Dynamic Metasurface Grouping for IRS Optimization in Massive MIMO Communications

Christian James Daniel

Submitted to the graduate degree program in Department of Electrical Engineering and Computer Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science.

Erik Perrins, Chairperson

Committee members

Taejoon Kim, Co-Chair

Morteza Hashemi

Date defended: Jan 6, 2022

The Thesis Committee for Christian James Daniel certifies
that this is the approved version of the following thesis :

Dynamic Metasurface Grouping for IRS Optimization in Massive MIMO Communications

Erik Perrins, Chairperson

Date approved: Jan 6, 2022

Abstract

Intelligent Reflecting Surfaces (IRSs) grant the ability to control what was once considered the uncontrollable part of wireless communications, the channel. These smart signal mirrors show promise to significantly improve the effective signal-to-noise-ratio (SNR) of cell-users when the line-of-sight (LOS) channel between the base station (BS) and user is blocked. IRSs use implementable optimized phase shifts that beamform a reflected signal around channel blockages, and because they are passive devices, they have the benefit of having low cost and low power consumption. Previous works have concluded that IRSs need several hundred elements to outperform relays [1]. Unfortunately, overhead and complexity costs related to optimizing these devices limit their scope to single-input single-output (SISO) systems [2]. With multiple-input multiple-output (MIMO) and Massive MIMO becoming crucial components to modern 5G and beyond networks [3], a way to mitigate these overhead costs and integrate IRS technology with the promising MIMO techniques is paramount for these devices to have a place within modern cell technologies. This thesis proposes an IRS element grouping scheme that greatly reduces the number of unique IRS phases that need to be calculated and sent to the IRS controller via the limited rate feedback channel and allows for the ideal number of groups to be obtained at the BS before data transmission. Three methods are proposed to design the phase shifts and element partitioning within our scheme to improve effective SNR in an IRS-aided system. In our simulations, it is shown that our best performing method is one that dynamically partitions the IRS elements into non-uniform groups based on information gathered from the reflected channel and then optimizes its phase shifts. This method successfully handles the overhead trade-off problem, and shows significant achievable rate improvement from previous works.

Acknowledgements

I would like to thank my wife Noelle. We share in each other's successes and I could not have finished this thesis without your support. That being said, my gift to you is that I won't make you read this long document. Thank you to my committee members and especially my research advisors Dr. Perrins and Dr. Kim for the advice and long hours required to make this thesis a reality. Thank you both for guiding me on the path to create research that I am proud of presenting.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Whether referred to as an intelligent reflecting surface (IRS) or one of its many other names: reconfigurable intelligent surface, software-controlled metasurface, large intelligent surface, start reflect-arrays, etc., IRSs are an exciting new technology consisting of a planar metasurface made up of a massive number of individual elements, also known as meta-atoms, that work together to effectively improve the wireless propagation environment [5]. Each of these meta-atoms has the ability to induce a phase shift that focuses and directs the reflected signal within the wireless environment to the intended user. IRS technology is now being more widely researched as recent developments in radio frequency micro electromechanical systems and metamaterials have led to implementable IRS systems that are capable of reconfiguring the phase shifts in real time [6]. IRS devices are now a feasible, yet unrefined, way to improve data rate performance, especially in systems affected by a line-of-sight (LOS) blockage. The leading benefit of IRS-aided systems over other techniques that improve the channel for better data rate performance, namely a relay, is that IRSs do not consume transmit power [1]. Instead, an IRS passively reflects an incident signal by adjusting each IRS element's capacitance and resistance through the use of a controller in response to the estimated wireless channel to jointly control the reflected signal's amplitude and phase shift [7]. By leveraging a massive number of these low-cost passive elements, an IRS can achieve significant signal-to-noise ratio (SNR) beamforming gain by constructively combining the reflected signals at the receiver. However, this gain comes at a cost. Namely, an increase in complexity overhead

due to the channel estimation and feedback time required to optimize many phase shifts. The time needed for overhead reduces the time left for data transmission which directly affects the maximum achievable data rate [2]. This creates a fundamental trade-off between creating effective SNR gain through the implementation of many optimized phase shifts and mitigating the overhead cost required to handle this implementation.

As the demand for higher data rates continues to surge as we approach 5G and beyond networks, multiple-input multiple-output (MIMO) and massive MIMO technology are at the forefront of current research and development. With the added complexity overhead of multiple signal paths from transmit antennas and receivers, solving the overhead trade-off problem is crucial to the effectiveness of IRS technology [8]. In this thesis we target this trade-off problem, and propose methods to reduce complexity while maintaining a high effective SNR to pave the way for massive MIMO applications.

1.2 Related Work

Many previous works have drawn the comparison between massive MIMO and IRS technology as massive MIMO takes advantage of many antennas to achieve spectral and energy efficiency performance, and IRSs utilize many reflecting elements to do the same [9]. Unfortunately, the SNR power-scaling properties of IRSs in practice are not as impactful as those in massive MIMO [10]. However, for 5G and beyond, IRSs will be critical additions towards the next generation of massive MIMO 2.0 [8], where IRSs populate the network environment to provide reliable assistance to the massive MIMO base stations. These IRS devices can be implemented on walls, billboards, building facades, etc. so that there is always an IRS within close proximity of the user [11], [12].

Although a few works propose IRSs with some active elements (connected to the baseband) [13], [14], many consider a completely passive surface [15], [16], [17] to be the most practical for energy efficiency and ease of implementation. With a passive surface comes issues with channel estimation and IRS reflection coefficient optimization without signal pro-

cessing at the IRS. Some works consider perfect channel state information (CSI) [18], [19] or statistical CSI [11] to simplify this problem. However, more realistic channel estimation methods have been proposed in works such as [17] and [20] that do not rely on CSI or active elements within an IRS.

After channel estimation, the base station has the ability to design the phase shifts for optimized reflection. One strategy used is alternating optimization between the base station transmit power allocation and IRS phase shift design as in [18] and [21]. Another method for optimization is designing the reflection coefficients via deep learning developed using channel information gathered over time [15], [16].

Complexity and overhead are the associated costs with these optimization methods, but many IRS works do not consider the practical cost of overhead when optimizing their data rate maximization problems. In [22], the hybrid beamforming performance is investigated in a cell network system model, but the associated overhead costs from MIMO techniques as opposed to single antenna transmission are not considered. In the experimental research proposed in [23], a prototype IRS is tested in a controlled LOS scenario and the effective SNR gain is modeled. While this work shows potential in IRS techniques, the cell-network implementation requires complexity considerations that this work does not investigate. The work in [24] puts into context the importance of overhead in IRS-aided systems, as there is an inherent complexity trade-off when increasing the number of transmit or receive antennas. More antennas will result in effective SNR gain while also leading to larger channel estimation and feedback costs.

The overhead trade-off problem for maximizing the achievable rate was formally proposed in [2] and the allocation of power and bandwidth resources problem was formulated within this framework. In this work it was shown that any rate improvements made with an increase of transmit and receive antennas or IRS elements are quickly offset by the reduction in data transmission time due to the extended length of time needed for channel estimation and feedback transmission. From this observation, it was postulated that the optimal operating

regime of an IRS system in terms of maximizing achievable data rate contains only one transmit and one receive antenna to minimize the number of transmitter-to-IRS and IRS-to-receiver channels to be estimated, and therefore the overhead. However, this is motivated by the overwhelming cost of complexity per IRS element. Now, if there was a way to reduce the cost of overhead per IRS element then there is more room for increased rate performance by utilizing a larger number of transmit antennas. In an attempt to mitigate this issue, several works including [25] and [26] attempt to model the optimized phase shifts as discrete and thus save on feedback overhead when transmitting the bits to the IRS controller.

Since the overhead is directly proportional to the number of channels to be estimated and the length of feedback transmission [2], another attempt to mitigate overhead introduced in [4] is to group nearby elements into rectangular sub-surfaces and only implement one phase shift per group. This reduces channel estimation by assuming one channel represents every element in a group and reduces feedback transmission by the number of phase shifts sent to the IRS controller. This scheme works to reduce the total overhead proportionally to the size of a group. In order to implement this effectively, the assumption is made that nearby elements have highly correlated channels, which allows each group of elements to share one phase shift. By optimizing the phase shift for each group, it was shown in [4] that this method achieves desirable achievable rate performance by significantly reducing overhead. The grouping method proposed in [4] has been built upon by other works such as [27] and [25] as an important way to reduce overhead. Along with this grouping method, [4] also presents the conclusion that for any IRS element grouping scheme, there exists an optimal grouping ratio (number of groups to number of IRS elements) that achieves the best trade-off between overhead reduction and effective SNR.

The previously proposed method for grouping shows the existence of an ideal grouping ratio, however, this ideal grouping is an averaged trend over many channel realizations. They utilize this trend by selecting a rule-of-thumb 2×2 size sub-surface to use for all simulations. While this choice of grouping increases achievable rate performance, the grouping chosen

is fixed and must be estimated before transmission. In fact, their method would require CSI before channel estimation to determine the ideal grouping at each channel realization. This requirement invalidates the benefit of the method in the first place, reduced channel estimation overhead. Therefore, implementation of the ideal grouping at every channel realization in their scheme is impractical. Because of the existence of this ideal grouping, there is currently untapped potential in IRS grouping schemes. This thesis will target this untapped potential by formulating a novel grouping scheme.

1.3 Overview of Methodology and Contributions

The methodologies and solutions of this paper are summarized as follows:

First, we formulate a narrowband transmission model for an IRS-aided system. Then we formulate an optimization problem to maximize effective SNR by optimizing IRS reflection coefficients. We then propose a scheme that utilizes IRS element sub-surface grouping that can achieve the ideal number of groups for each channel matrix. We leverage this knowledge to achieve the peak trade-off of achievable rate performance and overhead cost. This grouping scheme requires optimized phase shifts to maximize the effective SNR problem and reach our formulated upper bound. Therefore we develop three unique methods that each formulate a grouping strategy, and within the grouping framework, design the IRS reflection coefficients to improve effective SNR to target this upper bound.

The first two methods we develop for IRS reflection coefficient design build on our proposed scheme and take advantage of the combined IRS channel to improve effective SNR. One is explicitly solved while the other is a product of an alternating maximization algorithm. Next, we develop the third method that consists of an algorithm designed to sub-optimally allocate all IRS elements into groups efficiently and dynamically in order to approach the upper bound for effective SNR performance considering grouping. By relocating the elements fluidly from one group to another according to our formulated metric, we are able to greatly improve upon phase shift coherence within one group. In addition, we create another

algorithm that initializes the dynamic grouping algorithm, helps its convergence behavior, and shrinks the gap between the achieved effective SNR and its upper bound.

The following are the contributions for this thesis:

- Where the IRS element grouping technique introduced in [4] requires structured, rule-of-thumb grouping, we propose an alternate scheme that can take advantage of the ideal number of groups for each channel matrix, as well as designing the grouping for each individual IRS element in a way that improves effective SNR. Simulations show the performance of our methods greatly outperforms previous methods in terms of effective SNR and maximum achievable rate.

- Previous grouping methods [4] (other references) still have a theoretical ideal grouping ratio (IGR) of groups divided by IRS elements even though they cannot attain it. Our simulations show the convergence of the IGR and the ideal number of groups as the number of IRS elements increases for our methods as well as the reference method in [4].

- In previous works such as [2], it is asserted that overhead is an overwhelming cost for IRS implementation and thus IRSs are limited to SISO applications. However, our methods use variable-sized grouping to handle the trade-off between overhead and effective SNR gain with a focus on improving the achievable rate. Our simulations show that the proposed dynamic grouping method holds great potential to largely mitigate the high overhead associated with implementing IRS technology in massive MIMO systems.

1.4 Organization

The remaining chapters of this thesis are organized as follows. In Chapter 2 we establish the system and channel models. Chapter 3 introduces the concept of grouping as well as explaining what makes our proposed method unique. Chapter 3 provides the beginning of

our problem formulation for this thesis, as well as introducing the concept of IRS element grouping by discussing the scheme presented in a previous work. We then propose our own grouping scheme and explain why it is superior. In Chapter 4 we introduce the first of our three grouping methods, uniform phase grouping. In Chapter 5, the second grouping method, double phase grouping, is formulated and an algorithm is designed to improve effective SNR performance. In Chapter 6 our third and superior method, dynamic grouping, is proposed and our complex optimization problem is sub-optimally solved using a combination of two algorithms. Chapter 7 is our simulation chapter, with results and analysis. Finally, Chapter 8 is the location of our conclusions and proposals for future work.

1.5 Notations

The following notation is used in this thesis: A bold upper case $\mathbf{A}$ represents a matrix, a bold lower case $\mathbf{a}$ represents a vector, and a lower case a represents a scalar. If $a_{m,n}$ is an element of matrix $\mathbf{A}$, then the subscripts m and n designate the m-th row and the n-th column of $\mathbf{A}$, respectively. If b_i is an element of the vector $\mathbf{b}$ then the subscript i indicates that b_i is the i-th element of $\mathbf{b}$. The matrix $\mathbf{I}_M$ is the $M \times M$ identity matrix, the vector $\mathbf{e}_m$ is the m-th column of $\mathbf{I}_M$, and $\mathbf{1}_z, \mathbf{0}_Z$ are the all-ones column vector with length Z and the all-zeros column vector with length Z, respectively. In this work we use $\|\mathbf{A}\|_1$, $\|\mathbf{A}\|_2$, $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^*$, $\mathbf{A}^{\mathrm{H}}$, and $\lambda_1(\mathbf{A})$ to describe the 1-norm, 2-norm, transpose, conjugate, Hermitian (conjugate transpose), and the largest eigenvalue of $\mathbf{A}$ respectively. The superscript $\mathbf{A}^\star$ is saved for a solution to an optimization problem. We also use the $\otimes$ and $\odot$ which describe the Kronecker Product and Khatri-Rao product respectively. For the definition of Kronecker Product see Appendix A and for the definition of the Khatri-Rao product, see Appendix B. $\mathcal{CN}(0,1)$ is complex Gaussian noise with mean 0 and variance 1. The operator diag $(\mathbf{a})$ for a vector $\mathbf{a}$ of length M is defined as $[\mathbf{e}_1 \otimes a_1, \mathbf{e}_2 \otimes a_2, \cdots, \mathbf{e}_M \otimes a_M]$, the operator rank$(\mathbf{A})$ is the rank of matrix $\mathbf{A}$, and the operator divisors(a) represents the set of all factors $x_i \in \chi, \forall i$ where $\frac{a}{x_i}$ is always an integer. We also use $\angle\mathbf{a}$ to find the angle in radians (from $-\pi$ to π)

of each element of vector $\mathbf{a}$. The operator $\bigcup$ denotes the union of a set, and $\setminus$ denotes the set exclusion operator.

Chapter 2

System Model and Assumptions

Abstract

In this chapter, we introduce the Narrowband IRS System Model for this work, as well as defining the channel model and the pilot tone channel estimation method that is utilized.

2.1 System Model

We consider the single-user multiple-input single-output (MISO) downlink communication system model depicted in Fig 2.1 in a wireless cellular environment. A base station (BS) with N antennas transmits a single stream signal to a single antenna user where the direct link is completely blocked. Therefore, there is no line-of-sight (LOS) channel path and the signal only arrives at the receiver antenna after reflecting off of objects in the environment. An IRS with M elements is implemented within close proximity of the user to create a more favorable reflected channel [1]. Let $\mathbf{H} = [\mathbf{h}_1,\ \mathbf{h}_2, \cdots, \mathbf{h}_M]^{\mathrm{T}} \in \mathbb{C}^{M \times N}$ define the channel from the BS and the IRS, where $\mathbf{h}_m = [h_{m,1},\ h_{m,2}, \cdots, h_{m,N}]^{\mathrm{T}} \in \mathbb{C}^{N \times 1}$ contains the channel coefficients between the N transmit antennas to the m-th IRS element and let $\mathbf{g} = [g_1,\ g_2, \cdots, g_M]^{\mathrm{T}} \in \mathbb{C}^{M \times 1}$ define the reflected channel between the IRS and the user, where each g_m is the channel coefficient from the m-th IRS element to the receiver.

The IRS assists in the communication by inducing a phase shift at each of its M elements, which are individually implemented by the designated IRS controller. These phase phase shifts are designed at the BS and fed back to the IRS controller via the separate $h_F \in \mathbb{C}$

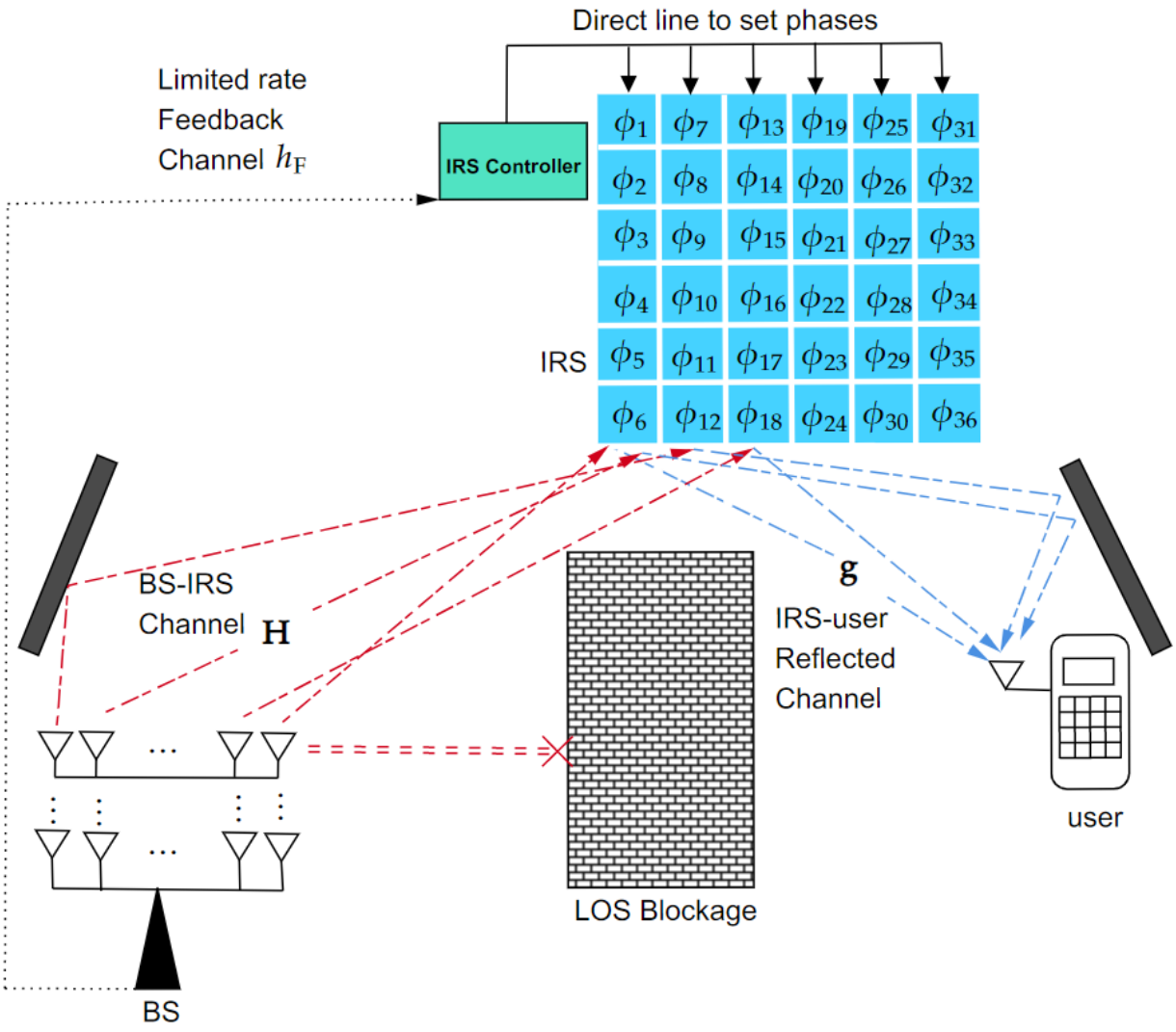

Figure 2.1: System model for downlink IRS-reflected channel

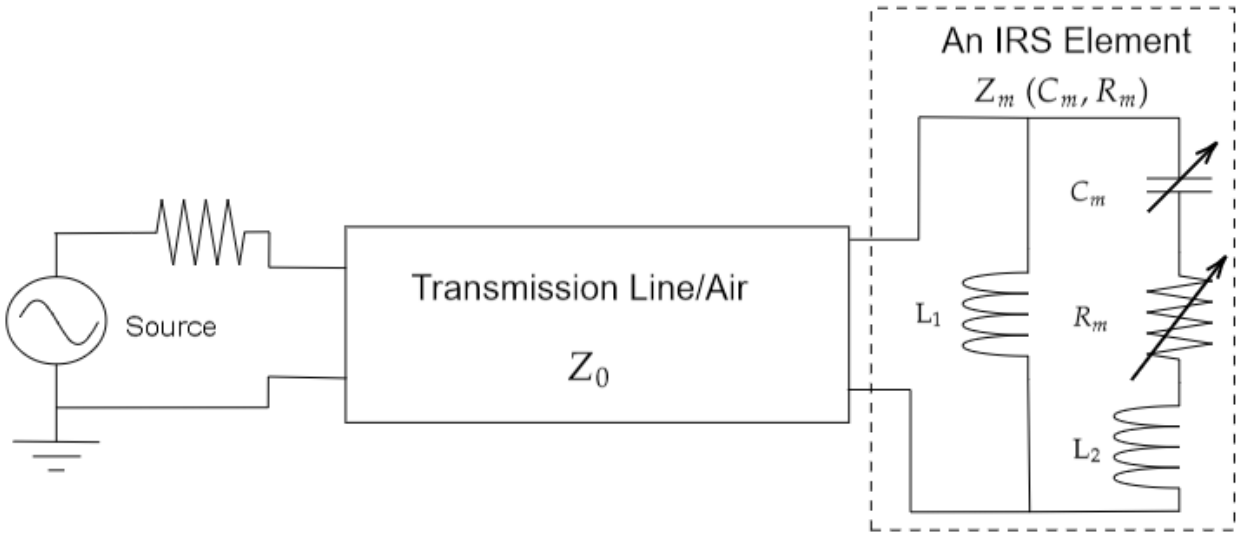

Figure 2.2: Transmission line equivalent circuit model of a single IRS element.

scalar limited-rate feedback channel. Because the IRS is completely passive, unlike a relay, it has no active receiving or transmitting capabilities and relies on the b_{F} feedback bits to implement the phase shift value applied at each element, M in total. These phase shifts direct the reflected signal paths towards the receiver and create a focused beampattern [1]. Essentially, an IRS can be imagined as a sort of analog beamformer in the middle of a channel, one that beamforms the signal but leaves the power unchanged.

The elements of an IRS are equally and tightly spaced on a 2-dimensional plane and constructed using a printed circuit board (PCB). Each element is composed of a metal patch on the top layer of the PCB dielectric substrate and a full metal sheet on the bottom layer [28]. In addition, a semiconductor device, which varies the impedance of the IRS element by controlling its biasing voltage, is embedded into the top layer metal patch so that the element response can be dynamically tuned in real time without needing to alter its geometric parameters [28]. The equivalent circuit model for each IRS element m is shown in Fig. 2.2.

The impedance of each element is varied by the semiconductor device according to the following equation:

$$Z_m\left(C_m, R_m\right) = \frac{j\omega L_1\left(j\omega L_2 + \frac{1}{j\omega C_m} + R_m\right)}{j\omega L_1 + \left(j\omega L_2 + \frac{1}{j\omega C_m} + R_m\right)}, \tag{2.1}$$

where $L_1, L_2, C_m, R_m,$ and ω denote the bottom layer inductance, top layer inductance, effective capacitance, effective resistance, and angular frequency of the incident signal, respectively. The reflection coefficient of each IRS element ϕ_m is a fraction of the impedance discontinuity between the element impedance in (2.1) and the free space impedance Z_0 as given in

$$\phi_m = \frac{Z_m\left(C_m, R_m\right) - Z_0}{Z_m\left(C_m, R_m\right) + Z_0}.$$

To simplify calculations in this thesis, instead of representing the reflection coefficient as a function of capacitance and resistance, we simply consider it as a function of an amplitude $\beta_m \in [0, 1]$ and a phase shift $\theta_m \in [0, 2\pi)$ of the form $\phi_m = \beta_m e^{j\theta_m}$, where $|\phi_m| \leq 1,\ \forall m$.

In practice, the phase shift θ_m is sent through a feedback channel from the transmitter to the IRS, and is therefore limited in its precision by the number of bits dedicated to its transmission. The value of the phase shift is chosen from 2^{b_F} different phase values in $[0, 2\pi)$, where b_F is the number of bits designated for phase shift selection. Therefore $\theta_m \in \{0, \Delta\theta, \ldots, (2^{b_\mathrm{F}} - 1)\Delta\theta\}$, where $\Delta\theta = 2\pi/2^{b_\mathrm{F}}$. As in previous works, we assume continuous values of θ_m for our optimization and use a large enough value of b_F to therefore assume that the negative effects of discretization are negligible [28]. For example, with $b_\mathrm{F} = 16$ the set of discreet phase shifts available for selection would contain $2^{16} = 65536$ possible values, for a precision of $\Delta\theta = 3\pi \times 10^{-5}$. We also assume that $\beta_m \in \{0, 1\}$ which is to say that the IRS has the ability to either fully reflect or fully absorb, as those are the only two states needed in this work. We use $\beta_m = 1$ for data transmission by assuming no attenuation due to the resistance of IRS elements and $\beta_m = 0$ to aid in ON/OFF channel estimation techniques [4].

2.2 Transmission Model

The input-output channel model of this system can be given by

$$y = s\mathbf{q}^{\mathrm{H}}\mathbf{v}_{\phi} + w, \tag{2.2}$$

where s is the transmit symbol satisfying the power constraint $E\left[|s|^2\right] \leq p$, $\mathbf{q} \in \mathbb{C}^{N\times 1}$ is our unit-norm ($\|\mathbf{q}\|_2 = 1$) transmit beamforming vector and $w \sim \mathcal{CN}(0,1)$. Here we use the vector $\mathbf{v}_{\phi}$ to represent the effective channel from the BS to the user. We can then say that the effective combined scalar channel from the n-th transmit antenna to the user is given by

$$v_n = \sum_{m=1}^{M} h_{mn} g_m \phi_m = \left[h_{1n}g_1, h_{2n}g_2, \cdots, h_{Mn}g_M\right]\boldsymbol{\phi},$$

where $\boldsymbol{\phi} = \left[\phi_1 \cdots \phi_M\right]^T \in \mathbb{C}^{M\times 1}$. Collecting all of $v_n, 1 \leq n \leq N$ into a vector $\mathbf{v}_{\phi} \in \mathbb{C}^{N\times 1}$ gives

$$\begin{aligned}
\mathbf{v}_{\phi} &= \left[\mathbf{h}_1 \otimes g_1, \mathbf{h}_2 \otimes g_2, \cdots, \mathbf{h}_M \otimes g_M\right]\boldsymbol{\phi} && (2.3)\\
&= \left(\mathbf{H}^T \odot \mathbf{g}^T\right)\boldsymbol{\phi} && (2.4)\\
&= \mathrm{vec}\left(\mathbf{g}^T\boldsymbol{\Phi}\mathbf{H}\right), && (2.5)
\end{aligned}$$

where $\boldsymbol{\Phi} = \mathrm{diag}(\boldsymbol{\phi})$ and the equality in (2.5) follows from the identity $\mathrm{vec}(\mathbf{A}\,\mathrm{diag}\,(\mathbf{d})\,\mathbf{B}) = \left(\mathbf{B}^T \odot \mathbf{A}\right)\mathbf{d}$. Focusing on the composite model in (2.4), one can rewrite the combined channel vector as

$$\mathbf{v}_{\phi} = \mathbf{V}_r\boldsymbol{\phi},$$

where we say the $\mathbf{H}$ and $\mathbf{g}$ combined channel is defined as $\mathbf{V}_{\mathrm{r}} = \left[\mathbf{v}_{\mathrm{r},1}, \mathbf{v}_{\mathrm{r},2}, \cdots, \mathbf{v}_{\mathrm{r},M}\right] = \left(\mathbf{H}^T \odot \mathbf{g}^T\right) \in \mathbb{C}^{N\times M}$ and $\mathbf{v}_{\mathrm{r},m} = \mathbf{h}_m \otimes g_m$. The effective channel input-output model is thus

given by

$$y = s\mathbf{q}^{\mathrm{H}}\mathbf{V}_r\boldsymbol{\phi} + w. \tag{2.6}$$

2.3 Pilot Tone Transmission and Channel Estimation

The transmission model in (2.6) gives context to how the IRS phase shifts are applied to the transmitted signal, and it also allows for a channel estimation technique to be created for our system model, which we derive in this chapter.

We use an IRS element on/off ($\phi_{\mathrm{ON}} = 1$ or $\phi_{\mathrm{OFF}} = 0$) scheme [4] to determine each channel vector $\mathbf{v}_{r,m}$ of $\mathbf{V}_{\mathrm{r}}$. By turning on only the m-th meta-atom and setting its reflection coefficient to 1 (i.e. $\boldsymbol{\phi} = \mathbf{e}_m$) the received signal is of the form

$$y = \mathbf{f}^H\mathbf{v}_{r,m} + w,$$

where $\mathbf{f} \in \mathbb{C}^{N\times 1}$ is a pilot vector. To estimate $\mathbf{v}_{r,m} \in \mathbb{C}^{N\times 1}$ with this technique, it follows that the transmitter sends a sequence of pilot vectors $\mathbf{F} = [\mathbf{f}_1 \cdots \mathbf{f}_T] \in \mathbb{C}^{N\times T}$ that is known to the receiver such that $\mathbf{F}\mathbf{F}^H = \mathbf{I}_T$. This orthogonality principle helps to simplify our solution. Then, the sounded signal after T channel uses is

$$\mathbf{y}_m = \mathbf{F}^H\mathbf{v}_{r,m} + \mathbf{w}_m.$$

We use the least squares (LS) channel estimation method as a practical case for our system model. The LS estimate for $\mathbf{v}_{r,m}$ is given by

$$\hat{\mathbf{v}}_{r,m} = \left(\mathbf{F}\mathbf{F}^H\right)^{-1}\mathbf{F}\mathbf{y}_m = \mathbf{F}\mathbf{y}_m,$$

where the last equality is due to the fact that $\mathbf{F}\mathbf{F}^H = \mathbf{I}_T$. Now, repeating the channel

sounding for each column of $\mathbf{V}_r$, its composite received signal is equivalently written as

$$\mathbf{Y} = \mathbf{F}^H \mathbf{V}_r + \mathbf{W},$$

where $\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_M] \in \mathbb{C}^{T \times M}$ and $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_M] \in \mathbb{C}^{T \times M}$. Thus, its LS estimate is

$$\hat{\mathbf{V}}_r = \mathbf{F}\mathbf{Y}.$$

To produce an accurate estimate, this method requires at least N pilot tones ($N \leq T$) to estimate the channel, and that the time needed is proportional to the number of IRS elements in

$$T_0 \left(NM + 1 \right), \tag{2.7}$$

where T_0 is the time required to send one pilot tone, and the extra " $+ 1$" is from the pilot tone needed to estimate the BS to IRS feedback channel shown in Fig. 2.1. In this system model, we also set aside a designated frequency band for robust and accurate feedback communication.

2.4 Channel Model

We follow a geometric channel model in this thesis. We assume the case where the IRS has a direct LOS path to the BS and to the user, because this has been observed to be the best scenario for IRS aided communication to improve performance [23]. We consider a quasi-static block fading channel model so that within one fading block the channels stay approximately constant. The channel parameters were adapted from [4] and can be found in Chapter 7.1. It is also important to note that the transmit antenna array elements are evenly spaced half of a wavelength apart in two dimensions to take the form of a standard uniform rectangular array (URA). In some previous works [29], the IRS was structured similar to an

antenna array, where each element is spaced half of a wavelength apart $\left(d_{\mathrm{BS}} = \frac{\lambda}{2}\right)$, however, in practical IRSs the elements are packed much closer together [4]. In this thesis, we define $d_{\mathrm{IRS}} = \frac{\lambda}{8}$.

The appropriate geometric channel between the BS and the IRS can be represented as

$$\mathbf{H} = \mathbf{A}_{\mathrm{R}}\mathbf{D}\mathbf{A}_{\mathrm{T}}^{\mathrm{H}}, \tag{2.8}$$

where the subscripts R and T refer to the receiver-end of the channel and transmit-end of the channel, respectively. The matrices $\mathbf{A}_{\mathrm{R}}$ and $\mathbf{A}_{\mathrm{T}}$ are defined as

$$\mathbf{A}_{\mathrm{R}} = \left[\mathbf{a}_{\mathrm{R}}\left(\varphi_{\mathrm{R},1}, \psi_{\mathrm{R},1}\right), \ldots, \mathbf{a}_{\mathrm{R}}\left(\varphi_{\mathrm{R},N_{\mathrm{ch}}}, \psi_{\mathrm{R},N_{\mathrm{ch}}}\right)\right] \in \mathbb{C}^{M \times N_{\mathrm{ch}}}, \tag{2.9}$$

$$\mathbf{A}_{\mathrm{T}} = \left[\mathbf{a}_{\mathrm{T}}\left(\varphi_{\mathrm{T},1}, \psi_{\mathrm{T},1}\right), \ldots, \mathbf{a}_{\mathrm{T}}\left(\varphi_{\mathrm{T},N_{\mathrm{ch}}}, \psi_{\mathrm{T},N_{\mathrm{ch}}}\right)\right] \in \mathbb{C}^{N \times N_{\mathrm{ch}}}, \tag{2.10}$$

where φ and ψ are the elevation and azimuth angles, respectively and N_{ch} is the number of channel paths. The subscripts of φ and ψ, R and T, describe the angles of arrival and departure, respectively. The array response of an array element consists of an incremental phase shift (in the frequency domain) caused by a shift in the spatial domain defined by the reference plane in Fig. 2.3. This variable phase shift is of the form $e^{j\theta_w}$, where θ_w is the phase offset output from the function $w\left(x, y, \varphi, \psi\right)$. This function is designed based on a URA aligned in the y-z plane as seen in Fig. 2.3 and changes in accordance with the selected array element and the associated angle pair [4]. The receive-end and transmit-end array response vectors are

$$\mathbf{a}_{\mathrm{R}}\left(\varphi, \psi\right) = \left[\mathrm{e}^{jw_{\mathrm{R}}(1,1,\varphi,\psi)}, \ldots, \mathrm{e}^{jw_{\mathrm{R}}(m_x,m_y,\varphi,\psi)}, \ldots, \mathrm{e}^{jw(M_x,M_y,\varphi,\psi)}\right]^{\mathrm{T}} \in \mathbb{C}^{M \times 1}, \tag{2.11}$$

$$\mathbf{a}_{\mathrm{T}}\left(\varphi, \psi\right) = \left[\mathrm{e}^{jw_{\mathrm{T}}(1,1,\varphi,\psi)}, \ldots, \mathrm{e}^{jw_{\mathrm{T}}(n_x,n_y,\varphi,\psi)}, \ldots, \mathrm{e}^{jw_{\mathrm{T}}(N_x,N_y,\varphi,\psi)}\right]^{\mathrm{T}} \in \mathbb{C}^{N \times 1}, \tag{2.12}$$

where M_x and M_y are the width and height dimensions of the IRS URA ($M = M_x \times M_y$) and N_x and N_y are the width and height dimensions of the URA at the BS ($N = N_x \times N_y$).

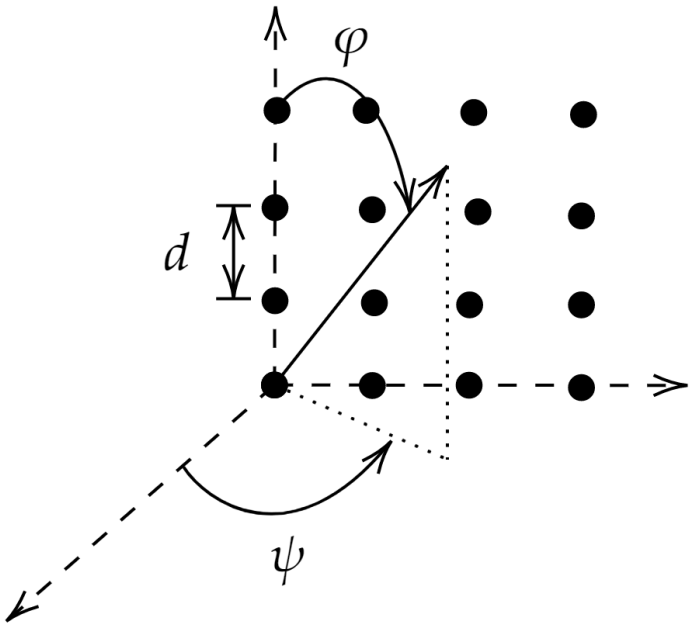

Figure 2.3: The URA response vector is defined for arrays oriented in the y-z plane, which is illustrated here. The distance between elements d, the elevation angle $\varphi \in [0, \pi]$, and the azimuth angle $\psi \in [-\pi, \pi]$ are also shown.

The value of the phase offset function w is calculated with one of the following equations:

$$w_{\mathrm{R}}(x, y, \varphi, \psi) = \frac{2\pi}{\lambda} d_{\mathrm{IRS}} \left((x-1) \sin\varphi \sin\psi + (y-1) \cos\varphi \right), \tag{2.13}$$

$$w_{\mathrm{T}}(x, y, \varphi, \psi) = \frac{2\pi}{\lambda} d_{\mathrm{BS}} \left((x-1) \sin\varphi \sin\psi + (y-1) \cos\varphi \right). \tag{2.14}$$

These array response values are based on the elevation and azimuth angle of the channel path and also the number of spatial shifts by separation distance d from the antenna (or IRS element) placed in the origin of the reference plane, shown in Fig. 2.3. In the channel equation (2.8) the matrix $\mathbf{D}$ is represented by

$$\mathbf{D} = \mathrm{diag}\left([\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_{N_{\mathrm{ch}}}] \right) \in \mathbb{C}^{N_{\mathrm{ch}} \times N_{\mathrm{ch}}}, \tag{2.15}$$

where α is the complex circularly symmetric Gaussian path gain for each channel path.

The channel vector $\mathbf{g}$ in (2.4) is similarly defined as $\mathbf{H}$ above but takes the following

form:

$$\mathbf{g} = \mathbf{d}_{\mathrm{g}}^{\mathrm{T}} \mathbf{A}_{\mathrm{g}}^{\mathrm{H}}, \tag{2.16}$$

where

$$\mathbf{d}_{\mathrm{g}} = [\alpha_{g,1}, \alpha_{g,2}, \alpha_{g,3}, \ldots, \alpha_{g,N_{\mathrm{ch}}}]^{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{ch}} \times 1}, \tag{2.17}$$

and $\mathbf{A}_{\mathrm{g}} \in \mathbb{C}^{M \times N_{\mathrm{ch}}}$ is constructed as the transmit array response matrix as in (2.10) above, but describes the array response in the model where the IRS is the "transmit-side" and the user is the receive-side:

$$\mathbf{A}_{\mathrm{g}} = [\mathbf{a}_{\mathrm{g}}(\varphi_{\mathrm{g},1}, \psi_{\mathrm{g},1}), \ldots, \mathbf{a}_{\mathrm{g}}(\varphi_{\mathrm{g},N_{\mathrm{ch}}}, \psi_{\mathrm{g},N_{\mathrm{ch}}})] \in \mathbb{C}^{M \times N_{\mathrm{ch}}}, \tag{2.18}$$

and the array response vector for the $\mathbf{g}$ channel is defined as in (2.11) where

$$\mathbf{a}_{\mathrm{g}}(\varphi, \psi) = \left[\mathrm{e}^{jw_{\mathrm{g}}(1,1,\varphi,\psi)}, \ldots, \mathrm{e}^{jw_{\mathrm{g}}(m_x,m_y,\varphi,\psi)}, \ldots, \mathrm{e}^{jw_{\mathrm{g}}(M_x,M_y,\varphi,\psi)}\right]^{\mathrm{T}} \in \mathbb{C}^{M \times 1}, \tag{2.19}$$

and finally the phase offset function is defined as:

$$w_{\mathrm{g}}(x, y, \varphi, \psi) = \frac{2\pi}{\lambda} d_{\mathrm{IRS}}((x-1)\sin\varphi\sin\psi + (y-1)\cos\varphi). \tag{2.20}$$

The azimuth angles of arrival and departure for both channels $\mathbf{H}$ and $\mathbf{g}$ are taken from the uniform distribution $[-180°, 180°]$, and the elevation angles of arrival and departure are taken from the uniform distribution $[-90°, 90°]$. Note for the channel $\mathbf{g}$, the angles of arrival are irrelevant due to the single antenna nature of the user.

An accurate and detailed channel model is important to properly simulate and test the practical impact that our methods have on the performance of this transmission model. We see in the next chapter that IRS phase shifts are directly optimized and grouped based on

the information gathered from the channel and the coherence between elements. Therefore the mathematical relationships between elements set in the channel model must be accurate for grouping to have a proper effect.

Chapter 3

Preliminary Problem Formulation and Grouping

Abstract

In this chapter we start by establishing the upper bound for the effective SNR that is pursued by our optimization methods, as well as constructing the overhead-aware achievable rate equation that this work attempts to maximize. Next, we reference the grouping scheme used in a previous work to combat overhead, and point out its room for improvement in terms of efficient grouping. Then we propose an alternate scheme with the capacity for designing the ideal number of groups.

3.1 Preliminary Problem Formulation

The target of this work is to maximize the effective SNR by optimizing the phase shift values of the IRS elements, which brings us to the optimization problem:

$$\max_{\mathbf{q},\boldsymbol{\phi}} \left|\mathbf{q}^{\mathrm{H}}\mathbf{V}_{\mathrm{r}}\boldsymbol{\phi}\right|^2 \tag{3.1a}$$

$$s.t.\ \|\mathbf{q}\|_2 = 1 \text{ and } |\phi_m| \leq 1,\ \forall m. \tag{3.1b}$$

It follows logically that the constraint on $\boldsymbol{\phi}$ could be set to equality without affecting optimality, henceforth in this work the constraints in (3.1b), both are set to equality, yielding

an equivalent problem

$$\max_{\mathbf{q},\boldsymbol{\phi}} \left|\mathbf{q}^{\mathrm{H}}\mathbf{V}_{\mathrm{r}}\boldsymbol{\phi}\right|^2 \tag{3.2a}$$

$$s.t.\ \|\mathbf{q}\|_2 = 1 \text{ and } |\phi_m| = 1,\ \forall m. \tag{3.2b}$$

Using previously formulated optimization techniques such as [2] and [30], the problem in (3.2) can be simplified. First, we define the singular value decomposition of the combined channel matrix as

$$\mathbf{V}_{\mathrm{r}} = \sum_{i=1}^{\gamma_{\mathrm{r}}} \lambda_i\left(\mathbf{V}_{\mathrm{r}}\right)\mathbf{u}_i\mathbf{r}_i^{\mathrm{H}},$$

where $\gamma_r = \mathrm{rank}\left(\mathbf{V}_{\mathrm{r}}\right)$, $\mathbf{u}_i \in \mathbb{C}^{N\times 1}$ is the i-th largest left singular vector, and $\mathbf{r}_i \in \mathbb{C}^{M\times 1}$ is the i-th largest right singular vector. Now, utilizing optimal beamforming techniques, the objective in (3.2) can be upper bounded by

$$\left|\mathbf{q}^{\mathrm{H}}\mathbf{V}_{\mathrm{r}}\boldsymbol{\phi}\right|^2 \le \lambda_1^2\left(\mathbf{V}_{\mathrm{r}}\right)\left|\mathbf{r}_1^{\mathrm{H}}\boldsymbol{\phi}\right|^2, \tag{3.3}$$

where the equality holds when $\mathbf{q}$ is set as the largest left singular vector of $\mathbf{V}_{\mathrm{r}}$. Therefore, we have

$$\mathbf{q}^{\star} = \mathbf{u}_1. \tag{3.4}$$

Now (3.3) can be further upper bounded as

$$\lambda_1^2\left(\mathbf{V}_{\mathrm{r}}\right)\left|\mathbf{r}_1^{\mathrm{H}}\boldsymbol{\phi}\right|^2 \le \lambda_1^2\left(\mathbf{V}_{\mathrm{r}}\right)\|\mathbf{r}_1\|_1^2, \tag{3.5}$$

where equality is achieved when

$$\boldsymbol{\phi}^{\star} = \mathrm{e}^{j\angle\mathbf{r}_1}. \tag{3.6}$$

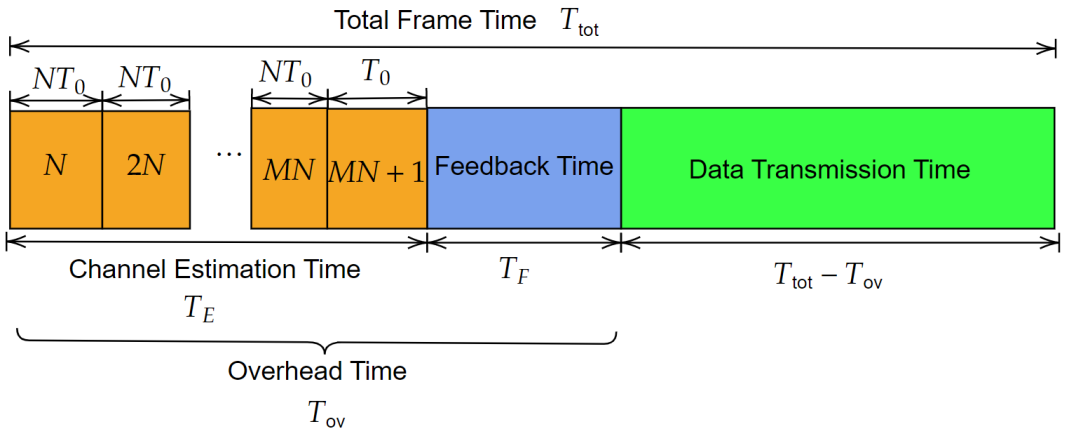

Figure 3.1: Model of total frame time showing its composition of overhead and data transmission.

This solution is similar to the equal gain transmission solution as formulated in [30], thus we will refer to our solution as the equal gain reflection (EGR) solution. Thus, in (3.5), we have shown the achievable upper bound to effective SNR in the given system model, utilizing two established solution techniques. However, this SNR cannot be fully taken advantage of because of the trade-off that exists when dealing with the overhead required to implement $\boldsymbol{\phi}^{\star}$ at the IRS. There exists a greater channel estimation and feedback cost that contribute to the overhead, which leads to an overall smaller achievable data rate, which is the main measure we use to gauge performance of our signal model. The overhead-aware achievable rate that is used is defined as:

$$R = \left(1 - \frac{T_{\text{ov}}}{T_{\text{tot}}}\right) B \log_2 \left(1 + \frac{p \left|\mathbf{q}^{\text{H}} \mathbf{V}_r \boldsymbol{\phi}\right|^2}{B N_0}\right), \tag{3.7}$$

where T_{tot} is the total frame time, T_{ov} is the time required for overhead, B is the bandwidth, p is the transmit power of the signal, and N_0 is the noise power spectral density. As T_{ov} has a large impact on the overall achievable rate, it is just as important to keep T_{ov} low as it is to maximize the effective SNR in this function. The two overhead terms that make up $T_{\text{ov}} = T_{\text{F}} + T_{\text{E}}$ are the time required for feedback and time required for channel estimation, respectively, as made clear in Fig. 3.1. As we noted in Chapter 2.3, $T_{\text{E}} = (NM + 1)\, T_0$, and

the value of T_{F} is

$$T_{\mathrm{F}} = \frac{Mb_{\mathrm{F}}}{B_{\mathrm{F}}\log_2\left(1+\frac{p_{\mathrm{F}}|h_{\mathrm{F}}|^2}{N_0 B_{\mathrm{F}}}\right)}, \tag{3.8}$$

where B_{F} and p_{F} are the bandwidth reserved for reliable feedback and power dedicated to feedback transmission, respectively. Therefore the expanded overhead-aware achievable rate equation is

$$R = \left(1 - \frac{(NM+1)\,T_0}{T_{\mathrm{tot}}} - \frac{Mb_{\mathrm{F}}}{T_{\mathrm{tot}}B_{\mathrm{F}}\log_2\left(1+\frac{p_{\mathrm{F}}|h_{\mathrm{F}}|^2}{N_0 B_{\mathrm{F}}}\right)}\right) B\log_2\left(1+\frac{p\left|\mathbf{q}^{\mathrm{H}}\mathbf{V}_r\boldsymbol{\phi}\right|^2}{BN_0}\right). \tag{3.9}$$

Due to the multiplicative effects of overhead costs, [2] had a pessimistic view on the potency of IRS aided systems, even so far as asserting that IRSs were not effective in increasing the achievable rate in any system with larger than one transmit antenna and one receive antenna when implementing a fully optimized $\boldsymbol{\phi}^\star$ solution as in (3.6). This thesis attempts to identify the sub-optimal trade-off between effective SNR gain and overhead reduction to achieve superior rate performance.

3.2 Existing Grouping Method

To combat this overhead trade-off issue, [4] introduced a method that significantly reduced overhead without much SNR loss. The concept was to take advantage of an IRSs closely packed elements and the assumption that nearby elements are highly correlated. Instead of assigning an optimized reflection coefficient to every IRS element, they proposed that nearby elements were combined into one of K evenly sized groups that could share one reflection coefficient, optimized for the group. This method attained a net achievable rate increase due to marginal effective SNR loss with a much smaller overhead cost than without grouping.

They accomplished these improvements by not only reducing the number of bit groups required for feedback, but also reducing the size of the channel matrix $\mathbf{V}_{\mathrm{r}}$ that needs to be estimated, as each grouped shared an estimated channel. This causes the values of M in the

overhead term of (3.9) to be reduced to K. Instead of estimating $\mathbf{V}_{\mathrm{r}} = [\mathbf{v}_{\mathrm{r},1}, \mathbf{v}_{\mathrm{r},2}, \cdots \mathbf{v}_{\mathrm{r},M}]$, only some channel vectors needed to be estimated in [4], i.e. $\widetilde{\mathbf{V}}_{\mathrm{r}} = \left[\mathbf{v}_{\mathrm{r},1}, \mathbf{v}_{\mathrm{r},Z}, \cdots \mathbf{v}_{\mathrm{r},Z(K-1)}\right]$ where $Z = \frac{M}{K}$ is the number of elements per group, with the important constraint that Z is an integer. Once the channel matrix $\widetilde{\mathbf{V}}_{\mathrm{r}}$ is estimated, the optimization problem is given by

$$\max_{\tilde{\mathbf{q}}, \tilde{\boldsymbol{\phi}}} \left| \tilde{\mathbf{q}}^{\mathrm{H}} \widetilde{\mathbf{V}}_{\mathrm{r}} \tilde{\boldsymbol{\phi}} \right|^2 \tag{3.10a}$$

$$s.t.\ \|\tilde{\mathbf{q}}\|_2 = 1 \text{ and } \left|\tilde{\phi}_k\right| = 1,\ \forall k, \tag{3.10b}$$

where $\tilde{\mathbf{q}} \in \mathbb{C}^{N \times 1}$ and $\tilde{\boldsymbol{\phi}} \in \mathbb{C}^{K \times 1}$ can be shown to be solved by the same solution method as in (3.4) and (3.6), where the sub-sampled channel matrix $\widetilde{\mathbf{V}}_{\mathrm{r}}$ is used to optimize the beamforming and reflection coefficient vectors. Then these solution vectors are applied to the actual combined channel matrix $\mathbf{V}_{\mathrm{r}}$ which gives the output effective SNR. This achieves a much higher achievable rate than without grouping (due to the large reduction in overhead costs), but there is a significant reduction in effective SNR due to $\tilde{\mathbf{q}}$ not being optimal for the full dimensional channel matrix and the obvious reduction in effective SNR due to the effects of grouping and the loss of degrees of optimization freedom.

Another aspect of their grouping strategy that allows for improvement is the choice of best grouping ratio $\rho = \frac{K}{M}$. As shown in [4] there exists an ideal grouping ratio (IGR) where the trade-off between achievable effective SNR gain and overhead is properly handled to improve achievable rate. Unfortunately, because a grouping ratio would need to be chosen before channel estimation takes place, there is no way to determine the IGR without the channel state information (CSI) corresponding to the full dimensional channel matrix prior to channel estimation. Therefore the IGR is unattainable in practice. Instead, a rule-of-thumb $\rho = \frac{1}{4}$ was suggested in [4] to achieve some benefits of grouping [4]. This inspired the creation of our proposed strategy, which is able to determine the ideal grouping ratio $\rho^\star$ at every channel realization.

3.3 Proposed Grouping Scheme - General Principles

Our proposed grouping scheme takes advantage of the IGR by utilizing the entire $\mathbf{V}_{\mathrm{r}}$ channel matrix at every channel realization, thereby having the information present to implement the ideal grouping. In order to have access this information, it is assumed that the entire channel must be estimated, which means our grouping does not reduce T_{E}, but it still can considerably reduce T_{F} by sending phase shift information with designated groupings, reducing T_{F} by a factor of $\rho^{\star}$ and keeping overhead as low as possible. We postulate that with the given overhead-aware model, any marginal overhead cost increase over the grouping method in [4] is more than made up for by the effective SNR gain related to our proposed reflection coefficient optimization methods within each grouping strategy. Another incidental way to increase effective SNR in our solution is by optimizing the beamforming vector according to (3.4). Because we have more information in our full channel $\mathbf{V}_{\mathrm{r}}$ we can beamform more accurately. With this $\mathbf{q}^{\star}$ solution assumed in all of our proposed methods, we can attempt to achieve the upper bound in (3.5) with our $\boldsymbol{\phi}^{\star}$ solutions, constrained of course by the grouping of the IRS elements. In the following chapters we propose 3 unique methods that offer sub-optimal solutions to replace $\boldsymbol{\phi}^{\star}$ vector, leveraging the information contained in the full channel matrix.

Chapter 4

Method 1 – Uniform Phase Grouping

Abstract

In this chapter, we take the first step in implementing the proposed grouping scheme from the previous chapter and formulate it as an effective SNR maximization problem. We also develop our first, and simplest of the three grouped reflection coefficient design methods.

4.1 Grouped Phase Optimization

The goal of this optimization method is to find $\boldsymbol{\phi}^{\star}$ that maximizes (3.2), but we are limited by the number of reflection coefficients K that are available to be optimized. To account for the reduction in optimization freedom, we adjust the constraint into our model and define

$$\boldsymbol{\phi} = \left(\tilde{\boldsymbol{\phi}} \otimes \mathbf{1}_Z\right), \tag{4.1}$$

where $\tilde{\boldsymbol{\phi}} \in \mathbb{C}^{K\times 1}$ is the vector containing the reflection coefficients for each group k, and $\mathbf{1}_Z$ is the length Z vector containing all ones that acts to uniformly distribute $\tilde{\phi}_k$ to each element in the group. The effective implemented reflection coefficients considering grouping are shown on the example IRS in Fig. 4.1.

In order to formulate the maximization problem to optimize $\tilde{\boldsymbol{\phi}}$ we must first adapt the

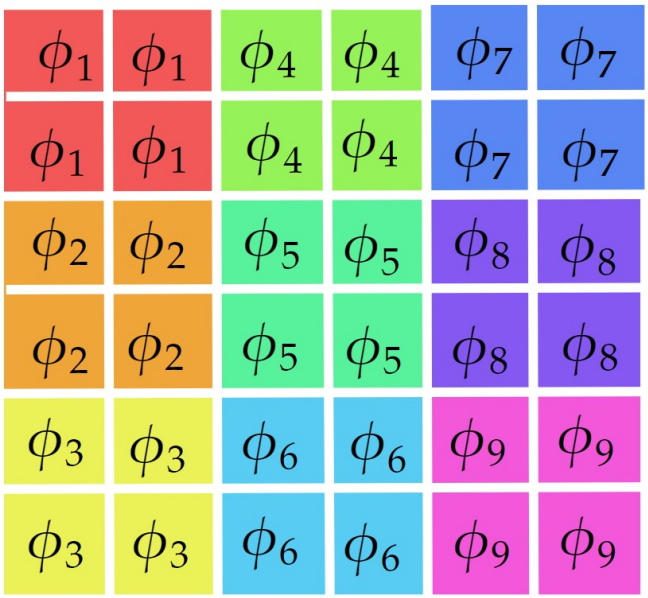

Figure 4.1: Example of the Method 1 grouping implementation on 6x6 IRS Array with group size $K = 9$ and elements per group $Z = 4$.

problem from (3.1), using our new definition of $\boldsymbol{\phi}$ in (4.1):

$$\max_{\mathbf{q},\tilde{\boldsymbol{\phi}}} \left|\mathbf{q}^{\mathrm{H}}\mathbf{V}_{\mathrm{r}}\left(\tilde{\boldsymbol{\phi}} \otimes \mathbf{1}_Z\right)\right|^2 \tag{4.2a}$$

$$s.t.\ \|\mathbf{q}\|_2 = 1 \text{ and } \left|\tilde{\phi}_k\right| = 1,\ \forall k. \tag{4.2b}$$

We see in (4.2a) that the optimization of $\mathbf{q}^\star$ does not depend on $\tilde{\boldsymbol{\phi}}^\star$ and is thus the same as (3.4). Therefore our focused effective SNR maximization problem is reduced to

$$\max_{\tilde{\boldsymbol{\phi}}} \left|\mathbf{r}_1^{\mathrm{H}}\left(\tilde{\boldsymbol{\phi}} \otimes \mathbf{1}_Z\right)\right|^2 \tag{4.3a}$$

$$s.t.\ \left|\tilde{\phi}_k\right| = 1,\ \forall k. \tag{4.3b}$$

This problem can be expanded to better understand the effect of each individual $\tilde{\phi}_k$ as

follows:

$$\max_{\tilde{\boldsymbol{\phi}}} \left| \tilde{\phi}_1 \sum_{i=1}^{Z} r_i^* + \tilde{\phi}_2 \sum_{i=1}^{Z} r_{i+Z}^* + \cdots + \tilde{\phi}_K \sum_{i=1}^{Z} r_{i+(K-1)Z}^* \right| \tag{4.4a}$$

$$s.t. \left| \tilde{\phi}_k \right| = 1, \ \forall k. \tag{4.4b}$$

This maximization problem can be simplified as the product of two vectors

$$\max_{\tilde{\boldsymbol{\phi}}} \left| \tilde{\mathbf{r}}_{\mathrm{M1}}^{\mathrm{H}} \tilde{\boldsymbol{\phi}} \right|^2 \tag{4.5a}$$

$$s.t. \left| \tilde{\phi}_k \right| = 1, \ \forall k, \tag{4.5b}$$

where our Method 1 grouped $\mathbf{r}_1$ vector is defined as

$$\tilde{\mathbf{r}}_{\mathrm{M1}} = \left[\sum_{i=1}^{Z} r_i, \sum_{i=1}^{Z} r_{i+Z}, \cdots, \sum_{i=1}^{Z} r_{i+(K-1)Z} \right]^{\mathrm{T}} \in \mathbb{C}^{K \times 1}. \tag{4.6}$$

From (4.5a) it is now clear that we can still use our EGR solution in (3.6), except where the phases of $\tilde{\phi}_k$ match the sums of the grouped elements of the $\mathbf{r}_1$ vector as seen in

$$\tilde{\phi}_k^{\star} = \mathrm{e}^{j\angle\left(\sum_{i=1}^{Z} r_{i+(k-1)Z}\right)}, \ \forall k. \tag{4.7}$$

This optimal solution to (4.3a) now achieves the effective SNR upper bound of

$$\left| \mathbf{q}^{\mathrm{H}} \mathbf{V}_{\mathrm{r}} \left(\tilde{\boldsymbol{\phi}} \otimes \mathbf{1}_Z \right) \right|^2 \leq \lambda_1^2 \left(\mathbf{V}_{\mathrm{r}} \right) \left\| \tilde{\mathbf{r}}_{\mathrm{M1}} \right\|_1^2. \tag{4.8}$$

This upper bound varies with the chosen grouping ratio which further motivates the importance of utilizing the IGR for these methods. We expect the upper bound to monotonically increase as ρ increases (where the upper bound is largest when there is no grouping, i.e. the upper bound in (3.5)), but in Chapter 4.2 we see that when considering overhead, there exists a trade-off for ρ in the achievable rate.

4.2 Rate Equation and Overhead Effects

To conclude this chapter, we model the feedback overhead that was reduced and formulate a modified achievable rate equation. Because grouping reduces the number of distinct phase shifts that need to be sent over the feedback channel, the bits required by the IRS controller are reduced as well, from Mb_{F} bits to only $M\rho b_{\mathrm{F}}$ bits. With all of the above considered, the modified achievable rate equation is defined as

$$\left(1-\frac{(NM+1)\,T_0}{T_{\mathrm{tot}}}-\frac{M\rho b_{\mathrm{F}}}{T_{\mathrm{tot}}B_{\mathrm{F}}\log_2\left(1+\frac{p_{\mathrm{F}}|h_{\mathrm{F}}|^2}{N_0B_{\mathrm{F}}}\right)}\right)B\log_2\left(1+\frac{p\left|\mathbf{q}^{\mathrm{H}}\mathbf{V}_r\left(\tilde{\boldsymbol{\phi}}\otimes\mathbf{1}_Z\right)\right|^2}{BN_0}\right). \quad (4.9)$$

This grouping method along with $\tilde{\boldsymbol{\phi}}$ optimization successfully improves effective SNR while also reducing T_{F}, which increases the overall achievable rate which we can see in Chapter 7.

Chapter 5

Method 2 – Double Phase Grouping

Abstract

In this chapter, we present our second method for generating effective SNR gain by leveraging a second, smaller phase shift kernel alongside $\tilde{\boldsymbol{\phi}}$ that we introduce as $\boldsymbol{\varepsilon}$. We formulate and show that an alternating maximization algorithm can be used to sub-optimally solve this problem.

5.1 Formulation: Method 2

From (4.2a) in our formulation of Method 1, we assume that the IRS reflection coefficients are exactly repeated for each element in the group. In grouping Method 2, we propose a strategy that increases the degrees of freedom for optimization by introducing a new kernel, $\boldsymbol{\varepsilon} \in \mathbb{C}^{Z\times 1}$ that replaces the $\mathbf{1}_Z$ vector in (4.2a) but contains phases just like $\tilde{\boldsymbol{\phi}}$. Due to the amplitude's invariance to phase rotation, these two vectors can be implemented simultaneously through the Kronecker product. Therefore our $|\phi_m| = 1, \forall m$ constraint remains unbroken. Now each IRS element has a reflection coefficient that is a product of two separate phases, as shown in Fig. 5.1. Once again, this optimization strategy utilizes the full channel $\mathbf{V}_{\mathrm{r}}$, so we again use $\mathbf{q}^\star$ in (3.4) and can set up our new effective SNR maximization problem as

$$\max_{\tilde{\boldsymbol{\phi}},\boldsymbol{\varepsilon}} \left|\mathbf{r}_1^{\mathrm{H}}\left(\tilde{\boldsymbol{\phi}} \otimes \boldsymbol{\varepsilon}\right)\right|^2 \tag{5.1a}$$

$$s.t. \left|\tilde{\phi}_k\right| = 1 \text{ and } |\boldsymbol{\varepsilon}_z| = 1, \ \forall k, z. \tag{5.1b}$$

Figure 5.1: Example of Method 2 grouping implementation on 6x6 IRS Array with group number K = 9, and elements per group Z = 4.

Unfortunately the problem in (5.2a) is not directly solvable, as the solution for $\tilde{\boldsymbol{\phi}}$ depends on $\boldsymbol{\varepsilon}$ and vice versa. However, we can form two optimally solvable sub-problems, one where $\boldsymbol{\varepsilon}$ is given and the maximization problem is optimally solved for $\tilde{\boldsymbol{\phi}}$ and one where $\tilde{\boldsymbol{\phi}}$ is given and is optimally solved for $\boldsymbol{\varepsilon}$. Similarly to (4.4a), the maximization problem in (5.1a), can first be expanded to

$$\max_{\tilde{\boldsymbol{\phi}}} \left| \tilde{\phi}_1 \sum_{z=1}^{Z} \varepsilon_z r_z^* + \tilde{\phi}_2 \sum_{z=1}^{Z} \varepsilon_z r_{z+Z}^* + \cdots + \tilde{\phi}_K \sum_{z=1}^{Z} \varepsilon_z r_{z+(K-1)Z}^* \right|^2 \tag{5.2a}$$

$$s.t. \ \left| \tilde{\phi}_k \right| = 1, \ \forall k, \tag{5.2b}$$

where $\boldsymbol{\varepsilon}$ is fixed. To solve this problem for $\tilde{\boldsymbol{\phi}}$, we have a straightforward optimal EGR solution when $\boldsymbol{\varepsilon}$ is fixed:

$$\tilde{\phi}_k^\star = \mathrm{e}^{j\angle\left(\sum_{z=1}^{Z} \varepsilon_z^* r_{z+(k-1)Z}\right)}, \ \forall k. \tag{5.3}$$

We can also reformulate (5.2a) to make the solution for $\boldsymbol{\varepsilon}^\star$ more obvious, as in

$$\max_{\boldsymbol{\varepsilon}} \left| \varepsilon_1 \sum_{k=1}^{K} \tilde{\phi}_k r^*_{1+(k-1)Z} + \varepsilon_2 \sum_{k=1}^{K} \tilde{\phi}_k r^*_{2+(k-1)Z} + \cdots \varepsilon_Z \sum_{k=1}^{K} \tilde{\phi}_k r^*_{Z+(k-1)Z} \right|^2 \tag{5.4a}$$

$$s.t.\ |\varepsilon_z| = 1,\ \forall z, \tag{5.4b}$$

where $\tilde{\boldsymbol{\phi}}$ is fixed. Now it is clear that the EGR solution can be used for $\boldsymbol{\varepsilon}^\star$ as well. When $\tilde{\boldsymbol{\phi}}$ is fixed we can say

$$\varepsilon_z^\star = \mathrm{e}^{j\angle\left(\sum_{k=1}^{K} \tilde{\phi}_k^* r_{z+(k-1)Z}\right)},\ \forall z. \tag{5.5}$$

5.2 Algorithm

Now that we have demonstrated that we have two sub-problems that can be optimally solved, we can use alternating optimization as described in Algorithm 1 to converge on a sub-optimal solution. In our algorithm, a random initial phase vector $\boldsymbol{\varepsilon}_0$ is used to find the first optimal $\tilde{\boldsymbol{\phi}}$ according to the solution in (5.3). Then the output vector $\tilde{\boldsymbol{\phi}}^\star$ is used as an input to the solution in (5.5) to find the optimal $\boldsymbol{\varepsilon}$. This output vector $\boldsymbol{\varepsilon}^\star$ can then be used as an input to (5.3) again, and thus begins the process of repeating this optimization. Since both functions are optimally solvable, the process of effective SNR gain improvement is monotonically increasing every iteration until significant convergence, as seen in Fig. 5.2.

Algorithm 1 Method 2 Alternating Optimization

Input: $\mathbf{r}_1, K, M, \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_0$
Output: $\tilde{\boldsymbol{\phi}}^\star, \boldsymbol{\varepsilon}^\star$
iteration = 0
repeat
  iteration = iteration +1
  Fix $\boldsymbol{\varepsilon}$ and maximize effective SNR with $\tilde{\boldsymbol{\phi}}$ according to (5.3)
  Fix $\tilde{\boldsymbol{\phi}}$ and maximize effective SNR with $\boldsymbol{\varepsilon}$ according to (5.5)
until Convergence: $\left|\mathbf{r}_1^{\mathrm{H}}\left(\tilde{\boldsymbol{\phi}} \otimes \boldsymbol{\varepsilon}\right)\right|^2_{(\text{iteration})} - \left|\mathbf{r}_1^{\mathrm{H}}\left(\tilde{\boldsymbol{\phi}} \otimes \boldsymbol{\varepsilon}\right)\right|^2_{(\text{iteration}-1)} \leq 10^{-7}$
$\tilde{\boldsymbol{\phi}}^\star = \tilde{\boldsymbol{\phi}}, \boldsymbol{\varepsilon}^\star = \boldsymbol{\varepsilon}$.

While in Method 1 we derived a strict upper bound for the effective SNR, for this method we cannot derive an explicit upper bound solution for Algorithm 1, we only have the convergence point which is shown in Fig. 5.2. We set convergence when the difference in effective SNR from one iteration to the next is less than 10^{-7}. Note that the loose upper bound in (3.5) still applies for all of our proposed methods.

5.3 Rate Equation and Overhead Effects

Since this grouping method requires that an additional Z phases are sent over the feedback channel, corresponding to the length Z kernel $\boldsymbol{\varepsilon}$, the T_{F} in the overhead model increases by Zb_{F} bits. Thus the feedback overhead term for Method 2 is defined as

$$T_{\mathrm{F}} = \frac{(M\rho + Z)\, b_{\mathrm{F}}}{B_{\mathrm{F}} \log_2\left(1 + \frac{p_{\mathrm{F}}|h_{\mathrm{F}}|^2}{N_0 B_{\mathrm{F}}}\right)} \tag{5.6}$$

The total overhead of this method increases over Method 1. However, any costs from increased overhead were overcome by significant effective SNR gain improvements with our alternating optimization. With given ρ our achievable rate equation for Method 2 is given by:

$$\left(1 - \frac{(NM+1)\, T_0}{T_{\mathrm{tot}}} - \frac{(K+Z)\, b_{\mathrm{F}}}{T_{\mathrm{tot}} B_{\mathrm{F}} \log_2\left(1 + \frac{p_{\mathrm{F}}|h_{\mathrm{F}}|^2}{N_0 B_{\mathrm{F}}}\right)}\right) B \log_2\left(1 + \frac{p\left|\mathbf{q}^{\mathrm{H}} \mathbf{V}_r \left(\tilde{\boldsymbol{\phi}}^{\star} \otimes \boldsymbol{\varepsilon}^{\star}\right)\right|^2}{BN_0}\right). \tag{5.7}$$

Remark 1 *Even though the model and Fig. 5.1 show examples where $Z \leq K$, it is still possible for $\boldsymbol{\varepsilon}$ to contain more elements than $\tilde{\boldsymbol{\phi}}$. In this case, there are more meta-atoms per group than there are groups in the IRS. Because the sizes of $\tilde{\boldsymbol{\phi}}$ and $\boldsymbol{\varepsilon}$, K and Z respectively, are complementary around the point $K = Z$ (for any set of values of K and Z there exists another set where K and Z are swapped), these vectors can be considered interchangeable in terms of which describes the number of groups and which describes the size of the groups.*

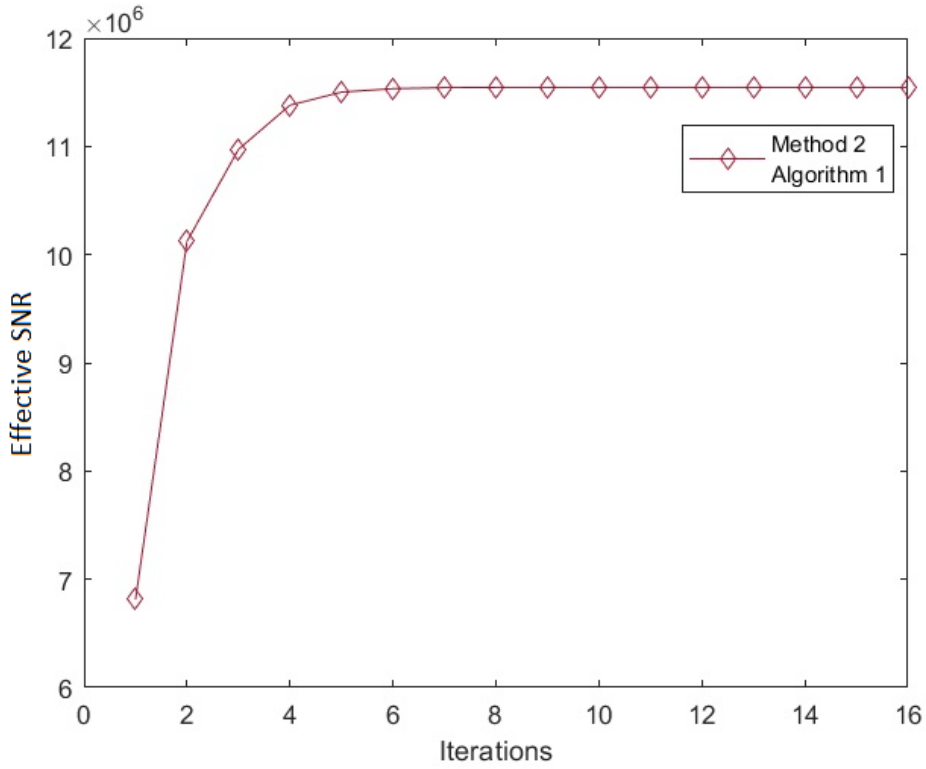

Figure 5.2: Effective SNR performance and convergence of Method 2 algorithm with respect to number of iterations. $K = 6, M = 36$. Here we execute the algorithm for a single channel realization according to the channel model in Chapter 2.4.

Because $\tilde{\boldsymbol{\phi}}$ and $\boldsymbol{\varepsilon}$ are interchangeable, it would be logical to assume that the effective SNR would be the same no matter the order of optimization. Furthermore, the numerator of the overhead term T_{F} in (5.7) *does not change if the values of K and Z are swapped. This means that the overhead for Method 2's achievable rate equation is symmetrical (and convex parabolic) in nature about the center point $Z = K$ or $\rho = 1/\sqrt{M}$ (if this point exists).*

Method 2 was formulated to highlight the relative improvements in effective SNR over Method 1 that can be achieved in our scheme when additional room for flexible grouping is created. In the following chapter we attempt to advance this idea to its limits and open up a substantial amount of room for flexible grouping design by relaxing two important assumptions, that groups are made up of nearby IRS elements and that groups must be equally sized.

Chapter 6

Method 3 – Dynamic Grouping and Phase Optimization

Abstract

In this chapter, we present our third method for generating improved effective SNR gain which proposes a new generalized framework that allows for dynamically-sized and distributed groups. A rate-improving algorithm is developed to efficiently partition the IRS elements into coherent groups. Another algorithm is introduced that initializes the first, and improves both effective SNR performance and iterations until convergence.

6.1 Problem Formulation

In this chapter, we design a way to free up even more degrees of freedom that allow us to significantly improve the effective SNR at the cost of a small increase of feedback overhead. Before this chapter, the assumption was made that all IRS groups were of equal size and the elements were close together, but we relinquish those assumptions to further increase effective SNR. This framework is a more general form of the formulation of Method 1, so the implementation of Method 1 is a special case of Method 3's general model. The proposed grouping method is depicted in Fig. 6.1. We see that each element has full freedom of choice, it can take the phase shift of any group, without consideration of location or group size. While it is true that element correlation exists within close elements of an IRS [4], by structuring the groups with dynamic sizes and locations we are allowing the elements to be partitioned in a way that captures any coherence of optimized phases ϕ_m. As such, to fit the model presented in Method 3, we replace Z with Z_k as the dynamic number of elements

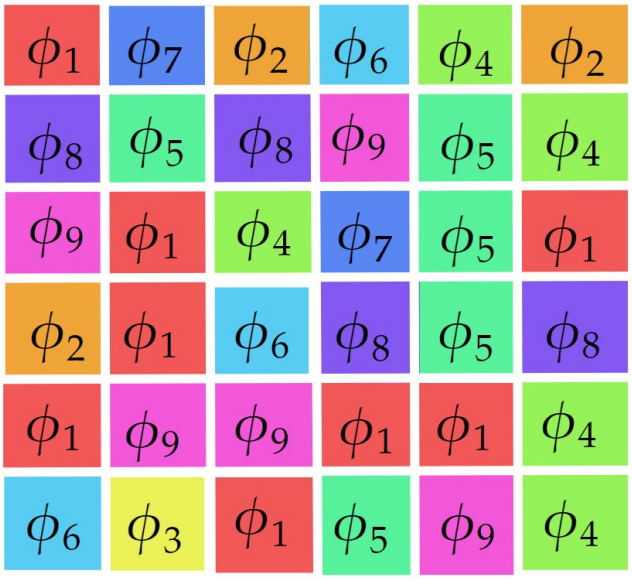

Figure 6.1: Example of Method 3 dynamic grouping implementation on 6x6 IRS Array with dynamic group size and variable number of elements per group.

in the k-th group. Looking back at the formulation for (4.8) and (4.6), we can adjust these problems to fit our new dynamic group size organization:

$$\max_{\tilde{\boldsymbol{\phi}}} \left| \tilde{\mathbf{r}}_{\mathrm{M3}}^{\mathrm{H}} \tilde{\boldsymbol{\phi}} \right|^2 \tag{6.1a}$$

$$s.t. \left| \tilde{\phi}_k \right| = 1, \ \forall k, \tag{6.1b}$$

where our Method 3 grouped $\mathbf{r}_1$ vector is defined

$$\tilde{\mathbf{r}}_{\mathrm{M3}} = \left[\sum_{i \in S_1} r_i, \sum_{i \in S_2} r_i, \cdots, \sum_{i \in S_K} r_i \right]^{\mathrm{T}} \in \mathbb{C}^{K \times 1}, \tag{6.2}$$

and $S_k \subset \{1, 2, \cdots, M\}$ denotes the set of IRS elements distributed into group k (which are no longer required to be adjacent to one another) that share one reflection coefficient, and form one disjoint union within $\{1, 2, \cdots, M\}$, i.e. $\bigcup_{k=1}^{K} S_k = \{1, 2, \cdots, M\}$. We can again find an explicit upper bound, achievable with an EGR solution, as in (4.8) to understand

the performance of this strategy:

$$\left|\mathbf{q}^{\mathrm{H}}\mathbf{V}_{\mathrm{r}}\boldsymbol{\phi}_{\mathrm{M3}}\right|^2 \leq \lambda_1^2\left(\mathbf{V}_{\mathrm{r}}\right)\left\|\tilde{\mathbf{r}}_{\mathrm{M3}}\right\|_1^2. \tag{6.3}$$

Our $\boldsymbol{\phi}_{\mathrm{M3}}^{\star}$ solution vector is always designed to match the phases of the elements of the vector $\tilde{\mathbf{r}}_{\mathrm{M3}}$, and because we define

$$\left\|\tilde{\mathbf{r}}_{\mathrm{M3}}\right\|_1 = \sum_{k=1}^{K}\left|\sum_{i\in S_k} r_i\right|, \tag{6.4}$$

by maximizing $\left\|\tilde{\mathbf{r}}_{\mathrm{M3}}\right\|_1^2$ we are essentially finding the best way to combine elements of $\mathbf{r}_1$ into K groups (of any size). This optimized grouping maximizes the sum of the magnitudes of all of the groups, which in turn significantly increases the effective SNR. This grouping optimization is presented in the following problem:

$$\max_{\{S_1,S_2,\cdots,S_K\}} \left\|\tilde{\mathbf{r}}_{\mathrm{M3}}\right\|_1^2 \tag{6.5a}$$

$$s.t.\ Z_k = \left|S_k\right|,\ \forall k, \tag{6.5b}$$

$$\sum_{k=1}^{K} Z_k = M, \tag{6.5c}$$

$$\bigcup_{k=1}^{K} S_k = \{1, 2, \cdots, M\}. \tag{6.5d}$$

where $|S_k|$ is defined as the cardinality of the set S_k and $\bigcup_{k=1}^{K} S_k$ is the disjoint union equal to the set of all m. The maximization problem above prepares the way for Algorithm 2, which attempts to find near-optimal partitioning for the sets of elements S_k as in (6.5a).

6.2 Method 3 Algorithm

The element allocation problem in (6.5a) is particularly complex. In order to try every combination of IRS elements and group sizes, the number of different combinations possible

is given by the Stirling Number of the Second Kind [31]:

$$\text{Number of Combinations} = \frac{1}{(K)!}\sum_{k=0}^{K}(-1)^{K-k}\begin{pmatrix} K \\ k \end{pmatrix}k^{M}. \tag{6.6}$$

Because IRSs are expected to contain a massive amount of elements [1], this number of combinations becomes extremely impractical. For instance, with only 16 IRS elements and 4 groups, the number of possible combinations becomes 1.7×10^8. In current research and development, there are IRSs that have been created and tested with 1100 elements [23]. Even if only 4 groups are used for this real-life example, that still means there are an incredibly large 7.68×10^{660} possible combinations [23]. Therefore it is paramount to find an expedited solution method to achieve the optimal sets $\{S_1^\star, S_2^\star, \cdots, S_K^\star\}$, or sets that perform close to optimally without an exhaustive search. Our algorithm outputs a near optimal solution, one that approaches the maximization of (6.5a) iteratively. We established in (6.5a) that our maximization depends on the magnitude of each group, therefore we use this as our metric for improving the effective SNR, defining the function f_1 as the following function:

$$f_1\left(S_k\right) = \left|\sum_{i \in S_k} \mathbf{r}_{1,i}\right|. \tag{6.7}$$

The proposed algorithm begins by taking in as input the K non-empty initial sets $\{S_{0,1}, S_{0,2}, \cdots, S_{0,K}\}$ as input to start the optimization. The initial steps of Algorithm 2 are straightforward, as the process consists of checking every IRS element to see if it is better to transfer it to another group, or to keep it where it is now. First the group "location" of the m-th IRS element is found and its previous set S_k is captured as the temporary set S_{tmp} before any changes are made (Steps 1-10). We then use our function f_1 to find the difference, according to our metric, that removing element m from S_{tmp} would cause. This is set as δ (Step 11). Then for every group $1:K$ (except for group k_{tmp}), the difference in the metric is measured, with and without the addition of m, which is defined as μ_k (Steps 12-14). As long

Algorithm 2 Method 3 Dynamic Grouping

Input: $\mathbf{r}_1, K, M, \{S_1, S_2, \cdots, S_K\} \leftarrow \{S_{0,1}, S_{0,2}, \cdots, S_{0,K}\}$

Output: $\{S_1^\star, S_2^\star, \cdots, S_K^\star\}$

iteration = 1
repeat
  iteration = iteration + 1
  for $m = 1 : M$ **do**
    for $k = 1 : K$ **do**
      if $m \in S_k$ **then**
        $S_{\text{tmp}} = S_k$
        $k_{\text{tmp}} = k$
      end if
    end for
    $\delta = f_1(S_{\text{tmp}}) - f_1(S_{\text{tmp}} \setminus \{m\})$
    for $k = \{1 : K\} \setminus \{k_{\text{tmp}}\}$ **do**
      $\mu_k = f_1(S_k \cup \{m\}) - f_1(S_k)$
    end for
    if $\max \mu_k \geq \delta$ **then**
      $k_{\text{best}} = \text{argmax}_i \, \mu_i$
      $S_{k_{\text{best}}} \leftarrow S_{k_{\text{best}}} \cup \{m\}$
      $S_{\text{tmp}} \leftarrow S_{\text{tmp}} \setminus \{m\}$
    end if
  end for
until Convergence: $\|\tilde{\mathbf{r}}_{\text{M3,iteration}}\|_1^2 - \|\tilde{\mathbf{r}}_{\text{M3,(iteration -1)}}\|_1^2 \leq 10^{-7}$ {According to (6.4)}
$\{S_1^\star, S_2^\star, \cdots, S_K^\star\} \leftarrow \{S_1, S_2, \cdots, S_K\}$

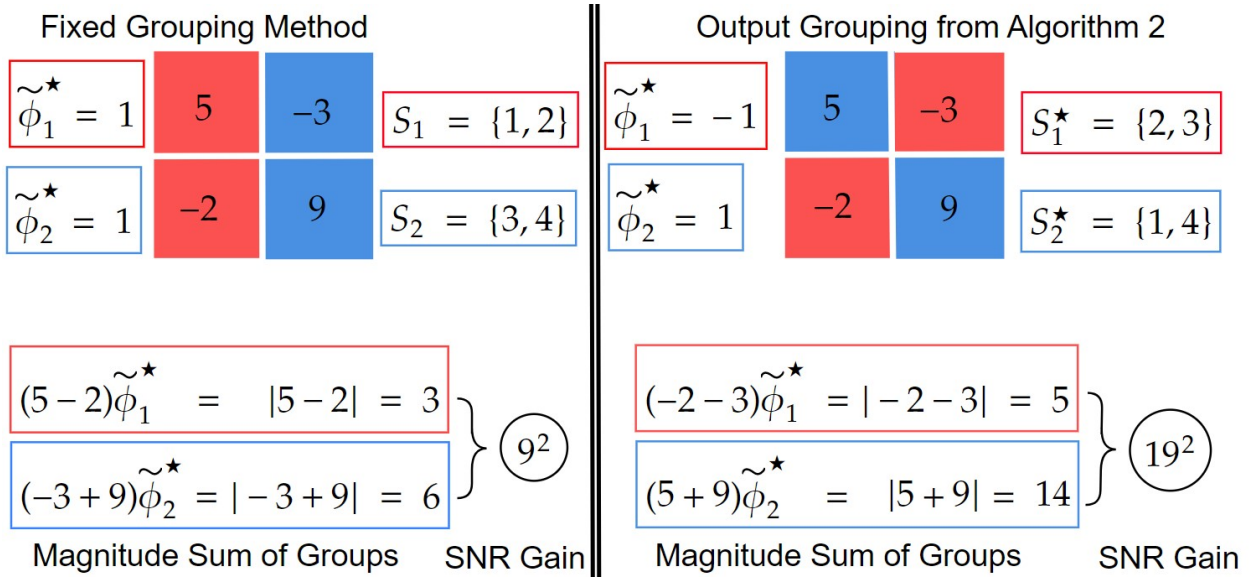

Figure 6.2: Method 3 toy example problem illustrating the effect of the efficient grouping of a 2x2 IRS on the sum of group magnitudes metric. On the left hand side is the default nearby grouping method and the right hand side is the optimal grouping that was output from Algorithm 2.

as one of the possible group switches out-benefits the cost δ (Step 14), then m is re-assigned (Steps 16-18). However, if δ is larger than $\mu_k, \forall k$, then m remains in its current group. This process is repeated for all m and then the set of $m = 1 : M$ is continuously repeated in the algorithm until we have convergence in our final metric, $\|\tilde{\mathbf{r}}_{\mathrm{M3}}\|_1^2$ (Step 21).

6.3 Illustrative Example

We use a toy example of a 2×2 IRS array, shown in Fig. 6.2 to illustrate the basic concepts that our algorithm uses to increase effective SNR. By understanding the effects of dynamic grouping we can apply these principles to a larger scale in Algorithm 2.

In this example, we are given four real values that represent the four $r_{1,m}$ values of the beamformed and reflected channel. Each element's $r_{1,m}$ value is listed in its corresponding element block. Remember, $\mathbf{r}_1$ is the largest singular vector of $\mathbf{V}_\mathrm{r}$ that we use because of the optimal beamformer multiplication $\mathbf{q}^\mathrm{H}\mathbf{V}_\mathrm{r}$ found in (3.4). These values have both a magnitude and a direction, negative or positive, that represents a simplified version of complex phase.

The formulation for our method states that grouping the $r_{1,m}$ values into K groups (in this case $K = 2$) according to the metric in (6.5a) gives us exceptional (in this case optimal) effective SNR. In the default first grouping in Fig. 6.2 labeled "Fixed Grouping Method," we assume a set grouping with 2×1 element blocks to represent a fixed grouping similar to [4]. To measure the effective SNR for this method, we find $\sum_{i \in S_k} r_{1,i} \tilde{\phi}_k$ for each group k and then take the sum of those values. We find the effective SNR of the "Fixed Grouping Method" in Fig. 6.2 to show the process. Note that in both grouping methods it is apparent that applying the optimized group phase shifts to each element of the group is equivalent to taking the magnitude of the sum of each group's $r_{1,m}$ values, where $m \in S_k$. This is consistent with complex values and we use this as a simplification in our calculations for Algorithm 2 instead of calculating out the optimized $\tilde{\phi}_k$ every time.

In Fig. 6.2 the "Output Grouping from Algorithm 2" is the designed (in this case optimal) resultant grouping from utilizing Algorithm 2 with the "Fixed Grouping Method" as input. To confirm the improved performance of Algorithm 2, we calculate the effective SNR to compare these methods. As we expect, the effective SNR is maximized in the new grouping, which we can check by comparing the effective SNR to the upper bound in

$$\left| \sum_{m=1}^{M} r_{1,m} \phi_m^\star \right|^2 = \left| \sum_{m=1}^{M} |r_{1,m}| \right|^2 = 19^2.$$

The steps to change the "Fixed Grouping Method" to the dynamic grouping through Algorithm 2 are shown in Figs. 6.3 and 6.4. First, we take the fixed grouping as an initial grouping input to Algorithm 2, shown as $[S_{0,1}, S_{0,2}]$. We separate the process into 4 separate steps, for each step selecting one of the elements $1, 2, 3, 4$ and evaluate that element for potential group changes. According to Algorithm 2, if $\delta < \mu_{k_{\text{other}}}$ (where in this case k_{other} denotes the other group that does not contain the selected element), then the selected element is moved to the other group. The output of Algorithm 2, $[S_1^\star, S_2^\star]$, is equal to the $[S_1, S_2]$ after Step 4 in Fig. 6.4. Note, we achieve convergence after the first iteration (Steps 1-4

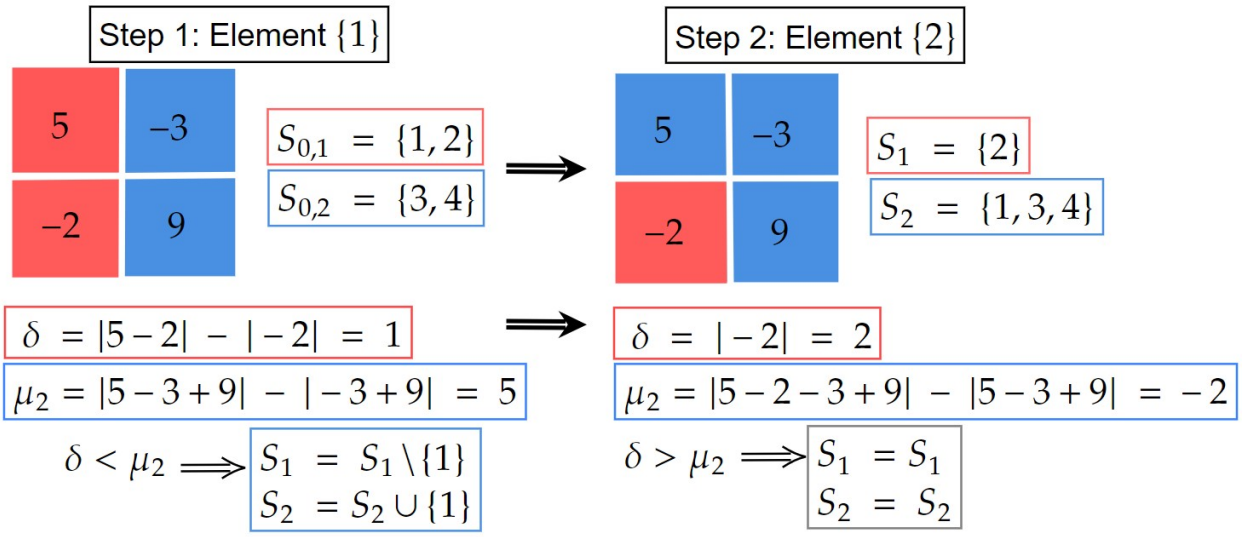

Figure 6.3: Illustrating Algorithm 2 using toy problem from Chapter 6.3. Steps 1-2.

in the figures), meaning the groups cannot be improved by further iterations. Typically, Algorithm 2 would repeat the process for another iteration to check convergence, but we have not included the subsequent iteration in additional figures because no changes to the groups occur.

6.4 Method 3 Initialization Algorithm

Any non-empty sets $\{S_{1,0}, S_{2,0}, \cdots, S_{K,0}\}$ with the property $\bigcup_{k=1}^{K} \mathrm{S_{k,0}} = \{1, 2, \cdots, M\}$ (where $\bigcup_{k=1}^{K} \mathrm{S_{k,0}}$ is a disjoint set) can be used as an initialization to Algorithm 2, however through testing we have determined that the improvement of this starting grouping affects not only the convergence of the algorithm, but it also improves the final effective SNR gain. This makes the initialization stage very important, and worthy of an algorithm of its own.

To prepare for Algorithm 3, we first formulate a method that brings Algorithm 2 closer to the upper bound in (6.3), i.e. where the elements are grouped in a way that minimizes their cancelled magnitude. Magnitude cancellation occurs when the magnitude of an element's $r_{1,m}$ values is out of phase with the other elements in its group. Therefore, the closer the phase values ($\angle r_{1,m}$) are within a group, the more we expect the magnitudes to work together

Algorithm 3 Initialize M3 Algorithm

Input: $\mathbf{r}_1, K, M$

Output: $\{S_{0,1}, S_{0,2}, \cdots, S_{0,K}\}$

$\boldsymbol{\theta} = \angle \mathbf{r}_1$
for $m = 1 : M$ **do**
  if $m = M$ **then**
    $\delta_{1 \leftrightarrow m} = \theta_1 - \theta_m$
  else
    $\delta_{(m+1) \leftrightarrow m} = \theta_{m+1} - \theta_m$
  end if
end for
$\mathcal{P} \leftarrow \{1 : M\}$
for $k = 1 : K$ **do**
  $x_k = \text{argmax}_{m \in \mathcal{P}} \, \delta_{(m+1) \leftrightarrow m}$ {Note: $\mathbf{x} = [x_1, x_2, \cdots, x_K]^{\text{T}}$}
  $\mathcal{P} \leftarrow \mathcal{P} \backslash \{x_k\}$
end for
$\mathbf{x} = \text{sort}\,(\mathbf{x})$ {Note: Ascending order}
for $k = 1 : K$ **do**
  if $k = K$ **then**
    $S_k \leftarrow \{x_1 : x_{k+1}\}$
  else
    $S_k \leftarrow \{(x_k + 1) : x_{k+1}\}$
  end if
end for
for $k_{\text{sel}} = 1 : K$ **do**
  $S'_{k_{\text{sel}}} \leftarrow \text{argmin}_{m_{\text{sel}} \in S_{k_{\text{sel}}}} |r_{1, m_{\text{sel}}}|$
end for
for $m = \{1 : M\} \setminus \left\{ \bigcup_{k=1}^{K} S'_k \right\}$ **do**
  for $k = 1 : K$ **do**
    $\mu_k = f_1\left(S'_k \cup \{m\}\right) - f_1\left(S'_k\right)$
  end for
  $k_{\text{best}} = \text{argmax}_i \, \mu_i$
  $S'_{k_{\text{best}}} \leftarrow S'_{k_{\text{best}}} \cup \{m\}$
end for
$\{S_{1,0}, S_{2,0}, \cdots, S_{K,0}\} \leftarrow \{S'_1, S'_2, \cdots, S'_K\}$

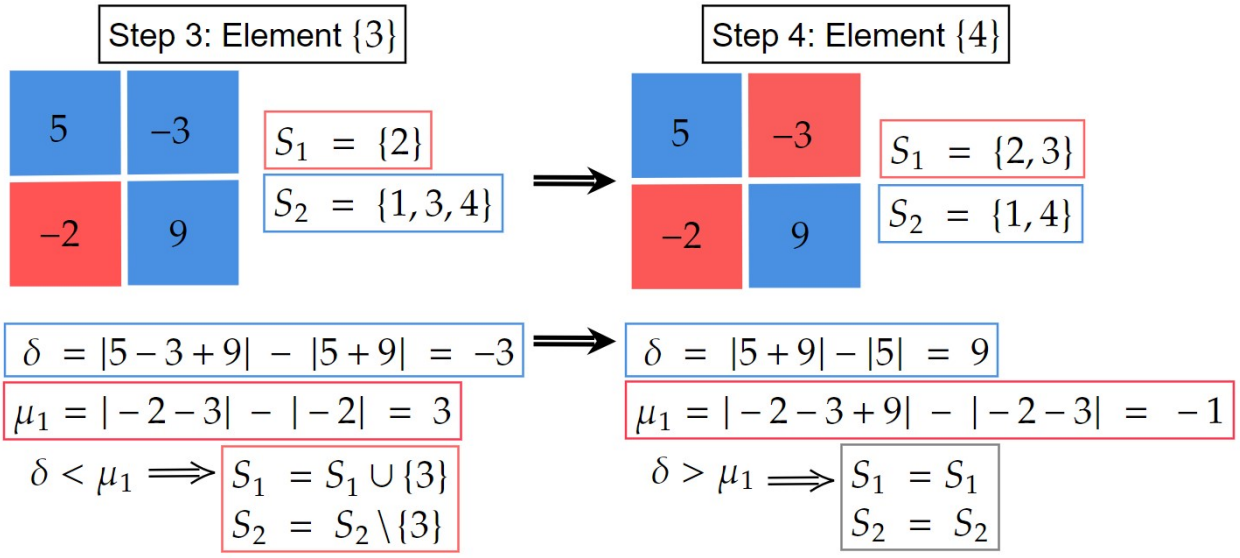

Figure 6.4: Illustrating Algorithm 2 using toy problem from Chapter 6.3. Steps 3-4.

towards effective SNR gain. The proposed method to mitigate magnitude cancellation begins by plotting the complex phase of each element's $r_{1,m}$ value, and then finding the largest K gaps between the complex phases, where $\delta_{(m+1)\leftrightarrow m}$ denotes the phase gap between element m and element $m+1$. We then define $\mathcal{P}$ as the set of all elements $1:M$. Next, the K largest complex phase gaps (1 for every selected group) are identified and the element corresponding to the beginning of the gap is stored as the variable x_k, where $\mathbf{x} = [x_1, x_2, \cdots, x_K]^{\mathrm{T}}$, and then x_k is removed from the set $\mathcal{P}$. The K element indices contained in $\mathbf{x}$ are then sorted from smallest to largest via sort($\mathbf{x}$) to be implemented in ascending order. This procedure is illustrated in Fig. 6.5 and is contained in Algorithm 3, Steps 1-14. Next, we insert K barriers in between these elements at the largest K phase gap values, δ as in Fig. 6.6. These barriers denote the boundaries for each group and correspond to Steps 15-21, where S_k is the set of elements within the group k. Thus we have created K intermediate groups where the phase values are all in-phase as much as possible, but we are not done. Within each of these groups, we examine the magnitudes of each element's $r_{1,m}$ value and then select the elements, m_{sel} with the K smallest magnitudes. At this point, all other elements from the groups are removed leaving one IRS element per group in our new set of groups S'_k (Steps 22-24). Then, the metric measurement method similar to Algorithm 2 is used to re-distribute

the rest of the remaining elements into groups (Steps 25-31), until we have the full initialized sets $\{S_{1,0}, S_{2,0}, \cdots, S_{K,0}\}$, where $\bigcup_{k=1}^{K} \mathrm{S_{k,0}} = \{1, 2, \cdots, M\}$ ($\bigcup_{k=1}^{K} \mathrm{S_{k,0}}$ is a disjoint set).

The additional re-distribution in Algorithm 3 is needed after the intermediate grouping stage because although the $r_{1,m}$ values are likely to have coherent phases at this stage, the grouping themselves are not ideal. For example, instead of allocating a single element to its own group as in Fig. 3, it may be better to combine two nearby groups and open up that partitioning to use elsewhere. The reason we select the smallest magnitude value instead of the largest magnitude to define the initial point within each group is that we would prefer to distribute the largest magnitude values according their effective SNR gain improvement, as they would have a larger effect than the smallest magnitude values.

Thus, we have proposed that because our optimization problem in (6.5a) is not convex in nature, our initial groupings have a significant effect on convergence and final effective SNR. In this sub-chapter, we proposed one potential solution to close the gap between the algorithm's effective SNR gain performance and the upper bound found in (6.3). Next, in Chapter 6.5 we test this assertion and compare the speeds of convergence with dedicated simulations.

6.5 Algorithm Performance Analysis

In order to test the performance of Algorithm 3, we simulated a channel,$\mathbf{V}_{\mathrm{r}}$, with $N = 4$ (2×2 BS) and $M = 36$ (6×6 IRS) modeled from Chapter 2.4 and set $\mathbf{r}_1$ as its largest right singular vector. with a random initial partitioning input as a baseline performance gauge to the initialization algorithm. We ran 500 channel realizations to determine average values.

As seen in Fig. 6.7, the initial grouping state design allows for greater effective SNR output, as well as earlier convergence to the upper bound of Method 3 derived in (6.3). Another important feature that saves on computational costs is the speed (number of iterations) at which the algorithm converges. Especially when working with massive IRSs in the future, individually running M allocation assignments every iteration becomes costly. In Fig. 6.8

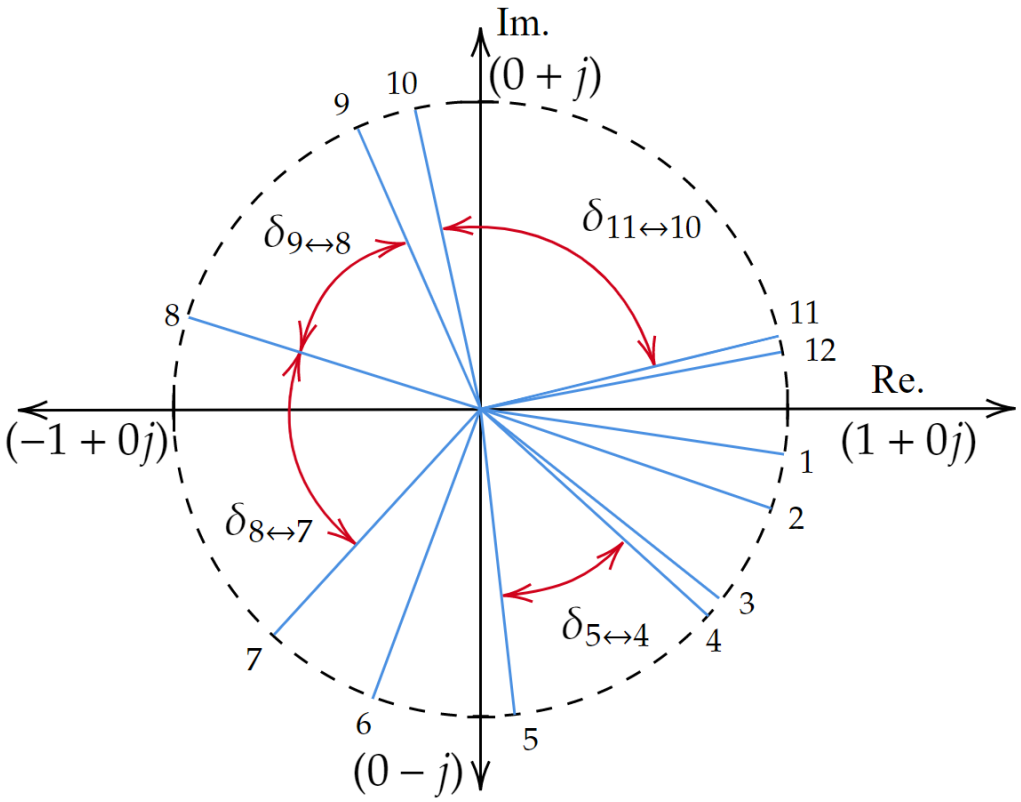

Figure 6.5: Illustrative representation of the initialization algorithm for Method 3, Algorithm 3 steps 1-10. Here the complex phases of $r_{1,m}$ for $M = 12$ are represented as blue lines with magnitude 1. In this case, $K = 4$, which corresponds to K separate δ values corresponding to the 4 largest gaps between phases.

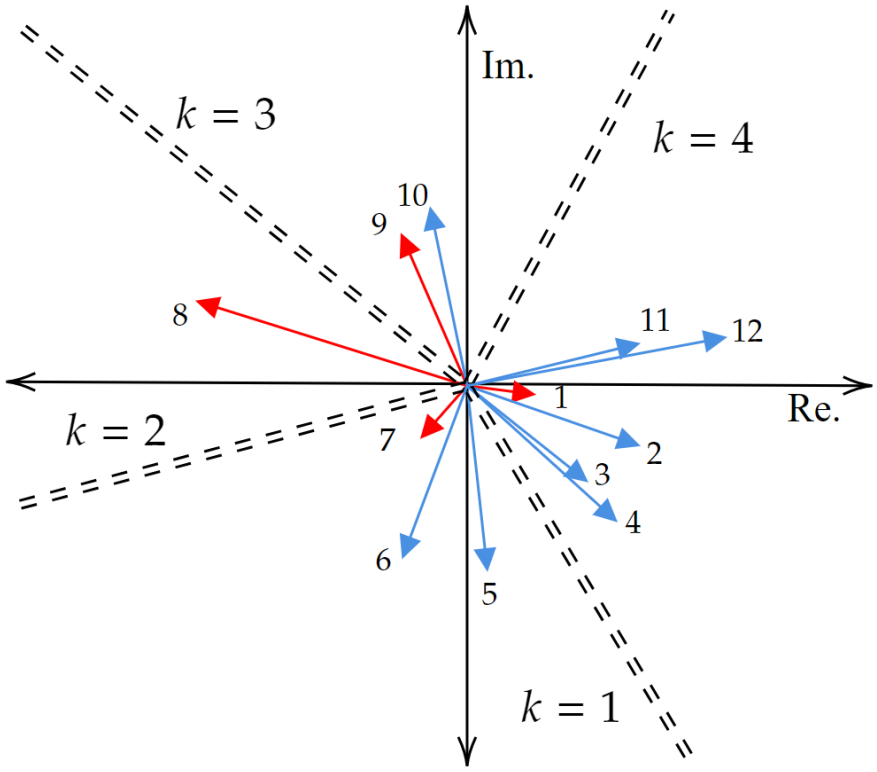

Figure 6.6: Illustrative representation of the initialization algorithm for Method 3, Algorithm 3 Steps 11-16. Here the magnitudes of $r_{1,m}$ are included as well as the complex phases. In this case, $K = 4$, where the grouping boundaries (shown in the figure as dashed lines) replace the δ locations in Fig. 6.5. The highlighted red values are the values of $r_{1,m}$ in each group k with the smallest magnitude.

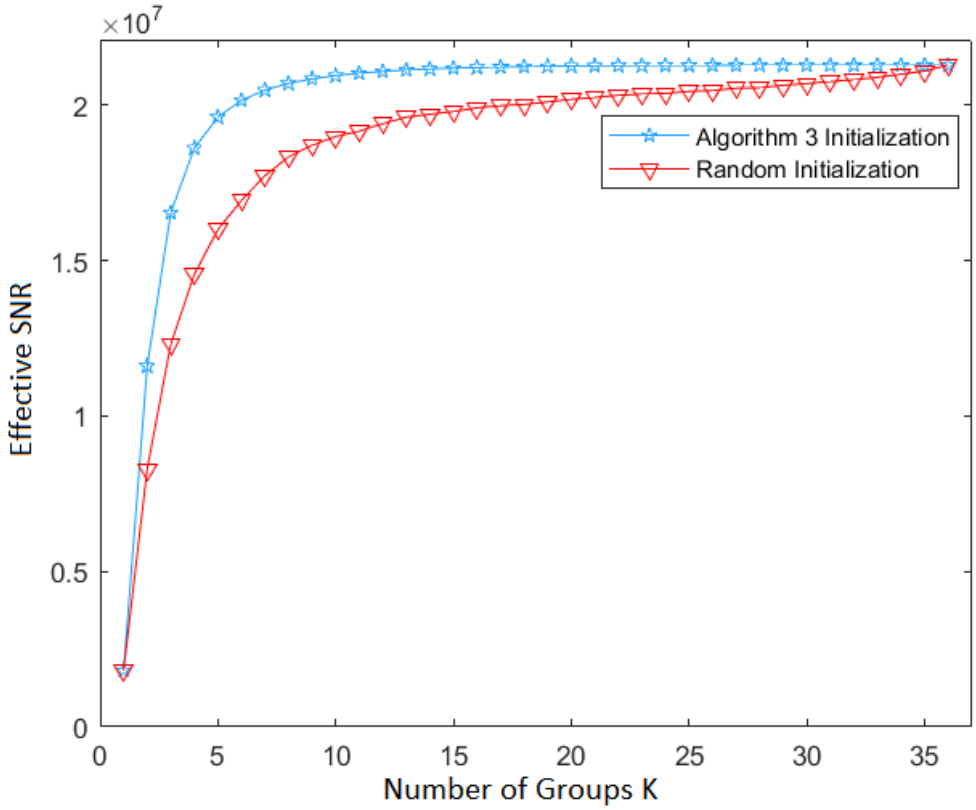

Figure 6.7: Comparing the effective SNR output of Algorithm 2 with and without the use of designed initial groups from Algorithm 3. Channel generated with $N = 4$ and $M = 36$.

we see that Algorithm 3 reduces the average number of iterations until convergence by about a factor of two.

6.6 Overhead Analysis

Since this grouping method requires an additional feedback transmission indicating the group that each IRS element belongs to, the number of required feedback bits increases from (5.7). With proper coding, we can communicate the location of each IRS element with a string of M coded bit strings of size $\lceil \log_2(M\rho) \rceil$ which increases the overhead as seen in (6.8). With given ρ, our achievable rate equation for Method 3 is given by:

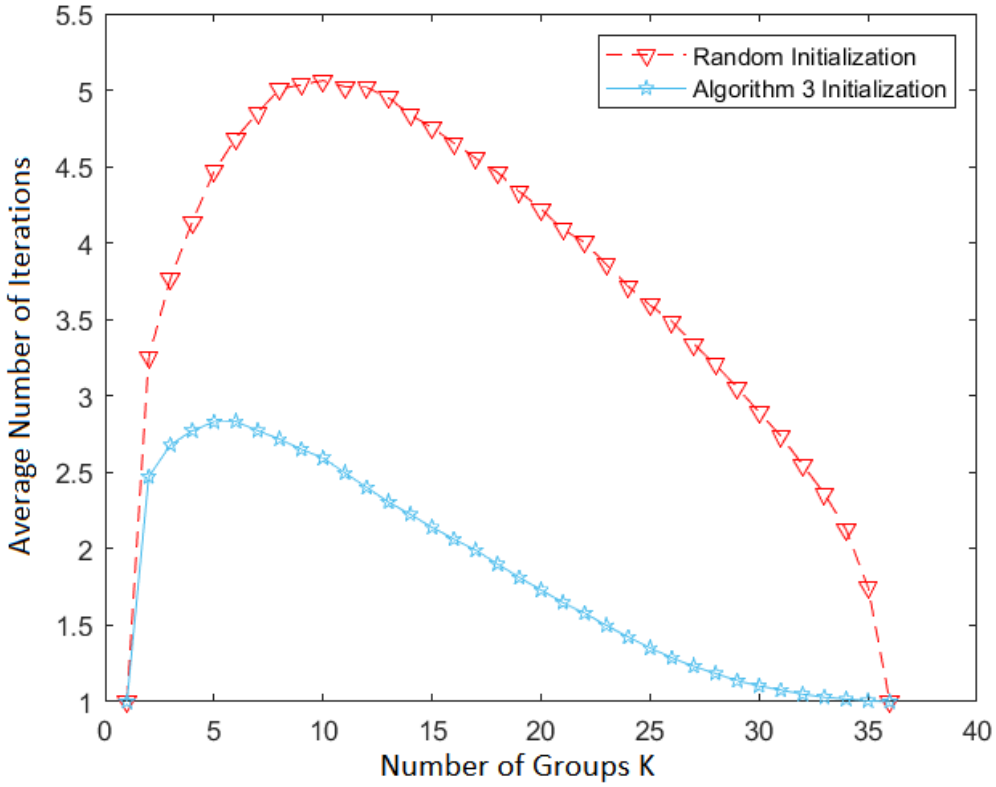

Figure 6.8: Comparing the average number of iterations until convergence of Algorithm 2 with and without the use of designed initial groups from Algorithm 3.

$$R = \left(1 - \frac{(NM+1)\,T_0}{T_{\text{tot}}} - \frac{M\rho b_{\text{F}} + \lceil \log_2(M\rho) \rceil M}{T_{\text{tot}} B_{\text{F}} \log_2\left(1 + \frac{p_{\text{F}}|h_{\text{F}}|^2}{N_0 B_{\text{F}}}\right)}\right) B \log_2\left(1 + \frac{p\left|\mathbf{q}^{\text{H}}\mathbf{V}_r \boldsymbol{\phi}^{\star}_{\text{M3}}\right|^2}{BN_0}\right), \quad (6.8)$$

where, like Method 1, the pre-log overhead term is monotonically increasing with ρ. Due to the behavior of overhead with respect to ρ and the fact that the output of our proposed algorithm approaches the upper bound on performance relatively quickly (as seen in Fig. 6.7), it seems that this method can likely take advantage of the lower number of groups to make the effective T_{F} lower than expected. This gives Method 3 a lot of potential to find the right trade-off between effective SNR gain and overhead costs to outperform the other methods. We investigate the comparative performance with simulations in Chapter 7 to observe whether the increase in effective SNR gain is worth the extra cost of overhead.

Chapter 7

Simulations

Abstract

In this chapter we introduce simulation setup to validate the performance of our proposed methods, including additions to our channel model, and necessary parameters that have to be chosen. In this chapter, we also observe the simulations and discuss the results of the proposed grouping schemes and reflection coefficient maximization problems in Chapters 4-6.

7.1 Simulation Setup

Before we present the simulations and results, it is necessary to describe the parameters used to perform the simulations in our models. The parameters in Table 7.1 are used for the simulations and are included here for reference, where $P_{\max}$ is the maximum power available to be allocated between p and p_{F} (i.e. $P_{\max} = p + p_{\mathrm{F}}$), and $B_{\max}$ is the maximum bandwidth available to be allocated between B and B_{F} (i.e. $B_{\max} = B + B_{\mathrm{F}}$). Note that in the simulations, the maximum achievable rate results are normalized by $B_{\max}$ with the associated units of bps/Hz.

T_{tot}	T_0	b_{F}	$P_{\max}$	$B_{\max}$	N_0
1.525 ms	$T_{\mathrm{tot}}10^{-5}$	16 bits	25 dBm	10 MHz	-75 dBm/Hz

Table 7.1: Parameters for Simulations

7.1.1 Channel Clustering

In our simulations, we consider the URA channel model in Section 2.1. The channel is modified to resemble a clustered channel model adapted from [31], where each cluster is made up of multiple subrays. Each mention of N_{ch} is replaced by $N_{\text{cluster}}N_{\text{subray}}$ whereby every azimuth and elevation angle in the previous model is treated as a cluster angle, and each has N_{subray} angles reflecting from the cluster. In this work we use $N_{\text{cluster}} = 8$ and $N_{\text{subray}} = 10$ and the angles are distributed as in [31]. The cluster angles are distributed as if they are the N_{ch} angle pairs in Chapter 2.4. The main difference in angle distribution when comparing the cluster angles to the subray angles is that the subray angles are uniformly distributed with angular spread of 5°, where the center angles of each subray cluster are defined by the angle pair corresponding to each N_{cluster}. This model, in effect, makes the channel more correlated as we see in a realistic practical setting and lends itself to methods of optimization such as ours that allocate elements based on channel information.

7.1.2 Grouping Ratio Optimization

We have established in this thesis that the IGR $\rho^\star$ can be set before data transmission unlike the existing grouping method in [4]. Therefore in our simulations we implement the IGR at each channel realization to show the potential ideal performance of our designed methods. In testing, we found that the achievable rate maximization problem with respect to ρ is not strictly convex for each channel realization, therefore it cannot be explicitly solved. Therefore, in our simulations we simply perform an exhaustive search over all possible values of ρ to find the IGR. In Methods 1 and 2 and the existing grouping method in [4] we use the set of $\rho_\kappa = \frac{\kappa}{M}, \forall \kappa \in \text{divisors}(M)$ and Method 3 searches over all values of $\rho_m = \frac{m}{M}, \forall m$.

7.1.3 Bandwidth and Power

The resources $B, B_{\mathrm{F}}, p,$ and p_{F} are allocated according to the methodologies introduced in [2]. Therein the resource allocation solves the optimization problem:

$$\max_{p,p_{\mathrm{F}},B,B_{\mathrm{F}}} R \tag{7.1a}$$

$$s.t.\ p + p_{\mathrm{F}} \leq P_{\max}, \quad B + B_{\mathrm{F}} \leq B_{\max} \tag{7.1b}$$

$$p \geq 0, \quad p_{\mathrm{F}} \geq 0, \quad B \geq 0, \quad B_{\mathrm{F}} \geq 0 \tag{7.1c}$$

$$T_{\mathrm{F}} \leq 1 - \frac{T_{\mathrm{E}}}{T_{\mathrm{tot}}}. \tag{7.1d}$$

While this problem is not jointly convex, it was reformulated to be jointly convex by taking the logarithm of R and through the following linear transformations (which preserve convexity): $p_{\mathrm{F}} = P_{\max} - p$ and $B_{\mathrm{F}} = B_{\max} - B$. These steps come together in the following maximization problem

$$\max_{p,B} \log R \tag{7.2a}$$

$$s.t.\ 0 \leq p \leq P_{\max} - p_{\mathrm{F}}, \quad 0 \leq B \leq B_{\max} - B_{\mathrm{F}} \tag{7.2b}$$

$$T_{\mathrm{F}} \leq 1 - \frac{T_E}{T_{\mathrm{tot}}}. \tag{7.2c}$$

From this problem they derive two explicit convex solutions, one with respect to B when p is fixed and the other solution with respect to p when B is fixed. In this work we utilize these solutions for our optimized B and p at each channel realization.

7.2 Benchmarks

The first simulation, found in Fig. 7.1, compares the resulting effective SNR from the three proposed methods in this thesis along with the method in [4] as a reference baseline. The upper bound plot in this figure represents the maximum achievable SNR based on the

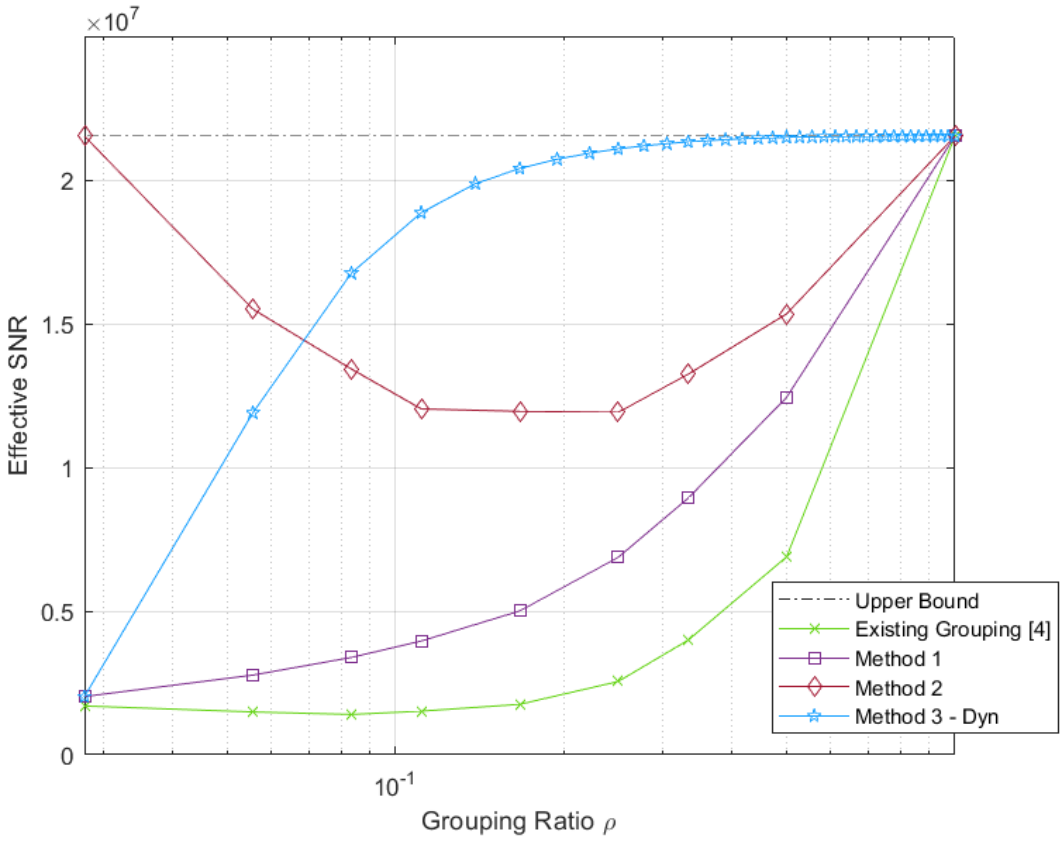

Figure 7.1: Comparison of the 3 Proposed Methods in terms of effective SNR, with the existing grouping from [4] as a baseline and the size parameters: $M = 36$ and $N = 4$.

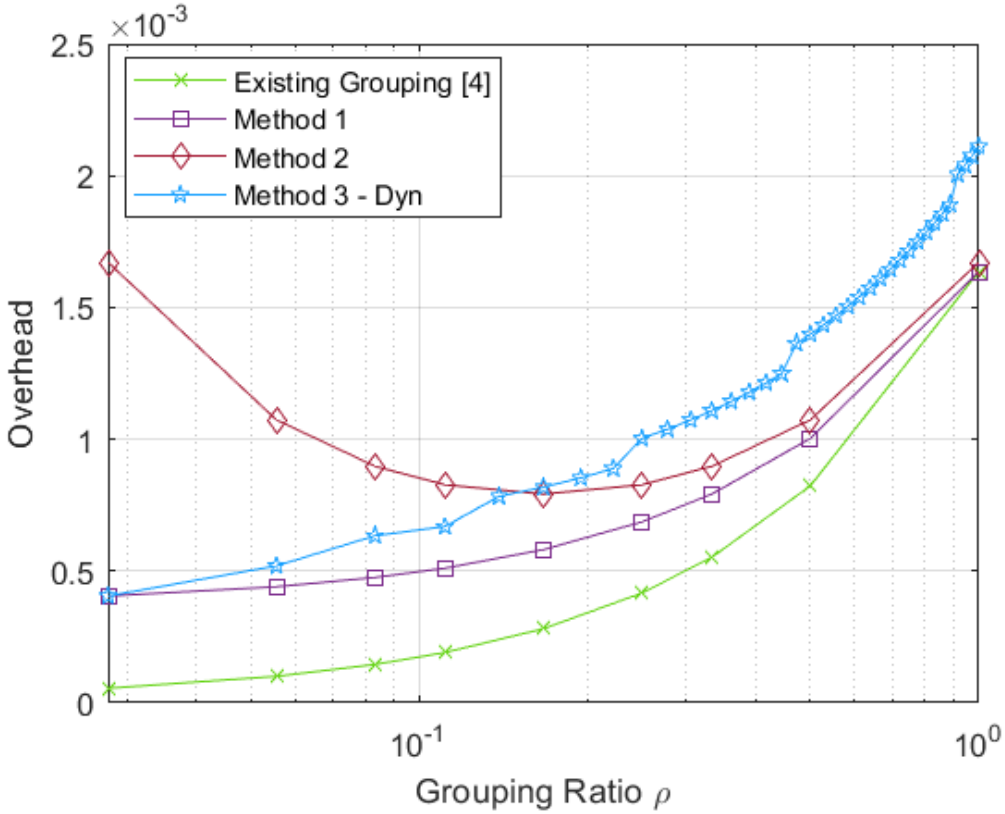

Figure 7.2: Comparison of the 3 Proposed Methods in terms of overhead with fixed resource allocation, with the existing grouping method as a baseline and $M = 36$
Note: The stair-stepping nature of the Method 3 curve is a result of the $\lceil \log_2(M\rho) \rceil M$ term from (6.8) that causes uneven breakpoints at each ρ.

bound in (3.5) and for most methods this is only achievable when $K = M$. Based on our formulations, we expect that the effective SNR improves for each successive method, as the purpose for creating more freedom of optimization is to get closer to this upper bound in (3.5) with smaller ρ, and this figure confirms that hypothesis. Specifically, Method 3 clearly shows the best SNR performance and converges to the upper bound the fastest among all methods with respect to the grouping ratio ρ. Note the symmetry of the effective SNR curve for Method 2 with respect to the center point $\rho = \frac{6}{36}$. As described in Chapter 5.3, this is due to the reciprocal nature of the two vector optimization. We see that the effective SNR of Method 2 is always at a minimum at the center point defined as $\rho_c = \frac{\sqrt{M}}{M}$.

In Fig. 7.2 we see that the overhead is monotonically increasing with grouping ratio ρ for

all but one of the curves. Note that though Method 2 is the exception to this observation, we see that the overhead plot shows the same symmetry as in Fig. 7.1, and the overhead is monotonically increasing after the center point ρ_c. Therefore, for Method 2, we can essentially define the range for the implementable grouping ratio as $\rho_c \leq \rho \leq 1$, because any grouping ratio lower than ρ_c can be described by a corresponding ρ with the same effective SNR and overhead on the opposite side of ρ_c.

In a general sense, effective SNR gain with a smaller ρ is more valuable in practice because it corresponds to a lower overhead cost. That is to say given two values of effective SNR that are identical within a method, the value at a lower ρ is preferred, because it follows that the achievable rate performance is improved with smaller overhead (with the exception of the symmetry exhibited in Method 2). Method 3 clearly demonstrates this observation as its superior achievable rate performance is a product of its rapid convergence to the effective SNR upper bound with respect to ρ. Because Method 3 has the largest relative overhead costs with respect to ρ we should expect the IGR chosen for this method to be significantly smaller than the other methods thus a lower effective overhead.

Additionally, Method 3 has the ability to set $K \in \{1 : M\}$ unlike other methods, this allows for better selectivity in selecting the precise IGR. In an extreme example, Fig. 7.3 shows the effective SNR curves with respect to ρ when $M = 25$. However, in this case, the set of $\rho \in \left\{\frac{1}{25}, \frac{5}{25}, \frac{25}{25}\right\}$ only has 3 values for the other methods to choose from for their IGR. As Method 3 has 25 values of ρ to choose from for its IGR it is able to take advantage of the ideal trade-off where other methods cannot. This introduces limitations on IRS construction for systems using other grouping methods, which is a problem that has not been investigated in previous works. Due to the integer limitation on Z, most grouping methods require specially-sized IRSs that allow for the most divisions corresponding to ρ points. This is something to consider when constructing IRS arrays. However, with Method 3, any size and shaped IRS array can be dynamically grouped, giving more reasons why our method is useful.

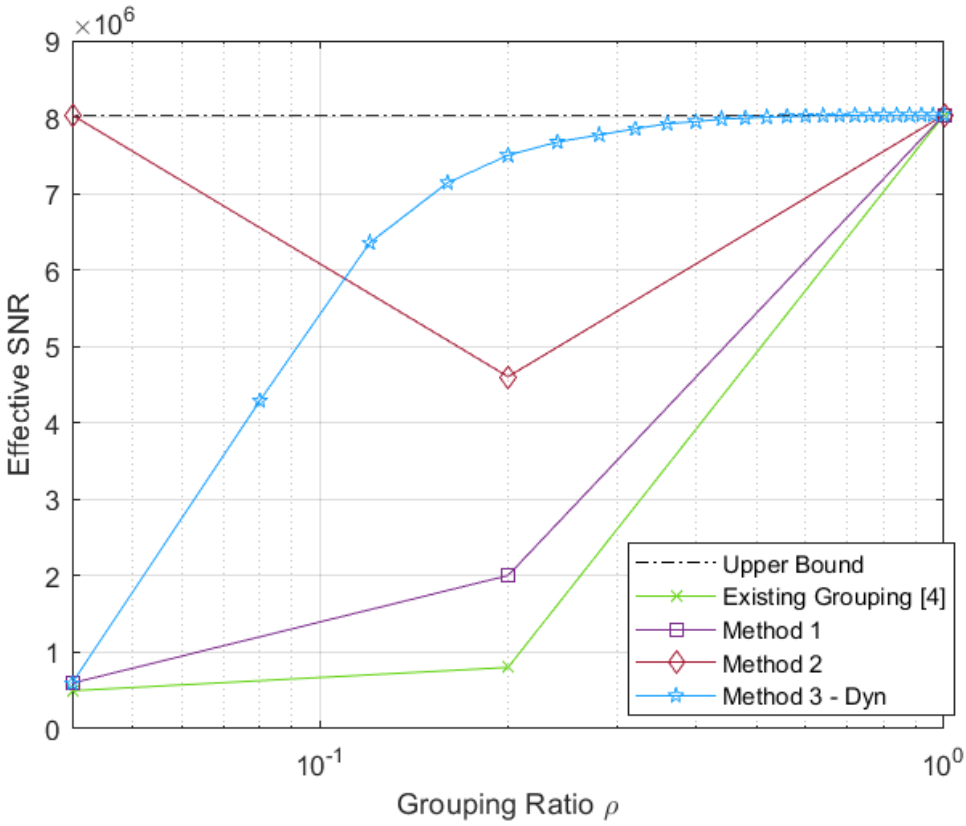

Figure 7.3: Comparison of the 3 Proposed Methods in terms of effective SNR, with the existing grouping from [4] as a baseline and the size parameters: $M = 25$ and $N = 4$.

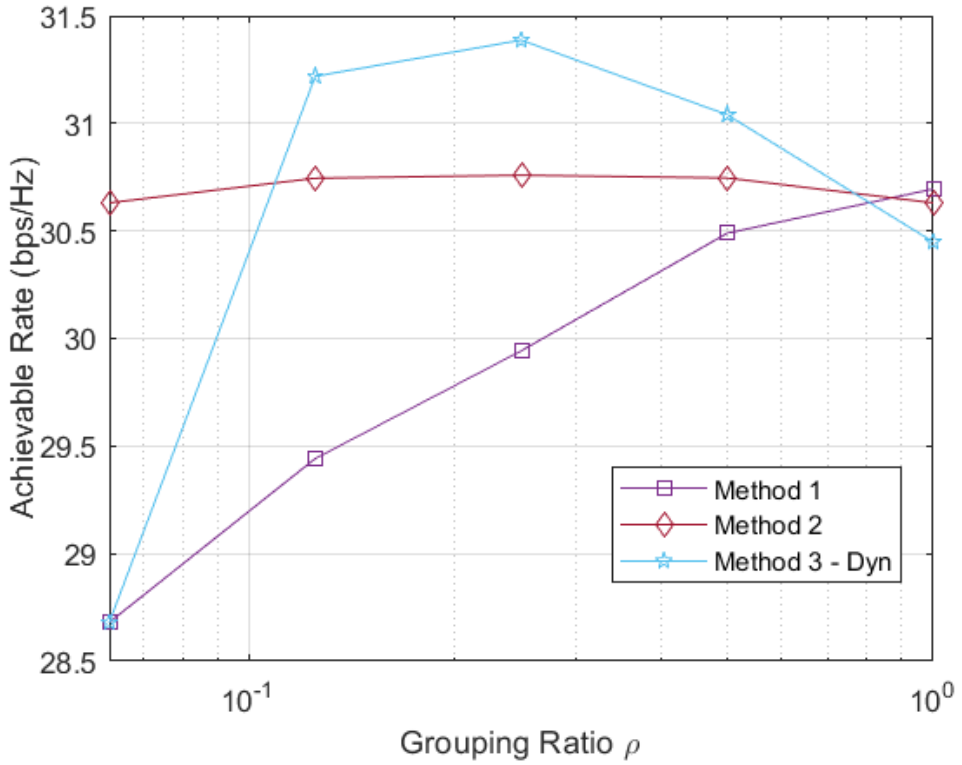

Figure 7.4: Rate with respect to Grouping Ratio, $M = 16$.

7.3 Convergence of Ideal Grouping Ratio

This work is the first of its kind to fully investigate the effect of the IGR on maximum achievable performance and convergence, because before our formulation, the IGR was unattainable. In the previous grouping work [4], the achievable rate with respect to the grouping ratio was shown to be concave when averaged over many channel realizations. To analyze the effects of a variable IRS size on the performance trade-off with respect to grouping ratio we present Figs. 7.4-7.6 which show the achievable rate with respect to grouping ratio for $M = 16$, 64, and 256. When averaged over 500 channel realizations we see the concave behavior of the curves and that each method has a single IGR that varies with M. After contrasting Figs. 7.4-7.6, it is clear that after increasing M, the IGR for all curves decreases and the curves themselves become steeper. This indicates that with large M, achieving the IGR is more important for achievable rate performance than in small M scenarios, because

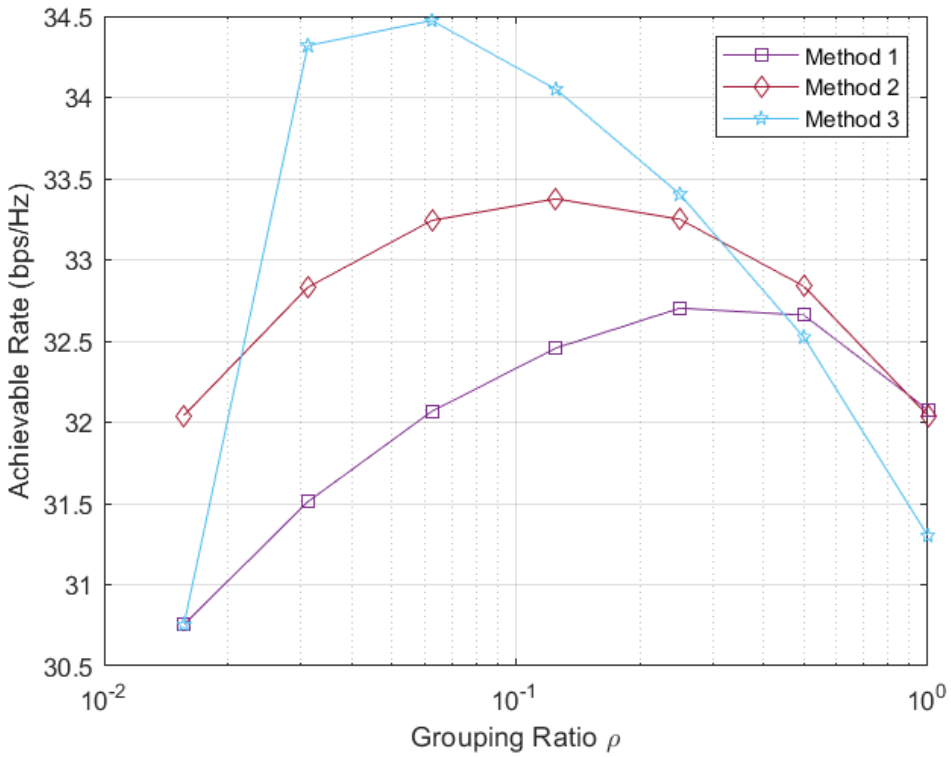

Figure 7.5: Rate with respect to Grouping Ratio, $M = 64$.

there is a bigger gap between the achievable rate of the IGR and the adjacent achievable rate values. This has implications for massive MIMO and for massive IRS system integration, as M increases, ideal grouping becomes paramount in mitigating overhead.

We observed in Figs. 7.4-7.6 that the IGR decreased as M increased. This seems to indicate some convergent behavior of the IGR, which we investigate further in the next simulations. In Fig. 7.7, we simulate the selected IGR of various grouping methods (averaged over 500 channel realizations) as M increases. In this figure, it is clear that Method 3 groups the most efficiently compared to the other methods, but as M increases, the denominator of ρ increases, so we also simulate the ideal number of groups $M\rho^{\star}$ to investigate its convergent behavior as well. In Fig. 7.8 we isolate the grouping from the increasing M to identify convergent behaviors. We observe that Methods 1 and 3 converge to a certain value of K as M increases, and the other two methods continue to change as M increases. Notably, Method 2 converges to the ideal K where $K = \sqrt{M}$. This follows from Figs. 7.4-7.6 where

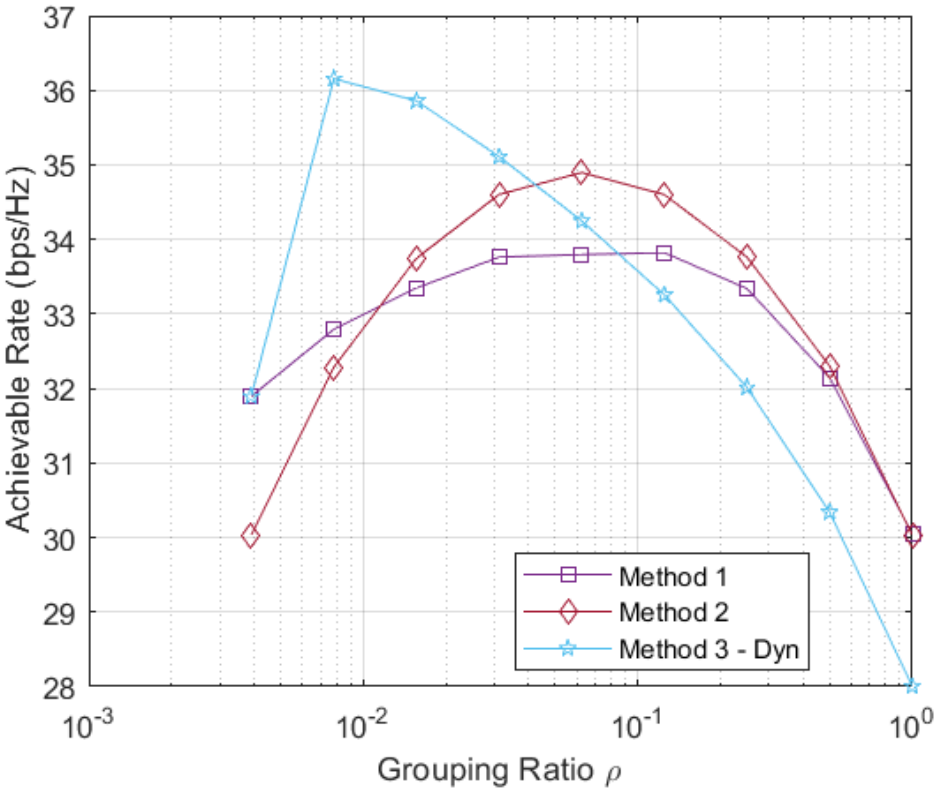

Figure 7.6: Rate with respect to Grouping Ratio, $M = 256$.

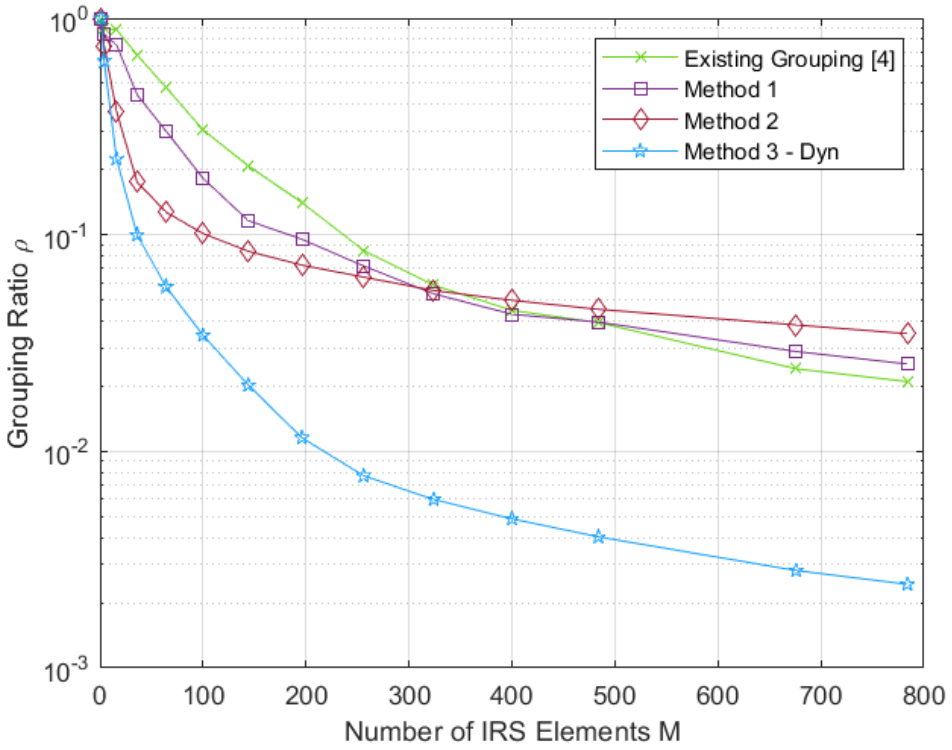

Figure 7.7: Convergence of the IGR for different methods, $N = 9$.

the ρ_c was always the point with the highest achievable rate. This is also a hindrance when we consider massive MIMO and massive IRS integrated systems, as the minimum overhead is limited by the overhead at the center point ρ_c.

The ideal K curve for Method 3 in Fig. 7.8 also has important implications for massive MIMO and massive IRS integrated systems. We observe that the ideal K for Method 3 converges to $K = 2$. This efficient grouping is needed to combat the large overhead associated with this method. However, with very large M and N, the convergence to $K = 2$ is the simplest case of grouping and has the benefit of lower complexity partitioning at the BS. This implies that the best way to improve Method 3 is to further optimize its algorithm to converge faster to the upper bound, leading to a larger achievable rate, which we investigate in the next simulation.

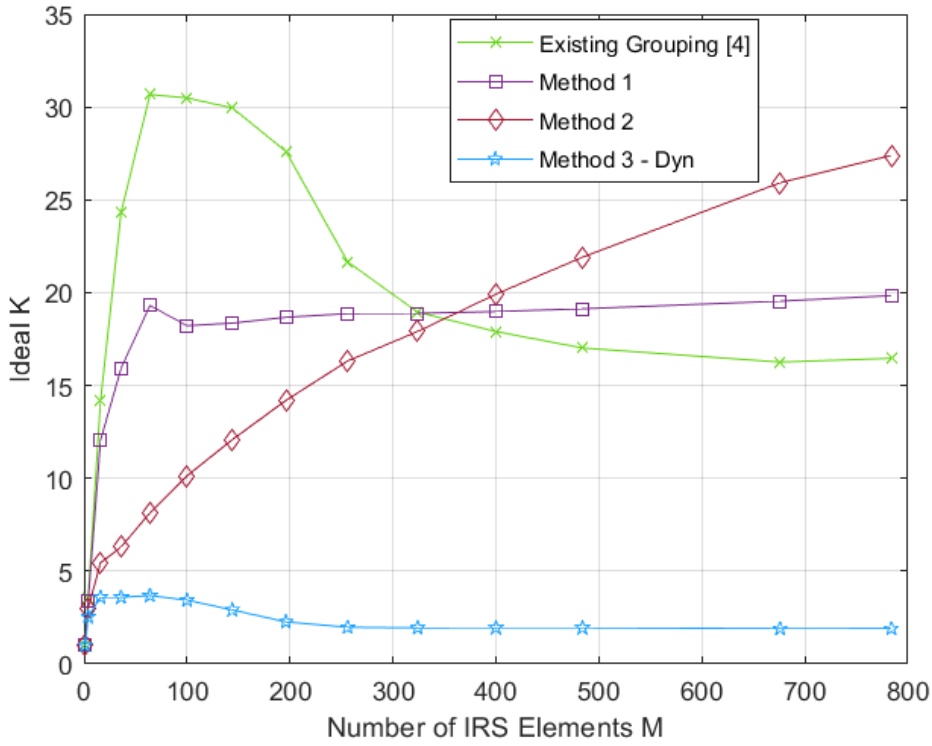

Figure 7.8: Convergence of the ideal number of groups K for different methods, $N = 9$.

7.4 Achievable Rate

The maximum achievable rate curves in Fig. 7.9 are arguably the most important to discuss for this thesis because maximizing achievable rate is one of the main motivations for our proposed grouping scheme. Several baseline curves are added to this simulation to better understand the relative performance of the proposed methods. The Reflecting Surface (RS) Mirror curve is simulating the performance of a non-intelligent reflecting surface. That is to say the surface is perfectly reflecting without intelligently directing its phase shifts, i.e. $\boldsymbol{\phi} = \mathbf{1}_M$. The RS in this case is acting as a signal "mirror" where any signal reflects according to Snell's Law [5]. Because the RS mirror does not require feedback to implement phase shifts and only needs to estimate the $\mathbf{v}_\phi$ channel from (2.3), the overhead is greatly reduced in this framework. For this simulation we redefine the overhead parameters from (2.7) to fit this model. We say that $T_{\mathrm{E}} = NT_0$, meaning that only N pilot tones are required

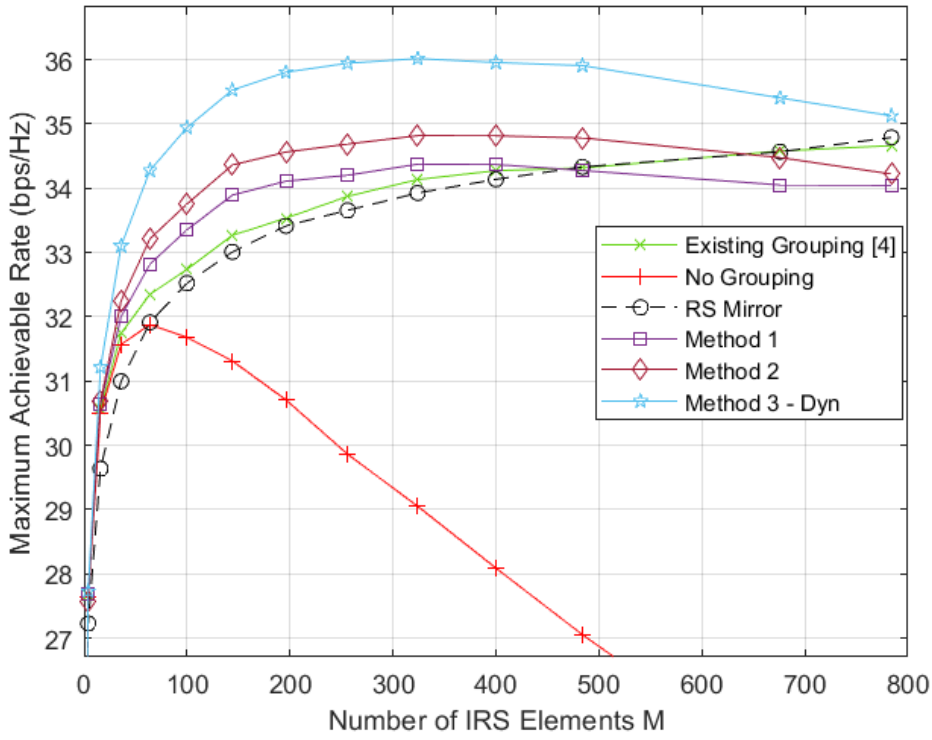

Figure 7.9: Maximum achievable rate in bps/Hz for proposed grouping methods and baseline curves, $N = 9$.

for channel estimation and $T_{\mathrm{F}} = 0$ because feedback is not required in this method. The other baseline curve introduced in Fig. 7.9 is the "No Grouping" method. This is the method investigated in [2] and is oblivious to grouping in general. It is used to show how much of a difference the proposed methods can make on the final maximum achievable rate compared to a default method that suffers from extreme overhead loss. The "No Grouping" method uses the following overhead parameters in its overhead model

$$T_{\mathrm{E}} = T_0 \left(NM + 1\right) \quad \text{and} \quad T_{\mathrm{F}} = \frac{b_{\mathrm{F}} M}{B_{\mathrm{F}} \log_2 \left(1 + \frac{p_{\mathrm{F}} |h_{\mathrm{F}}|^2}{B_{\mathrm{F}} N_0}\right)},$$

where T_{E} is the same as in the proposed methods and T_{F} represents the bits for a full phase shift sent for every IRS element M. Therefore, what these curves attempt to show are the two extremes of complexity and overhead when it comes to IRS communication. The "RS Mirror" method shows what is possible without feedback or IRS phase optimization and the "No Grouping" method shows what happens when overhead mitigation is not implemented at all.

As seen in Fig. 7.9, the IRS implementation without grouping has poor achievable rate performance as M increases compared to the other methods, which confirms the work of [2] and the need for grouping as a way to mitigate overhead for massive MIMO/massive IRS systems. The "Existing Grouping" method from [4] has the benefit of the lowest overhead of grouping methods, so it is always increasing as M increases, but only slightly outperforms the "RS Mirror" curve. Additionally, it is still not immune to overhead degradation as it performs worse than the "RS Mirror" curve as M becomes very large. All of the proposed methods have a peak performance number of IRS elements for its maximum achievable rate before they succumb to overhead costs. In this simulation, that point is $M = 324$. Interestingly, this peak M is shared among all proposed methods. This indicates that the ideal M for these proposed methods depends more on the number of antennas at the BS and the channel parameters than the method simulated.

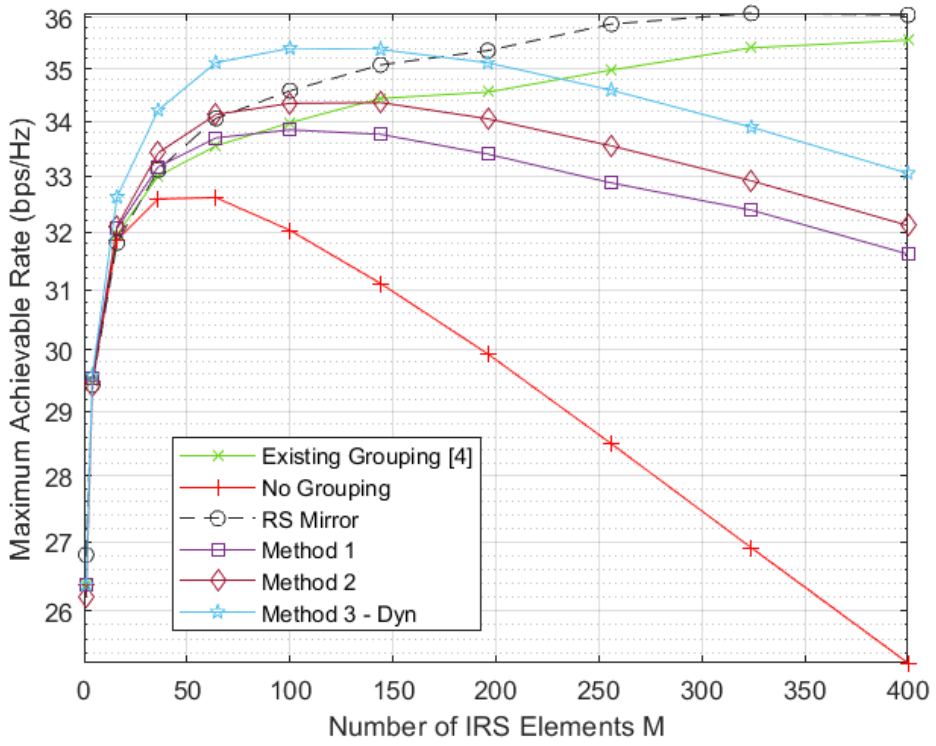

Figure 7.10: Maximum achievable rate in bps/Hz for proposed grouping methods and baseline curves, $N = 36$.

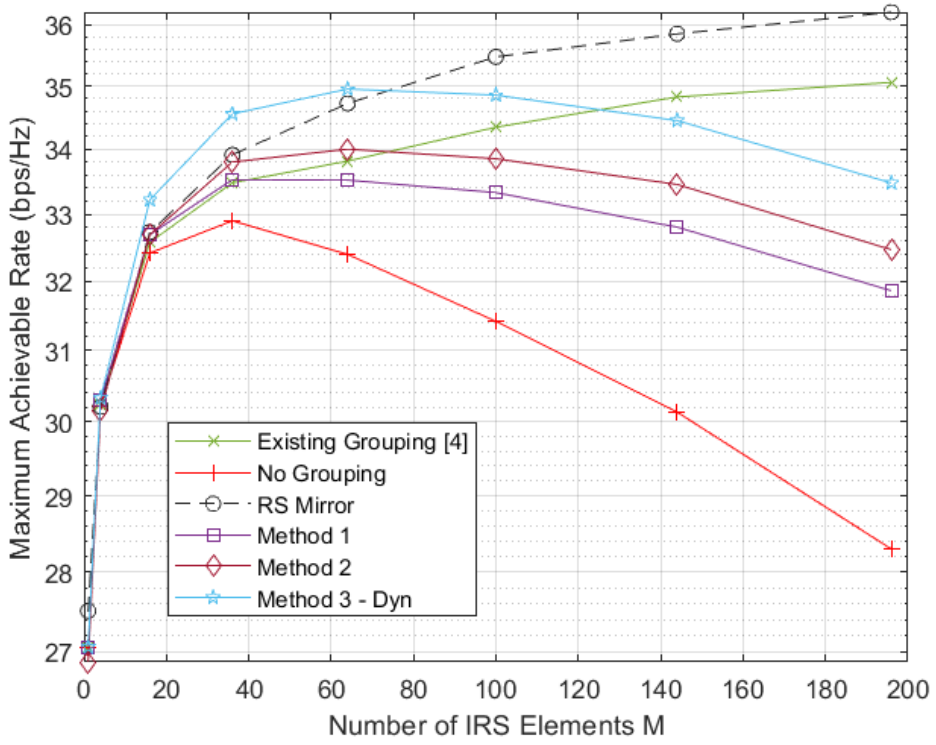

Figure 7.11: Maximum achievable rate in bps/Hz for proposed grouping methods and baseline curves, $N = 64$.

To illustrate scenarios with large N, i.e. approaching massive MIMO, we investigate the performance of $N = 36$ and $N = 64$ in Figs. 7.10-7.11. In these figures, we see that the proposed Method 3 is still the best performing method in terms of maximum achievable rate. The major downside of this method that hinders its performance at large M and N is the fixed T_{E} overhead set by the channel estimation. As we increase N, every method, including Method 3, eventually gives way to overwhelming overhead costs so that RS Mirror (which requires the least amount of overhead) becomes the best performing method. Since this thesis only considers the simple case of LS channel estimation, there is certainly room for improvement in practice such as implementing more efficient state-of-art IRS channel estimation techniques such as those proposed in [20] and [29] in order to reduce T_{E}. As overhead cost decreases, numerical results indicate that the proposed Method 3 is effectively implementable in larger and larger massive MIMO or massive IRS systems for grouping and reflection coefficient optimization.

In our final simulation in Fig. 7.12 we plot the effective SNR of all of the same curves in Fig. 7.9. We have already established that Method 3 is the best performing method in terms of effective SNR, but this figure shows that there is still room to be improved in our Method 3 algorithms. From our framework for the "No Grouping" method, it is clear that the effective SNR performance is equivalent to the upper bound in (3.5). Thus its curve in Fig. 7.12 represents this value with respect to M. Therefore, the gap in effective SNR between "No Grouping" and Method 3 is the potential performance increase to be implemented. With improvements to Method 3 that reduce overhead or improve the convergence of Algorithm 2 to the upper bound, this gap in effective SNR can be diminished.

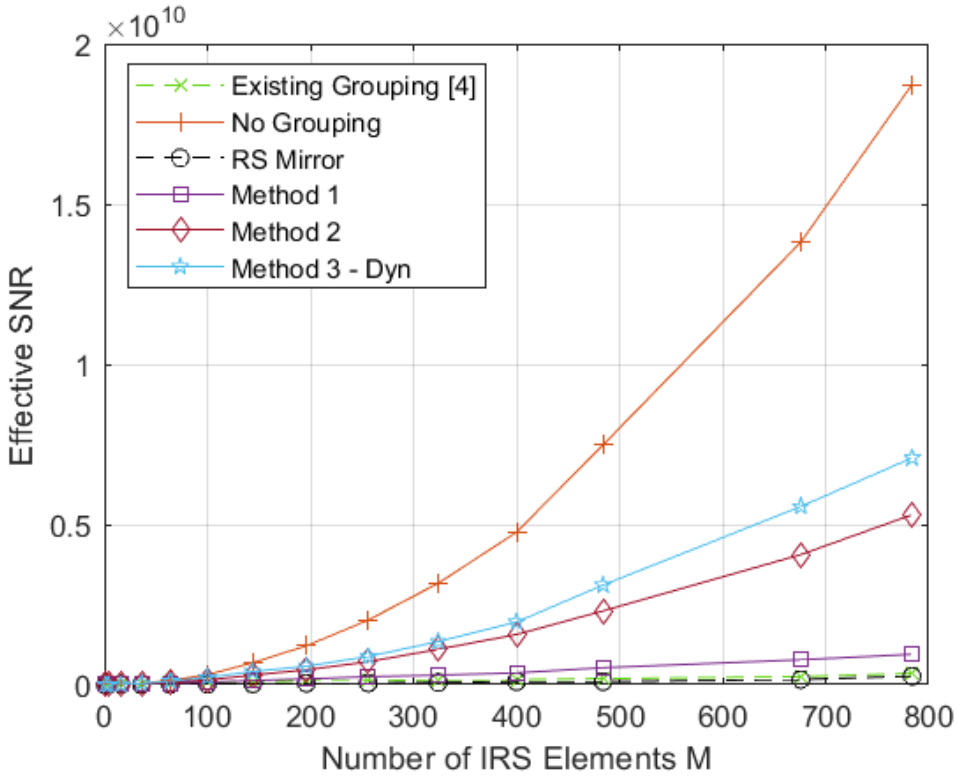

Figure 7.12: Effective SNR for proposed grouping methods and baseline curves with respect to M, $N = 9$.

Chapter 8

Conclusions

8.1 Conclusions

This thesis set out to prove that our proposed IRS element grouping scheme alleviates overhead limitations of IRS technology and thus opens up IRS use for high complexity massive MIMO applications. Current state-of-art works have used IRS element grouping to reduce immense overhead costs needed to optimize the reflection of an IRS-aided transmission, specifically by reducing the feedback communication and channel estimation time, but the number of groups had to be guessed at the BS and was not capable of being optimized to the wireless channel. In this thesis, we reformulated the grouping problem and showed that by leveraging the entire reflected channel matrix, we could design better element groupings that attained significantly improved effective SNR. We proposed three methods for grouping and reflection coefficient optimization, where the final dynamic grouping method shows the best performance. These dynamic element groupings change with each channel block and vary in group size and location to find the most efficient partitioning. While these grouping require more overhead to implement, our simulations show that the achievable rate performance trade-off is favorable, especially when considering the flexibility to choose the ideal grouping. Our simulations show that the proposed dynamic grouping method holds great potential to largely mitigate the high overhead associated with implementing IRS technology in massive MIMO systems.

8.2 Future Directions and Research

In this work we used an exhaustive search to find the best IGR for every channel. This is because, in practice, it is clear that the achievable rate as a function of ρ is not strictly convex for every channel realization and therefore not directly solvable. We leave the development of an efficient solution to find IGR for future research, which is needed in practice to keep calculation costs low but was not important to our simulations results. If no solvable optimization problem can be found, then in practice the convergence behavior of the IGR from Chapter 7 can be used to shrink the range of the exhaustive search.

The biggest hindrance for Method 3 is channel estimation overhead. If a channel estimation technique that utilized orthogonal pilot tones was developed to reduce T_{E} to $T_0\,(M+1)$, then numerical results indicate that the performance of Method 3 would have no issues outperforming the "RS Mirror" method for feasible massive MIMO or massive IRS implementations. Works such as [29] and [20] have proposed more efficient solutions, but there is room for improvement.

We established in Chapter 7 that Method 3 chooses a small IGR to minimize its high overhead. Therefore by improving the initialization algorithm, Algorithm 3, it is possible to decrease the gap to the upper bound with respect to ρ. The best way to do this would be to formulate an explicit solution that can find optimal grouping in one-shot rather than require the iterative solution proposed in this work. The initialization in Algorithm 3 considers the magnitude and angle of each IRS element's $r_{1,m}$ value to design an initial grouping. The magnitude and angle could be combined to form a "relative magnitude" between two IRS elements. The relative magnitude would be calculated by characterizing the fraction of the magnitude of r_{1,m_1} in the angle direction of another element's r_{1,m_2} complex value. By determining the relative magnitude between each IRS element and every other element it should theoretically be possible to use this information in a way to design the optimal groups in terms of effective SNR performance.

References

[1] E. Björnson, Özdogan, and E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?" *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 244–248, 2020.

[2] 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 Commun.*, vol. 20, no. 1, pp. 126–141, 2021.

[3] K. Zhi, C. Pan, G. Zhou, H. Ren, and K. Wang, "Analysis and optimization of ris-aided massive mimo systems with statistical csi," in *2021 IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, 2021, pp. 153–158.

[4] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets ofdm:protocol design and rate maximization," *IEEE Transactions on Commun.*, vol. 68, no. 7, pp. 4522–4535, 2020.

[5] Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 581–585, 2020.

[6] W. Cai, H. Li, M. Li, and Q. Liu, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light: Sci. Appl.*, vol. 3, no. 10, p. e218, Oct. 2014.

[7] ——, "Practical modeling and beamforming for intelligent reflecting surface aided wideband systems," *IEEE Wireless Commun. Letters*, vol. 24, no. 7, pp. 1568–1571, 2020.

[8] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive mimo 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 232–257, 2020.

[9] S. Hu, F. Rusek, and O. Edfors, "Beyond massive mimo: The potential of data transmission with large intelligent surfaces," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2746–2758, 2018.

[10] E. Björnson and L. Sanguinetti, "Demystifying the power scaling law of intelligent reflecting surfaces and metasurfaces," in *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2019, pp. 549–553.

[11] 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.

[12] 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.

[13] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Processing Letters*, vol. 27, pp. 905–909, 2020.

[14] A. Taha, Y. Zhang, F. B. Mismar, and A. Alkhateeb, "Deep reinforcement learning for intelligent reflecting surfaces: Towards standalone operation," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020, pp. 1–5.

[15] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," *IEEE Access*, vol. 9, pp. 44 304–44 321, 2021.

[16] Özdoğan and E. Björnson, "Deep learning-based phase reconfiguration for intelligent reflecting surfaces," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*, 2020, pp. 707–711.

[17] 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.

[18] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[19] H. Li, R. Liu, M. Liy, Q. Liu, and X. Li, "Irs-enhanced wideband mu-miso-ofdm communication systems," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, 2020, pp. 1–6.

[20] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6607–6620, 2020.

[21] Y. Yang, S. Zhang, and R. Zhang, "Irs-enhanced ofdm: Power allocation and passive array optimization," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[22] 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.

[23] X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, and E. Björnson, "Ris-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials," *IEEE Transactions on Communications*, pp. 1–1, 2021.

[24] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[25] 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.

[26] 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.

[27] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced ofdm: Channel estimation and reflection optimization," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 518–522, 2020.

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

[29] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced ofdm: Channel estimation and reflection optimization," *IEEE Wireless Commun. Letters*, vol. 9, no. 4, pp. 518–522, 2020.

[30] D. Love and R. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1102–1110, 2003.

[31] S. Park, A. Alkhateeb, and R. W. Heath Jr., "Dynamic subarrays for hybrid precoding in wideband mmwave mimo systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 5, pp. 2907–2920, 2017.

Appendix A

Definition of the Kronecker Product

$$\mathbf{a} = [a_1, a_2, a_3]^{\mathrm{T}}, \mathbf{b} = [b_1, b_2, b_3]^{\mathrm{T}}$$

$$\mathbf{a} \otimes \mathbf{b} = [a_1 b_1,\ a_1 b_2,\ a_1 b_3,\ \cdots,\ a_3 b_1,\ a_3 b_2,\ a_3 b_3]^{\mathrm{T}}$$

Appendix B

Definition of the Khatri-Rao Product

Here is the definition of the column-wise Khatri-Rao product used in this work:

$$\mathbf{C} = \left[\begin{array}{c|c|c} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{array}\right] = \left[\begin{array}{c|c|c} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{array}\right],$$

$$\mathbf{D} = \left[\begin{array}{c|c|c} \mathbf{d}_1 & \mathbf{d}_2 & \mathbf{d}_3 \end{array}\right] = \left[\begin{array}{c|c|c} d_{1,1} & d_{1,2} & d_{1,3} \\ d_{2,1} & d_{2,2} & d_{2,3} \\ d_{3,1} & d_{3,2} & d_{3,3} \end{array}\right],$$

where the Khatri-Rao product of these two matrices would be:

$$\begin{aligned} \mathbf{C} \odot \mathbf{D} &= \left[\begin{array}{c|c|c} \mathbf{c}_1 \otimes \mathbf{d}_1 & \mathbf{c}_2 \otimes \mathbf{d}_2 & \mathbf{c}_3 \otimes \mathbf{d}_3 \end{array}\right] \\ &= \left[\begin{array}{c|c|c} c_{1,1}d_{1,1} & c_{1,2}d_{1,2} & c_{1,3}d_{1,3} \\ c_{1,1}d_{2,1} & c_{1,2}d_{2,2} & c_{1,3}d_{2,3} \\ c_{1,1}d_{3,1} & c_{1,2}d_{3,2} & c_{1,3}d_{3,3} \\ \vdots & \vdots & \vdots \\ c_{3,1}d_{1,1} & c_{3,2}d_{1,2} & c_{3,3}d_{1,3} \\ c_{3,1}d_{2,1} & c_{3,2}d_{2,2} & c_{3,3}d_{2,3} \\ c_{3,1}d_{3,1} & c_{3,2}d_{3,2} & c_{3,3}d_{3,3} \end{array}\right] \end{aligned}$$