please refer to Appendix A.

In the case of fully connected architectures, it is always possible to find a matrix \( \Theta \) satisfying (46) with the quasi-Newton method, as shown in (42). Thus, the received signal power upper bound is tight in this case. However, when group connected architectures are considered, (46) can be satisfied only when a particular condition on \( u_{RI} \) and \( u_{IT} \) is met. More precisely, recalling that \( \Theta \) is a block diagonal matrix written as (9), (46) can be satisfied if and only if

\[ \|u_{RI,g}\| = \|u_{IT,g}\|, \forall g, \]  \hfill (47)

where \( u_{RI,g} \in \mathbb{C}^{N_G \times 1} \) and \( u_{IT,g} \in \mathbb{C}^{N_G \times 1} \) contain the \( N_G \) elements of \( u_{RI} \) and \( u_{IT} \) corresponding to the \( N_G \) RIS elements grouped into the \( g \)-th group, respectively.

Condition (47) is achievable only for particular realizations of the channels \( H_{RI} \) and \( H_{IT} \). To characterize the best grouping strategy for every channel realization, we introduce the unit vectors \( m_{RI} = [\|u_{RI,1}\|, \|u_{RI,2}\|, \ldots, \|u_{RI,G}\|] \) with \( m_{RI} \in \mathbb{R}^{1 \times G} \), and \( m_{IT} = [\|u_{IT,1}\|, \|u_{IT,2}\|, \ldots, \|u_{IT,G}\|]^T \) with \( m_{IT} \in \mathbb{R}^{G \times 1} \). Thus, fulfilling condition (47) is equivalent to making \( m_{RI} \) and \( m_{IT} \) equal.
Consequently, the quality of a specific grouping strategy can be measured with the cosine similarity between the unit vectors $\mathbf{m}_{RI}$ and $\mathbf{m}_{IT}$, defined as

$$\rho = \mathbf{m}_{RI} \mathbf{m}_{IT}. \quad (48)$$

Finally, given any realizations of $\mathbf{H}_{RI}$ and $\mathbf{H}_{IT}$, the best grouping strategy is defined as the one that maximizes this cosine similarity $\rho$. In the following, we propose two practical strategies to design the grouping strategy: The former, based on the channel statistics, maintains the grouping strategy fixed in time; the latter, based on the channel realization, adapts the grouping strategy on a per channel realization basis.

### A. Grouping Based on the Channel Statistics: Correlated Grouping and Uncorrelated Grouping

We propose two opposite families of grouping strategies based on the channel statistics, namely, correlated grouping and uncorrelated grouping. On the one hand, correlated grouping minimizes the maximum distance between antenna elements grouped together. Thus, antenna elements within the same group are likely to experience spatially correlated fading effects. On the other hand, uncorrelated grouping maximizes the minimum distance between antenna elements grouped together. In this way, antenna elements within the same group are expected to experience uncorrelated fading effects. A graphical representation of these two families of grouping strategies is shown in Fig. [2] where we consider RISs realized through uniform linear array (ULA) or uniform planar array (UPA). Here, different colors correspond to different groups.

In the case of RISs built with continuous-value impedance networks, uncorrelated grouping is expected to outperform correlated grouping. This is because uncorrelated grouping averages out the small-scale effects within each element of the vectors $\mathbf{m}_{RI}$ and $\mathbf{m}_{IT}$. Thus, the elements of $\mathbf{m}_{RI}$ tend to take on similar values to each other, as well as the elements of $\mathbf{m}_{IT}$. As a result, $\mathbf{m}_{RI}$
Algorithm 1: Grouping strategy optimization based on the channel realization.

**Input:** \(N_I, N_G, H_{RI}, H_{IT}\)

**Output:** The optimized grouping strategy \(s^*\)

1. \(i \leftarrow 0;\)
2. \(s_0\) is given by uncorrelated grouping;
3. **repeat**
   4. \(i \leftarrow i + 1;\)
   5. From \(s_{i-1}\), generate the set \(S\) containing all the \(\frac{N_I N_G (G-1)}{2}\) grouping strategies obtainable by swapping two elements;
   6. \(s_i\) is given by the best grouping strategy among the ones in \(S \cup s_{i-1}\);
4. **until** \(s_i = s_{i-1};\)
5. **return** \(s^* \leftarrow s_i;\)

and \(m_{IT}\) tend to be aligned, satisfying more easily condition (47). Conversely, when correlated grouping is employed, the elements in each group experience highly correlated fading effects. Thus, the \(g\)-th element of \(m_{RI}\) and \(m_{IT}\) depends on the fading experienced locally in the \(g\)-th group. Since all the elements of \(m_{RI}\) are likely to be different, and independent from the elements of \(m_{IT}\), condition (47) is more hardly satisfied.

**B. Grouping Based on the Channel Realization: Optimized Grouping**

Now, our goal is to find the grouping strategy that maximizes the cosine similarity \(\rho\) at every channel realization. In an RIS with \(N_I\) elements grouped into \(G\) groups, each having \(N_G = \frac{N_I}{G}\) elements, the number of possible grouping strategies is given by

\[
N_S = \frac{1}{G!} \left( \frac{N_I}{N_G} \right) \left( \frac{N_I - N_G}{N_G} \right) \cdots \left( \frac{N_G}{N_G} \right) = \frac{1}{G!} \frac{N_I!}{(N_G!)^G}. \tag{49}
\]

Due to the high value of \(N_S\), an exhaustive search among all the possible \(N_S\) grouping strategies would be too computationally expensive for medium-high values of \(N_I\). For instance, considering \(N_I = 32\), we have \(N_S > 10^8\) for every \(N_G\) non-trivial divisor of \(N_I\). Thus, to simplify this problem, we propose the iterative process shown in Algorithm 1. The initial grouping strategy, denoted as \(s_0\), is given by uncorrelated grouping. At each iteration, we generate all possible grouping strategies obtainable by swapping two elements in \(s_{i-1}\). Here, the operation of swapping
two elements consists in selecting two RIS elements belonging to two different groups and assigning each of them to the group of the other. The resulting set is denoted as $S$ and its cardinality is $\frac{N_1N_2(G-1)}{2}$. Then, $s_i$ is given by the grouping strategy which maximizes (48) among the ones in $S \cup s_{i-1}$. Finally, the convergence is considered reached when $s_i = s_{i-1}$. We refer to the grouping strategy returned by Algorithm 1 as the optimized grouping $s^*$. Note that the convergence of Algorithm 1 is guaranteed by the following two facts. First, the quality of the selected grouping strategy $s_i$ is increasing over iterations. This ensures that the same grouping strategy is never selected in multiple iterations; when this happens, the convergence is reached. Second, the total number of grouping strategies is limited.

V. NUMERICAL RESULTS

In this section, we evaluate the performance of an RIS-aided MIMO system with different discrete RIS configurations, each characterized by its number of elements $N_I$, group size $N_G$, grouping strategy, and number of resolution bits. The performance is measured in terms of received signal power, given by $P_R = P_T \|H_{RT} + H_{RI}\Theta H_{IT}\|^2$.

A. Dataset Description

Let us consider a three-dimensional coordinate system $(x, y, z)$, in which the $z$-axis represents the height above the ground in meters (m). The transmitter, the RIS, and the receiver are ULAs composed of $N_T$, $N_I$, and $N_R$ antennas, respectively, with half wavelength antenna spacing. We set $N_T = 4$ and $N_R = 2$, while considering several values of $N_I$ as specified in the following. The transmitter is located at $(5, -250, 25)$ and its antennas are arranged along the $x$-axis. The RIS is located at $(0, 0, 5)$ and its antennas are arranged along the $y$-axis. Finally, the receiver is located at $(5, 5, 1.5)$ with random orientation. The distance-dependent path loss is modeled as $L_{ij}(d_{ij}) = L_0 d_{ij}^{-\alpha_{ij}}$, where $L_0$ is the reference path loss at distance 1 m, $d_{ij}$ is the distance, and $\alpha_{ij}$ is the path loss exponent for $ij \in \{RT, RI, IT\}$. We set $L_0 = -30$ dB, $\alpha_{RT} = 4$, $\alpha_{RI} = 2.8$, $\alpha_{IT} = 2$, and $P_T = 10$ W.

We analyze two scenarios with different small-scale fading effects: i.i.d. Rayleigh fading and correlated fading. In the former scenario, the small-scale fading of all channels are assumed to be i.i.d. Rayleigh distributed. Thus, we have $H_{RT} \sim \mathcal{CN}(0, L_{RT}I)$, $H_{RI} \sim \mathcal{CN}(0, L_{RI}I)$ and $H_{IT} \sim \mathcal{CN}(0, L_{IT}I)$. In the latter scenario, we generate the small-scale effects with QuaDRiGa version 2.4, a MATLAB based statistical ray-tracing channel simulator [45]. An urban
macrocell propagation environment is simulated, with line-of-sight (LoS) in the transmitter-RIS link, and with non-line-of-sight (NLoS) in the transmitter-receiver and RIS-receiver links. The channel models “3GPP_38.901_UMa_LOS” and “3GPP_38.901_UMa_NLOS” have been used to simulate the LoS and NLoS channels, respectively. The ULAs at the transmitter, the RIS, and the receiver are modeled with the “3gpp-3d” array type and the carrier frequency is set to $f_c = 2.5$ GHz.

For both scenarios, offline learning stages have been carried out to define the optimal codebooks. In the case of scalar-discrete RISs with $B = 1$, $\psi_1$ has been obtained by averaging the outcomes of pattern search over 100 channel realizations. Otherwise, $k$-means clustering has been applied to a training dataset composed of 100 channel realizations for scalar-discrete RISs, and 500 channel realizations for vector-discrete RISs. Before applying $k$-means clustering, the optimal $X_I$ elements in the training datasets have been cleaned by removing possible outliers. We define outliers as elements more than four scaled median absolute deviation (MAD) from the median value, with the scaling factor set to $c = \left(\sqrt{2}\text{erfc}^{-1}(1/2)\right)^{-1}$, where erfc $(x)$ refers to the complementary error function. With this scaling factor, when a Gaussian distribution is cleaned, the scaled MAD approximates its standard deviation. In our alternating optimization processes, convergence is considered reached when the fractional decrease of the objective value is below $\epsilon = 10^{-3}$. Finally, the average received signal power has been computed for each RIS configuration using the Monte Carlo method.

B. Scalar-Discrete RISs

Let us now evaluate the received signal power in the case of scalar-discrete RISs. We start by analyzing the performance of the traditional single connected architecture, i.e., with group size $N_G = 1$. Fig. 3 shows the received signal power for different values of $N_I$, using from one to four resolution bits per phase shift. The performance of these scalar-discrete RISs is compared with the performance of continuous phase shift RISs and with $E[P_R]$, i.e., the average upper bound (45). We compute the received signal power of continuous phase shifts RIS by alternatively optimizing $\Theta$ and the vectors $g$ and $w$, as widely established in the literature [4], [46], [47]. After $g$ and $w$ are initialized to feasible values, the optimization process alternates between the following two steps until convergence is reached. With fixed $g$ and $w$, $\Theta$ is updated by setting $\theta_{n_I} = \text{arg}\left(gH_{RT}w - \text{arg}\left([gH_{RI}\_n_I][H_{IT}w\_n_I]\right)\right)$, $\forall n_I$. With fixed $\Theta$, $g$ and $w$ are updated as the dominant left and right singular vectors of the matrix $H_{RT} + H_{RI}\Theta H_{IT}$, respectively.
The performance of continuous phase shift RISs is approximately achieved by using four resolution bits per phase shift. Thus, a further increase in the resolution only adds more hardware complexity to the RIS, with no performance improvement. Furthermore, the upper bound (45) is well above the performance of continuous phase shift RISs, due to the limited diagonal structure of Θ in the case of single connected RISs. With correlated fading, we observe that the received signal power is higher than with i.i.d. Rayleigh fading channels. This is merely due to the dominant eigenmode transmission adopted, which benefits from highly correlated channels.

For the group and fully connected architectures, our novel approach presented in Subsection III-B is used to optimize scalar-discrete RISs. The 12 plots in Fig. 4 show the received signal power when group connected architectures (with group size $N_G \in \{2, 4, 8\}$) are used in the presence of i.i.d. Rayleigh fading channels and correlated channels. In the case of correlated channels, the three grouping strategies proposed are compared. Fig. 5 shows the received signal power achieved with fully connected architectures. In Fig. 4 and Fig. 5, the performance of scalar-discrete RISs is compared with the performance of continuous-value RISs and with the average upper bound $E[\bar{P}_R]$. We compute the received signal power of continuous-value RISs by extending the alternating optimization approach previously used for single connected continuous phase shift RISs. Now, with fixed $g$ and $w$, we update $Θ$ at each iteration by maximizing the objective $|g^H R_T w + g^H R_I Θ^H I_T w|$ with the quasi-Newton method as in [42].

Let us compare the performance of the three grouping strategies in the case of group connected RISs. As expected, uncorrelated grouping achieves a higher received signal power than correlated grouping, and optimized grouping is the best strategy among the three proposed. This holds for
Fig. 4. Average received signal power versus the number of RIS elements in group connected architectures. In the case of correlated fading, correlated grouping “CG”, uncorrelated grouping “UG”, and optimized grouping “OG” are compared.

all the group sizes $N_G$, and levels of resolution $B$ considered. Additionally, we observe that the higher the group size, the fewer resolution bits are necessary to reach the performance of continuous-value RISs. In particular, in the fully connected architecture, a single resolution bit is sufficient to obtain the same performance as the optimized continuous-value RISs, as shown in Fig. 5. The rationale behind this behavior is given by the necessary and sufficient condition (46) for $\Theta$ to achieve the upper bound. This condition consists of an underdetermined system of $N_I$ equations in $\frac{N_I(N_G+1)}{2}$ unknowns. Thus, a higher group size $N_G$ implies more degrees of freedom, and fewer resolution bits are required to satisfy it. Finally, we notice that the average upper bound $E[\bar{P}_R]$ is nearly achieved by using fully connected architectures. Further experiments revealed that this upper bound is perfectly tight in absence of the direct link $H_{RT}$.

It is worthwhile to clarify the difference between our findings and the results in [19]. In [19], the authors prove that single connected RISs with discrete phase shifts can achieve an signal-to-noise ratio (SNR) in the order of $O(N_I^2)$, regardless the number of bits $B \geq 1$ (see Proposition 1 in [19]). This means that a single resolution bit ($B = 1$) can achieve the same performance growth as $B = \infty$ in the limit $N_I \to \infty$. However, when a practical number of RIS elements
is considered, $B = 1$ cannot achieve the same performance as $B = \infty$. Furthermore, even if $B = 1$ and $B = \infty$ achieve the same asymptotic SNR growth, $B = 1$ causes a power loss of 3.9 dB \cite{22}. Conversely, in this work, we show that $B = 1$ can achieve the same exact performance as $B = \infty$ in fully connected architectures for any $N_I$, as represented in Fig. 5.

Let us now compare the performance obtained with different group sizes when the number of RIS elements $N_I$ is fixed. To this end, we optimize RISs with $N_I = 64$ elements considering all the possible group sizes, spanning from the single connected architecture ($N_G = 1$) to the fully connected ($N_G = 64$). In Fig. 6 the upper bound of the received signal power is reported together with the power maximized with different resolutions. We notice that the received signal power increases with the group size for every $B$ considered. Furthermore, the performance of discrete-value RISs with $B = 1$ approaches the performance of continuous-value RISs as the group size increases, converging to it in the fully connected architecture.

When $B$ bits are allocated to each $X_I$ element, the number of total bits employed scales with the number of $X_I$ elements as $B \frac{(N_G+1)}{2} N_I$. Fixing $N_I = 64$, we now investigate which are the RIS configurations, described by the pairs $(N_G, B)$, that maximize the received signal power when the number of total bits is limited. Fig. 7 shows how the single, group, and fully connected architectures with discrete reconfigurable impedances can provide the compromise between performance and hardware complexity. Here, each point represents the received signal power achievable with a specific RIS configuration, when $B \frac{(N_G+1)}{2} \times 64$ total resolution bits are available. Furthermore, the numeric labels above the points identify the group size $N_G$. In the case of correlated fading, the RIS configurations designed with optimized grouping achieve
Fig. 6. Average received signal power versus the group size. In the case of correlated fading, correlated grouping “CG”, uncorrelated grouping “UG”, and optimized grouping “OG” are compared.

Fig. 7. Average received signal power versus the number of total resolution bits. In the case of correlated fading, correlated grouping “CG”, uncorrelated grouping “UG”, and optimized grouping “OG” are compared.

a higher received signal power than their corresponding configurations using a grouping based on the channel statistics. Thus, we can conclude that optimized grouping is the best grouping strategy when scalar-discrete RISs are employed. Furthermore, among the two grouping strategies based on the channel statistics considered, uncorrelated grouping is the preferred one.

C. Vector-Discrete RISs

Let us analyze the performance of vector-discrete RISs, in comparison with scalar-discrete RISs. To this end, we set the number of resolution bits assigned to each block $X_{I,g}$ to $B_V = B \frac{N_G(N_G+1)}{2}$. This is equivalent, in terms of total bits employed, to assigning $B$ bits to each reactance element, allowing a comparison between scalar-discrete and vector-discrete RISs. Fig. 8 shows the performance achieved by the vector discretization strategy applied to group connected architectures with group size $N_G = 2$. Comparing this figure with the top of Fig. 4, where scalar discretization was used to optimize group connected RISs with $N_G = 2$, we can notice a significant performance improvement. In particular, in Fig. 8 $B_V = 9$ resolution bits per impedance network block are sufficient to approach the performance of continuous-value RISs. In terms of total resolution bits employed, $B_V = 9$ bits per impedance network block are equivalent.
In the case of correlated fading, correlated grouping "CG", uncorrelated grouping "UG", and optimized grouping "OG" are compared.

Finally, we compare the performance of RISs with different group sizes when the same number of total bits is employed. To this end, we fix the number of total bits used to $N_I$, which are required by the single connected architecture with $B = 1$. To use $N_I$ total bits, the feature space having $\frac{N_G(N_G+1)}{2}$ dimensions is quantized with $B_V = N_G$ resolution bits. Fig. 9 shows the received signal power achieved with different values of group size. Here, we can notice that higher group sizes achieve better performance than lower group sizes. This is due to the optimality condition (46), which has more degrees of freedom as the group size increases. Interestingly, correlated grouping achieves higher received signal power than uncorrelated grouping and optimized grouping. This is because with correlated grouping there is a higher correlation between the optimal reactance elements within each group $X_{I,g}$. Vector-quantization can benefit from this property, achieving better performance than in the uncorrelated grouping and optimized grouping cases, when a low resolution is used. In conclusion, differently from what observed for scalar-discrete RISs, correlated grouping is the preferred grouping strategy in the case of vector-discrete RISs used with a low number of total resolution bits.

D. Impact of the Grouping Strategy

Now, we investigate the impact of the grouping strategy on the received signal power in the case of RISs based on continuous-value impedance networks. To this end, we consider the correlated fading scenario. To quantify the received signal power increase in the presence of group and fully connected RISs compared to single connected RISs, we introduce the power gains of group and
fully connected RISs as $G_{\text{Group}} = \frac{P_{R_{\text{Group}}}}{P_{R_{\text{Single}}}}$ and $G_{\text{Fully}} = \frac{P_{R_{\text{Fully}}}}{P_{R_{\text{Single}}}}$, respectively. Here, $P_{R_{\text{Single}}}$, $P_{R_{\text{Group}}}$, and $P_{R_{\text{Fully}}}$ are the received signal powers in the presence of continuous-value RISs based on single, group, and fully connected impedance networks, respectively. The average power gains $G_{\text{Group}}$ (for group size $N_G \in \{2, 4\}$) and $G_{\text{Fully}}$ are shown in Fig. 10, where the three grouping strategies proposed are compared. Fig. 10 shows that, with a fixed group size $N_G$, the performance of the group connected architecture is highly dependent on the grouping strategy. As expected, we observe that uncorrelated grouping outperforms correlated grouping, and that optimized grouping is the best strategy among the three proposed. Optimized grouping with group size $N_G = 4$ nearly achieves the same performance as the fully connected architecture.

### E. Optimization Computational Complexity

To conclude our discussion, we compare RISs based on continuous-value impedance networks with RISs based on discrete-value impedance networks in terms of optimization computational complexity. On the one hand, when the quasi-Newton method is used to optimize the continuous values of $X_I$, the computational complexity of each iteration is $O\left(\frac{N_I(N_G+1)^2}{2}\right)$ [42]. On the other hand, when alternating optimization is used to optimize the discrete values of $X_I$, we distinguish two cases, namely scalar-discrete RISs and vector-discrete RISs. In the case of scalar-discrete RISs, the total number of search times in a complete iteration of the alternating optimization algorithm is $\frac{N_I(N_G+1)^2}{2}B$. Instead, in the case of vector-discrete RISs, the number of search times per iteration becomes $\frac{N_I}{N_G}2^{B_V}$. The values of these computational complexities are reported in Fig. 11. In the case of scalar-discrete RISs, we observe that our RIS design strategy based on discrete-value impedance networks is beneficial also in terms of optimization.
Fig. 10. Average power gains of the group and fully connected architectures over the single connected architectures in the case of correlated fading. Correlated grouping “CG”, uncorrelated grouping “UG”, and optimized grouping “OG” are compared.

computational complexity. Even with $B = 4$, the complexity is far less than the complexity of the quasi-Newton method for medium-high numbers of RIS elements.

VI. CONCLUSION

We propose a novel RIS design based on discrete-value reconfigurable impedance networks. The proposed design strategy extends the previous research on single connected RIS architectures with discrete phase shifts to group and fully connected RIS architectures, recently introduced in [42]. Our strategy is composed of two stages. Firstly, through the offline learning stage, we build the codebook containing the possible values for the reconfigurable impedances. Secondly, during the online deployment stage, we optimize the reconfigurable impedances with alternating optimization on the basis of the designed codebook. Two different approaches are developed, namely scalar-discrete and vector-discrete RISs, based on scalar and vector quantization, respectively. Vector-discrete RISs achieve higher performance than scalar-discrete RISs at the cost of increased optimization computational complexity. Numerical results show that a single resolution bit per reconfigurable impedance is sufficient to achieve the optimality in fully connected RISs.

We address the problem of grouping the RIS elements in the case of group connected architectures, still unexplored in the literature. We describe two approaches to design the grouping strategy: One based on the channel statistics, and the other, called optimized grouping, based on the channel realization. To optimize the grouping strategy on a per channel realization basis, we propose a practical iterative algorithm, since exhaustive search is prohibitive. By optimizing
the grouping strategy with this algorithm, RISs with group size 4 nearly achieve the same performance as fully connected RISs, with reduced hardware complexity.

Two future research avenues include the channel estimation and implementation circuits for group connected architectures. Channel estimation has been traditionally achieved for single connected RISs by exploiting the diagonal structure of the scattering matrix. Because of this property, the knowledge of the cascaded channel transmitter-RIS-receiver is sufficient to optimize the RIS, and the channel estimation problem can be highly simplified [4]. However, efficient channel estimation techniques for group connected architectures are still unexplored. A further research direction regards the implementation circuit design of group connected RISs. In this work, we assume that the RIS exhibits zero energy dissipation, as assumed in the vast majority of works on discrete phase shift single connected RISs. This allows us to optimize the discrete-value reactance matrix being agnostic about the RIS implementation circuit. However, in practical circuits, reconfigurable impedances include inevitable resistance values causing power loss, as modelled in [14], [19]. The equivalent circuit model proposed in [14] could be considered for designing group connected RISs based on discrete-value impedance networks in a future work.

APPENDIX

A. Proof of Proposition 7

In the absence of the direct link $H_{RT}$, the channel matrix between the transmitter and the receiver is given by $H = H_{RT} \Theta H_{IT}$. By applying the definition of induced matrix norm, the
received signal power upper bound can be written as

\[ P_R = \|H_{RI}\Theta H_{IT}\|^2 = \max_{\|x\|=1} \|H_{RI}\Theta H_{IT}x\|^2 \]  

(50)

\[ \leq \max_{\|x\|=1} \|H_{RI}\|^2 \|\Theta H_{IT}x\|^2 \]  

(51)

\[ \leq \max_{\|x\|=1} \|H_{RI}\|^2 \|\Theta H_{IT}\|^2 \|x\|^2 = \|H_{RI}\|^2 \|H_{IT}\|^2 = \bar{P}_R, \]  

(52)

Thus, the upper bound is achieved when both (51) and (52) are equalities. In (51), the equality holds when \(\Theta H_{IT}x\) is the dominant right singular vector of \(H_{RI}\), i.e., the dominant left singular vector of \(H_{RI}^H\). Similarly, in (52), the equality holds when \(x\) is the dominant right singular vector of the matrix \(\Theta H_{IT}\), i.e., \(\Theta H_{IT}x\) is the dominant left singular vector of the matrix \(\Theta H_{IT}\). Finally, we have that both (51) and (52) are equalities when \(\Theta H_{IT}x\) is the dominant left singular vector of both \(H_{RI}^H\) and \(\Theta H_{IT}\). To make the dominant left singular vectors of \(H_{RI}^H\) and \(\Theta H_{IT}\) equal, \(\Theta\) must satisfy \(u_{RI} = \Theta u_{IT}\), where \(u_{RI}\) and \(u_{IT}\) are the dominant left singular vectors of \(H_{RI}^H\) and \(H_{IT}\), respectively.

REFERENCES

[1] 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. 116753–116773, 2019.

[2] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” IEEE Communications Magazine, vol. 58, no. 1, pp. 106–112, 2020.

[3] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, “Reconfigurable intelligent surfaces: Principles and opportunities,” IEEE Communications Surveys Tutorials, vol. 23, no. 3, pp. 1546–1577, 2021.

[4] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, “Intelligent reflecting surface-aided wireless communications: A tutorial,” IEEE Transactions on Communications, vol. 69, no. 5, pp. 3313–3351, 2021.

[5] M.-A. Badiu and J. P. Coon, “Communication through a large reflecting surface with phase errors,” IEEE Wireless Communications Letters, vol. 9, no. 2, pp. 184–188, 2020.

[6] P. Xu, G. Chen, Z. Yang, and M. D. Renzo, “Reconfigurable intelligent surfaces-assisted communications with discrete phase shifts: How many quantization levels are required to achieve full diversity?” IEEE Wireless Communications Letters, vol. 10, no. 2, pp. 358–362, 2021.

[7] H. Zhang, B. Di, L. Song, and Z. Han, “Reconfigurable intelligent surfaces assisted communications with limited phase shifts: How many phase shifts are enough?” IEEE Transactions on Vehicular Technology, vol. 69, no. 4, pp. 4498–4502, 2020.

[8] D. Li, “Ergodic capacity of intelligent reflecting surface-assisted communication systems with phase errors,” IEEE Communications Letters, vol. 24, no. 8, pp. 1646–1650, 2020.

[9] 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.
[10] J. An, C. Xu, L. Gan, and L. Hanzo, “Low-complexity channel estimation and passive beamforming for RIS-assisted MIMO systems relying on discrete phase shifts,” *IEEE Transactions on Communications*, pp. 1–1, 2021.

[11] W. Yang, H. Li, M. Li, Y. Liu, and Q. Liu, “Channel estimation for practical IRS-assisted OFDM systems,” in *2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2021, pp. 1–6.

[12] J. Xu, W. Xu, and A. L. Swindlehurst, “Discrete phase shift design for practical large intelligent surface communication,” in *2019 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, 2019, pp. 1–5.

[13] C. You, B. Zheng, and R. Zhang, “Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming,” in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[14] 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.

[15] 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.

[16] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, “Weighted sum-rate maximization for intelligent reflecting surface enhanced wireless networks,” in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[17] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, “Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1809–1822, 2020.

[18] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, “Exploiting intelligent reflecting surfaces in NOMA networks: Joint beamforming optimization,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6884–6898, 2020.

[19] 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.

[20] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, “Intelligent reflecting surface enhanced wireless networks: Two-timescale beamforming optimization,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 2–17, 2021.

[21] ———, “Exploiting amplitude control in intelligent reflecting surface aided wireless communication with imperfect CSI,” *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 4216–4231, 2021.

[22] 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, 2020.

[23] M. Fu, Y. Zhou, Y. Shi, and K. B. Letaief, “Reconfigurable intelligent surface empowered downlink non-orthogonal multiple access,” *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3802–3817, 2021.

[24] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, “Energy efficient multi-user MISO communication using low resolution large intelligent surfaces,” in *2018 IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–6.

[25] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

[26] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, “Large intelligent surface-assisted wireless communication exploiting statistical CSI,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8238–8242, 2019.

[27] B. Zheng, Q. Wu, and R. Zhang, “Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA?” *IEEE Communications Letters*, vol. 24, no. 4, pp. 753–757, 2020.

[28] G. Yang, X. Xu, and Y.-C. Liang, “Intelligent reflecting surface assisted non-orthogonal multiple access,” in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, 2020, pp. 1–6.
[29] Z. Ding, R. Schober, and H. V. Poor, “On the impact of phase shifting designs on IRS-NOMA,” IEEE Wireless Communications Letters, vol. 9, no. 10, pp. 1596–1600, 2020.

[30] Z. Abdullah, G. Chen, S. Lambotharan, and J. A. Chambers, “A hybrid relay and intelligent reflecting surface network and its ergodic performance analysis,” IEEE Wireless Communications Letters, vol. 9, no. 10, pp. 1653–1657, 2020.

[31] X. Ma, Z. Chen, Y. Chi, W. Chen, L. Du, and Z. Li, “Channel estimation for intelligent reflecting surface enabled terahertz MIMO systems,” in 2020 IEEE International Conference on Communications Workshops (ICC Workshops), 2020, pp. 1–6.

[32] X. Ma, Z. Chen, W. Chen, Y. Chi, Z. Li, C. Han, and Q. Wen, “Intelligent reflecting surface enhanced indoor terahertz communication systems,” Nano Communication Networks, vol. 24, p. 100284, 2020.

[33] W. Chen, X. Ma, Z. Li, and N. Kuang, “Sum-rate maximization for intelligent reflecting surface based terahertz communication systems,” in 2019 IEEE/CIC International Conference on Communications Workshops in China (ICCC Workshops), 2019, pp. 153–157.

[34] Y. Lu and L. Dai, “Reconfigurable intelligent surface based hybrid precoding for THz communications,” arXiv preprint arXiv:2012.06261, 2020.

[35] W. Cai, H. Li, M. Li, and Q. Liu, “Practical modeling and beamforming for intelligent reflecting surface aided wideband systems,” IEEE Communications Letters, vol. 24, no. 7, pp. 1568–1571, 2020.

[36] Q. Wu and R. Zhang, “Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints,” IEEE Journal on Selected Areas in Communications, vol. 38, no. 8, pp. 1735–1748, 2020.

[37] Y. Zhao, B. Clerckx, and Z. Feng, “IRS-aided SWIPT: Joint waveform, active and passive beamforming design under nonlinear harvester model,” IEEE Transactions on Communications, pp. 1–1, 2021.

[38] S. Gong, Z. Yang, C. Xing, J. An, and L. Hanzo, “Beamforming optimization for intelligent reflecting surface-aided SWIPT IoT networks relying on discrete phase shifts,” IEEE Internet of Things Journal, vol. 8, no. 10, pp. 8585–8602, 2021.

[39] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. D. Renzo, C.-B. Chae, and L. Hanzo, “Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results,” IEEE Access, vol. 8, pp. 45 913–45 923, 2020.

[40] M. Dunna, C. Zhang, D. Sievenpiper, and D. Bharadia, “ScatterMIMO: Enabling virtual MIMO with smart surfaces,” in Proceedings of the 26th Annual International Conference on Mobile Computing and Networking, 2020, pp. 1–14.

[41] X. Yu, V. Jamali, D. Xu, D. W. K. Ng, and R. Schober, “Smart and reconfigurable wireless communications: From IRS modeling to algorithm design,” arXiv preprint arXiv:2103.07046, 2021.

[42] S. Shen, B. Clerckx, and R. Murch, “Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis,” IEEE Transactions on Wireless Communications, pp. 1–1, 2021.

[43] D. M. Pozar, Microwave engineering. John wiley & sons, 2011.

[44] B. Clerckx and C. Oestges, MIMO wireless networks: Channels, techniques and standards for multi-antenna, multi-user and multi-cell systems. Academic Press, 2013.

[45] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, “QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials,” IEEE Transactions on Antennas and Propagation, vol. 62, no. 6, pp. 3242–3256, 2014.

[46] 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.

[47] A. Zappone, M. Di Renzo, F. Shams, X. Qian, and M. Debbah, “Overhead-aware design of reconfigurable intelligent surfaces in smart radio environments,” IEEE Transactions on Wireless Communications, vol. 20, no. 1, pp. 126–141, 2021.