Multi-user Beamforming on Intelligent Reflecting Surface and Networks

Yihong Liu

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

School of Engineering
College of Science and Engineering
University of Glasgow

June 2022

Abstract

A new type of retrofitted low-cost material is recently proposed and made it possible to control (or program) part of the wireless channel around us. It is called the intelligent reflecting surface (IRS), also named reconfigurable intelligent surface (RIS) or metasurface. The metamaterial composed surface can realize selective EM properties by integrating artificially designed electronic elements that can be controlled by processors (e.g. field programmable gate array (FPGA)). Therefore, the wireless channel is controllable with such IRS posting on the ceiling and wall. Specifically, this functionality is realized by controlling the excitation and phase of each electronic element on the surface. The phase and amplitude of the EM wave impinging on the surface can be reflected in a designed manner of EM wave's superposition. From this view, the wireless transmission can be enhanced by focusing the signal power while mitigating the interference power.

Based on the promising characteristics of IRS, the thesis aims to systematically study the design of IRS to improve transmission performance, where the key performance indicator (KPI) includes signal power gain level, interference suppressed level, signal-to-interference plus noise ratio (SINR), angle accuracy of beamforming, energy efficiency (EE) and spectrum efficiency (SE). To improve the system performance, the model of IRS used for wireless transmission should be first investigated as a cornerstone for the following design. The design proposed in this thesis on the IRS will mainly include the beamforming design for multi-user (MU) for power maximizing and interference minimizing. Meanwhile, the thesis also tends to provide sufficient analysis to the insight into MU beamforming in terms of intrinsic system constraints. Moreover, efficient guidance for single IRS and multiple IRSs deployments is also one of the objectives to be carried out. Lastly, the thesis also aims to address and bridge the gap between theoretical assumptions and the physical ground truth to improve the proposed design with more practical concerns.

Specifically, in this thesis, the first contribution includes the fundamental modelling of a single IRS. Meanwhile, the multi-user (MU) beamforming and interference-free transmission are achieved with a single IRS. An optimization problem called multi-user linearly constrained minimum variance (MU-LCMV) beamformer is formulated under the criterion of

minimizing the overall received signal power subject to a certain level of power response at desired signal directions and arbitrary low power response at the interference directions. A closed-form amplitude-unconstrained phase-continuous (AUPC) solution is derived first, and then an amplitude-constrained phase-continuous (ACPC) solution is obtained by using sequential quadratic programming (SQP). The weights for multi-bit control or binary phase control are also obtained by quantization as amplitude-constrained phase-quantized (ACPQ) solution. It is interesting to find that given arbitrary solutions, the IRS beam pattern shows that to achieve MU (N pairs of transceivers, $N > 1$) transmission through a single surface, up to $N-1$ redundant beams are generated, significantly affecting power efficiency. The directions of the redundant beams are mathematically derived, which is important for inter-user interference management of IRS-aided wireless communication systems. Note that the redundant beams of IRS for MU transmission are intrinsic issues, not limited to configuration structure and optimization algorithms. The finding of redundant beams can be one of the most important contributions to this thesis. Simulation results verify the existence and accuracy of the redundant beam directions.

Another contribution of the thesis focuses on analysing and optimising the IRS networks, which are made up of multiple IRSs. The reachable sum rate upper bound is derived for IRS networks with different network graph topologies, where the limit of a single IRS is set as the fundamental cornerstone. In particular, the optimal condition for achieving sum-rate upper bound with one IRS in a closed-form function and the analytical condition to achieve interference-free transmission are derived, respectively. Leveraging this optimal condition, the multi-user multi-order-reflection (MUMOR) sum-rate upper bound of IRS network with different network topologies are obtained, where the linear graph (LG), complete graph (CG) and null graph (NG) topologies are considered. For these three topologies, the sum rate limit is applicable for an arbitrary number of reflections, arbitrary size, and the number of IRSs/reflectors in the networks. The channel estimation strategy is also discussed for intelligent reflecting surface (IRS) networks. Simulation results verify our theories and derivations and demonstrate that different network topologies' sum rate upper bounds are under a K-fold improvement given K-piece IRS.

The last contribution of the thesis is to analyze the system performance with the influence of practical effects. In particular, two practical effects, the mutual coupling and near-field effect, which exist objectively and bring bias to the theoretical model, are discussed and analyzed, respectively. Since the size of IRS can be embedded with hundreds of thousands of reflecting elements, the near-field effect is considered more common. To compensate, the design of near-field beamforming for MU with different distances to IRS is given. The redundant beam with a near-field effect is numerically analyzed. The discriminant is derived to

judge the existence of redundant beams, which can also improve the complexity of numerical computation for redundant beams. Simulation results validate that the mutual coupling mainly affects the side lobe level and near-field effect brings bias to the covered area by redundant beams. In addition, the approximation-free analysis in the near-field to the redundant beam position is found to be more precise than conventional approximation analysis in one order of magnitude.

To conclude, the work presented in this thesis provides insight into beamforming on IRS from the connections between the hardware components and the basic transmission models in physical layer transmission. The weight design, deployment of IRS networks and performance limitation analysis are fundamental and can potentially enhance any state-of-the-art wireless communication system ranging from transceiver design, system and architecture design, network deployment, and self-organizing-network operations.

Contents

List of Publications

Papers Included in The Thesis

Journal

- **Y. Liu**, L. Zhang and M. A. Imran, "Multi-user Beamforming and Transmission Based on Intelligent Reflecting Surface," in IEEE Transactions on Wireless Communications, doi: 10.1109/TWC.2022.3157808.
- **Y. Liu**, L. Zhang, F. Gao and M. A. Imran, "Intelligent Reflecting Surface Networks with Multi-Order-Reflection Effect: System Modelling and Critical Bounds," in IEEE Transactions on Communications, doi: 10.1109/TCOMM.2022.3202212.
- **Y. Liu**, L. Zhang, F. Gao and M. A. Imran, "Multi-user Beamforming of IRS with Near-Field Effect," in IEEE Communications Letters, drafting.
- H. Han, **Y. Liu** and L. Zhang, "On Half-Power Beamwidth of Intelligent Reflecting Surface," in IEEE Communications Letters, vol. 25, no. 4, pp. 1333-1337, April 2021, doi: 10.1109/LCOMM.2020.3046369.

Conference

- **Y. Liu**, L. Zhang, B. Yang, W. Guo and M. A. Imran, "Programmable Wireless Channel for Multi-User MIMO Transmission Using Meta-Surface," 2019 IEEE Global Communications Conference (GLOBECOM), 2019, pp. 1-6.
- **Y. Liu**, L. Zhang, P. V. Klaine and M. A. Imran, "Optimal Multi-user Transmission based on a Single Intelligent Reflecting Surface," 2021 IEEE 4th International Conference on Electronic Information and Communication Technology (ICEICT), 2021, pp. 1-4, doi: 10.1109/ICEICT53123.2021.9531152.

Papers Not Included in The Thesis

Journal

- X. Lin, L. Zhang, A. Tukmanov, **Y. Liu**, Qammer Abbasi and M. A. Imran, "On the Design of Broadbeam of Intelligent Reflecting Surface," in IEEE Transactions on Vehicular Technology, under review.

Conference

- S. Yang, H. Han, **Y. Liu**, W. Guo, and L. Zhang, “Intelligent Reflecting Surface-induced Randomness for mmWave Key Generation," in IEEE Global Communications 2022 Workshop, under review.
- Y. Zhang, L. Yang, Y. Wang, M. Wang, Y. Chen, L. Qiu, **Y. Liu**, G. Xue, and J. Yu, “Acoustic Lens for Sensing and Communication,” in 20th USENIX Symposium on Networked Systems Design and Implementation, drafting.

List of Figures

List of Tables

Glossary

2G 2nd generation. 1

4G 4th generation. 12

5G 5th generation. 1, 12, 114

6G 6th generation. 18

ACPC amplitude-constrained phase-continuous. 26, 27, 37, 46, 50, 63, 67, 69, 103, 115

ACPQ amplitude-constrained phase-quantized. 37, 46, 52, 63, 115

AF array factor. 54

AOA angle of arrival. 5, 9, 40–43, 56, 125

AOD angle of departure. 9, 40–43, 81, 82, 125

AP access point. 18

AUPC amplitude-unconstrained phase-continuous. 26, 27, 34, 46, 63, 67, 69, 115

AXPC amplitude-relaxed phase-continuous. 34, 46, 63, 69, 108, 111

BS base station. 14, 20–24, 40, 80, 81

C-LoS cascaded-nonline-of-sight. 25, 39, 40, 43–45, 74, 76, 83, 84, 86, 88, 92

CG complete graph. 28

CP cyclic prefix. 13

CS compressive sensing. 21

CSI channel state information. 13, 23, 29, 46, 79

D2D device-to-device. 23

dB decibels. 2

DC direct-current. 15, 16

EE energy efficiency. 13, 21–23, 25, 26, 70, 114–116

EM electromagnetic. 1, 14, 16–18, 20, 31, 40

FIFO far-in far-out. 96

FINO far-in near-out. 97

FIR finite impulse response. 8

FOR first-order-reflection. 25, 39, 40, 44

FPGA field programmable gate array. 17

FSPL free space path loss. 1–4, 31, 94, 111

GA genetic algorithm. 63

GHz giga hertz. 4, 108

HPBW half-power beamwidth. 20, 53–55, 67, 69, 115

IoT internet of things. 18

IRS intelligent reflecting surface. 14, 15, 17–29, 32, 35, 37–40, 44–47, 54, 69–71, 74, 79–83, 94–96, 108, 112, 114–118

ISAC integrated sensing and communication. 117

ISI inter-symbol interference. 13

LC inductor-capacitor. 52

LCMV linearly constrained minimum variance. 47, 48, 69

LG linear graph. 28

LoS line-of-sight. 4, 5, 11, 12, 29, 31, 39, 40, 42, 44

m-MIMO massive multiple-input multiple-output. 8, 13

MIMO multiple-input multiple-output. 8, 9, 11–14, 20, 114

MISO multi-input-single-output. 20, 22, 24

MMSE minimal mean-squared error. 21, 96

mmWave millimeter wave. 8, 12, 108

MOR multi-order-reflection. 23–25, 29, 44, 92, 94, 115, 117

MRC max ratio combining. 46, 47, 53, 54, 69, 72, 81, 103, 108, 115

MSE mean-squared error. 23

MU multi-user. xii, 8, 18, 19, 21–27, 46, 47, 52, 70, 95, 96, 99, 100, 108, 109, 111, 112, 114–116, 118

MU-LCMV multi-user linearly constrained minimum variance. 26, 27, 47, 63, 83, 99, 111, 114, 115

MUMOR multi-user multi-order-reflection. 24, 79, 115

MUMP multi-user multi-path. 116

MUSP multi-user single-path. 92, 116

NG null graph. 28

NIFO near-in far-out. 97

NINO near-in near-out. 97, 99

NLoS nonline-of-sight. 4, 11, 29, 40

NOMA non-orthogonal multiple access. 22

OFDM orthogonal frequency-division multiplexing. 13, 14, 20

OMA orthogonal multiple access. 22

PIN positive-intrinsic–negative. 15, 16, 37

PLE path loss exponent. 3, 4

QoS quality of service. 12–14, 18, 21–23, 46

RF radio frequency. 9, 12–14, 16, 19

RIS reconfigurable intelligent surface. 14, 15

Rx receiver. 1, 13, 23, 29–32, 39, 40, 69

SDMA space-division multiple access. 11, 12

SDR semidefinite relaxation. 21

SE spectral efficiency. 13, 23, 25, 26, 70, 114–116

SINR signal-to-interference-plus-noise ratio. 21, 23, 32, 53, 67, 100

SLL side-lobe level. 20

SNR signal-to-noise ratio. 18, 22, 94, 108, 115

SOR second-order-reflection. 39, 40, 44, 84

SQP sequential quadratic programming. 26, 27, 50, 51, 60, 63, 69, 103

STAR simultaneously transmitting and reflecting. 21, 22

STD standard deviation. 4

SU single user. 19, 20, 22–24

SUMOR single-user multi-order-reflection. 23

SUMP single-user multi-path. 92, 94, 116

SUSP single-user single-path. 116

SVD singular value decomposition. 10

THz terahertz wave. 20

Tx transmitter. 1, 13, 29–32, 39, 69

UE user. 23, 79–83

ULA uniform linear array. 5, 30, 33, 46, 54, 55, 98, 108

URA uniform rectangular array. 26, 33, 46, 55, 108

ZF zero forcing. 18, 21, 86, 88

List of Symbols

$(\cdot)^H$	The complex conjugate transpose
$(\cdot)^T$	The matrix (vector) transpose
$\angle$	The phase of a complex variable
$\hat{\mathbf{y}}_s$	The phased signal on IRS
λ	Wavelength of transmitted carrier
$\mathbb{C}$	Complex numbers
$\mathbb{R}$	Real numbers
$\mathbf{A_{in}}$	Matrix containing incident steering vector with respect to IRS
$\mathbf{A_{out}}$	Matrix containing exit steering vector with respect to IRS
$\mathbf{a}$	The steering vector
$\mathbf{A}_C$	Combined steering matrix
$\mathbf{a}_C$	The combined steering vector by incident steering vector and exit steering vector
$\mathbf{a}_{in}$	Incident steering vector with respect to IRS
$\mathbf{a}_{out}$	Exit steering vector with respect to IRS
$\mathbf{I}$	The identity matrix
$\mathbf{n}$	The noise vector
$\mathbf{s}$	The transmitted signal vector
$\mathbf{W}$	The weights matrix of an IRS
$\mathbf{w}$	The vectorized weights from an IRS weights matrix

$\mathbf{y}$ The received signal vector

$\mathscr{A}_r$ The effective aperture of the receiving antenna

$\mathscr{A}_t$ The effective aperture of the transmitting antenna

$\mathscr{L}$ The multipath number

$\mathscr{X}_\sigma$ The random variable brought by channel fading

$\odot$ The point-wise multiplication

$\mathbf{\Omega}_{in,i}$ The spatial information of incident directions from i-th transmitter

$\mathbf{\Omega}_{out,i}$ The spatial information of exit directions towards i-th receiver

${}^{X}P_{Y}$ The number of Y-permutations from a set S of X elements

ϕ_{in} Incident azimuth angle with respect to IRS

ϕ_{out} Exit azimuth angle with respect to IRS

σ The power of noise

$\tau(\theta)$ The considered time delay in angle domain

θ The angle symbol

θ_{in} Incident elevation angle with respect to IRS

θ_{out} Exit elevation angle with respect to IRS

$\triangleq$ The equivalent symbol in terms of optimization

Ξ_m The phase shifts applied on m-th element of IRS

c The speed of the light

d_0 The distance between the adjacent elements

d_{in} The distance from the reference point of transmitter to the reference point of IRS

d_{out} The distance from the reference point of IRS to the reference point of receiver

d_{tr} The distance between the transmitting antennas and receiving antennas

$diag(\cdot)$ The symbol for vectoring a matrix by taking its diagonal terms

$E\{\cdot\}$ The expectation operator.

f Frequency of transmitted carrier

K The total number of IRS pieces in the environment

l_{IRS} Path loss factor via IRS

M The number of elements

$m(t)$ The base band signal in time domain

N The number of transceiver pairs

P_r The received signal power

P_t The transmitted signal power

$s(t)$ The input signal in time domain

$\binom{X}{Y}$ The number of Y-combinations from a set S of X elements

Acknowledgements

First of all, I would like to express my deepest thanks to my primary supervisor, Prof. Lei Zhang, who contributes a lot of his time to revising my papers. Though rejections with harsh comments by reviewers always put me into self-doubt and depression, it is Prof. Lei Zhang's continuous and determined encouragement that helps me to combat these issues. Without a doubt, Prof. Zhang is the best supervisor I can ever follow in my research life and life research. My gratitude is beyond my words.

I must also thank Prof. Muhammad Ali Imran, my second supervisor, for contributing valuable and critical comments to my work on intelligent surfaces.

Moreover, I would like to thank my friends in room 724 of JWS Building for discussions before the abruption of Covid-19.

Finally, I would like to thank my family for any kind of suggestions, support and helps.

Chapter 1

Introduction

To achieve faster, smarter, safer communication in the future, several generations of wireless networking (2^{nd} generation (2G) to 5^{th} generation (5G)) have improved the achievable throughput and accessibility to diverse services [7]. The main targets of the wireless communication systems remain similar, i.e., how to efficiently utilize the scarce resources (especially spectrum) and how to deal with the unknown and uncontrollable signal propagation channels. Nevertheless, with the resource allocation based on all kinds of optimization criteria, the wireless channel plays the most critical and deterministic role in the design of all transmission techniques.

1.1 Background

For wireless transmission, the uncontrollable and unknown characteristics of the physical transmission channel originate from two main issues, which are channel path-loss and fading, respectively.

1.1.1 Channel Path-loss

Channel path-loss, the reduction (attenuation) in power density of an electromagnetic (EM) wave as it propagates through the transmission environment, is typically determined by the propagation environment, propagation medium and the distance between the transmitter (Tx) and the receiver (Rx). For the free space path loss (FSPL), where the EM wave propagates in a vacuum as the carrier, we have the relationship between the power of Rx and Tx concerning the transmission distance given by the Friis transmission equation [1]. The Friis formula can be written as

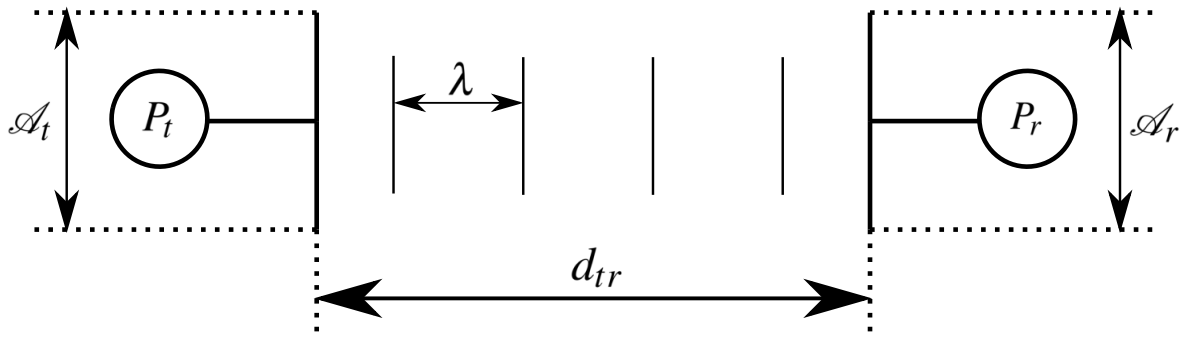


Figure 1.1: Friis transmission model [1].

$$\frac{P_r}{P_t} = \mathscr{A}_t \mathscr{A}_r \left(\frac{1}{d_{tr}\lambda}\right)^2, \tag{1.1}$$

where $\mathscr{A}_t$ is the effective aperture of the transmitting antenna, $\mathscr{A}_r$ is the effective aperture of the receiving antenna, λ is the carrier's wavelength, and d_{tr} is the distance between the transmitting antennas and receiving antennas, as shown in Fig. 1.1. To make the formula above hold, the distance d_{tr} must be large enough such that the antennas are in the far field of each other, i.e., $d_{tr} >> \lambda$.

Considering the single isotropic antennas for transmitting and receiving with unit gain, on the basis of consistency with thermodynamics [8], the effective aperture of the isotropic antenna can be derived as

$$\mathscr{A}_{iso} = \frac{\lambda^2}{4\pi}. \tag{1.2}$$

By substituting the effective aperture into equation (1.1), we can calculate FSPL as

$$\text{FSPL} = \frac{P_t}{P_r} = \left(\frac{4\pi d_{tr}}{\lambda}\right)^2. \tag{1.3}$$

It can be noticed that the power attenuation is based on distance and transmission frequency in the free space transmission environment. The path-loss is directly proportional to the square of the distance and is inversely proportional to the square of wavelength[1]. For most of the current literature, the path-loss is represented in decibels (dB) for convenience and hence

[1]Note that although distance dependence of path-loss is straightforward, the frequency dependence of path-loss does not come from the free space environment. The frequency dependence of path-loss is rather from the effective aperture since increasing the frequency result in a shorter wavelength and hence mitigate the effective aperture.

	$PL(d_0)$ [dB]	n	σ[dB]
Corridor	68	1.64	2.53
LoS hall	68	2.17	0.88
NLoS hall	68	3.01	1.55

Table 1.1: Statistical parameters of 60 GHz channel [2].

FSPL can be rewritten as

$$\begin{aligned} \mathrm{FSPL}(d_{tr}, f) &= 10\log_{10}\left(\left(\frac{4\pi d_{tr} f}{c}\right)^2\right) \\ &= 20\log_{10}\left(\frac{4\pi d_{tr} f}{c}\right) \\ &= 20\log_{10}(d_{tr}) + 20\log_{10}(f) + 20\log_{10}\left(\frac{4\pi}{c}\right) \\ &= 20\log_{10}(d_{tr}) + 20\log_{10}(f) - 147.55 \end{aligned}, \tag{1.4}$$

where c is the speed of the light, and f is the frequency of the carrier.

1.1.2 Channel Fading

For practical wireless transmission, the carrier transmitted within the environment typically experiences the multi-path effect [9]. The multi-path effect is caused by objects within the transmission environment that reflect the EM wave randomly since the objects in the environment are usually randomly placed. Therefore, the carrier impinging on these objects can be reflected/scattered randomly to the Rx via multiple paths. Since reflected/scattered carriers passing through these paths have different EM phases, carriers can add either constructively or destructively at the Rx, which brings the source of channel fading. In particular, channel fading is introduced as a class of random variation on the attenuation of the carrier signal. Other factors, e.g., transmission medium absorption, different weather conditions [10], etc., may also contribute to the channel fading.

To predict the practical path-loss with fading, the fading is modelled as a random variable in the path-loss model, which can be written as

$$\mathrm{PL}(d_{tr}, f) = 20\log_{10}(d_{tr}) + 20\log_{10}(f) + \mathscr{X}_\sigma - 147.55, \tag{1.5}$$

where $\mathscr{X}_\sigma$ is the random variable brought by channel fading, following typical distributions. Considering certain terrain types encountered in the practical transmission environment, the path loss exponent (PLE) model with correction factors are leveraged. Correction factors are

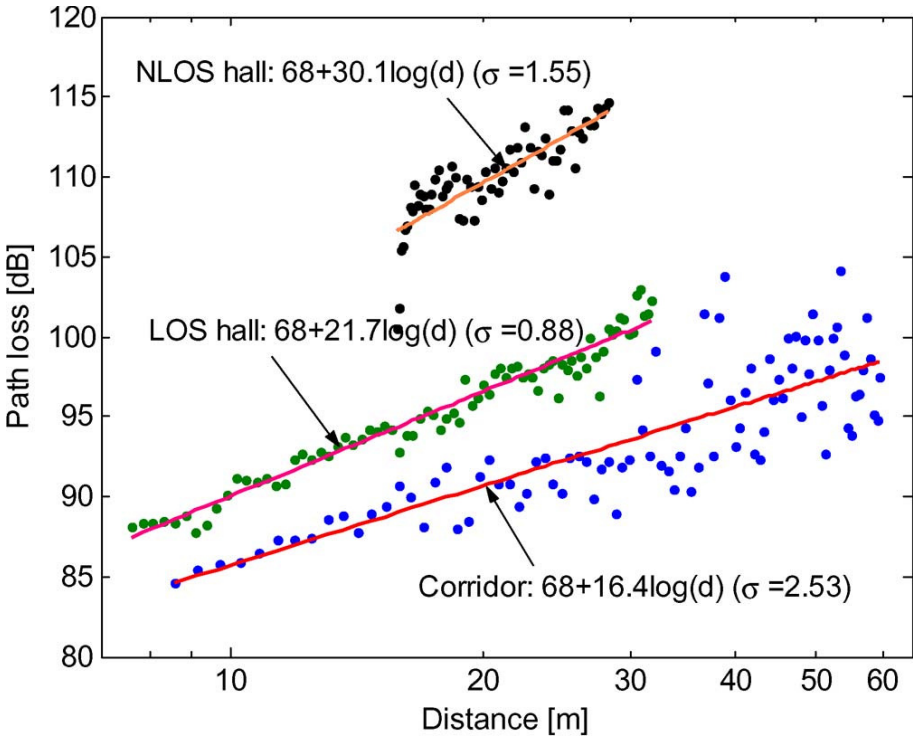


Figure 1.2: The measurement on path loss of 60 GHz carrier [2].

introduced in the exponent model to fit with the actual experiment measurement [11–13]. For instance, one of the indoor PLE model is given by [2] for 60 giga hertz (GHz) carrier from measurements, which is written as

$$PL(d_{tr}) = PL_0(d_0) + 10n\log_{10}\frac{d_{tr}}{d_0} - S_\sigma. \tag{1.6}$$

In equation (1.6), n is the path loss exponent that characterizes the increasing rate of path loss with the increase of separation d_{tr} between the Txs and the Rxs. The PL_0 denotes the FSPL at a reference distance d_0, which is often chosen as 1m for indoor environments. Parameter σ is the standard deviation (STD) of variations S_σ of the local average of the signal power. The minus sign before S_σ is taken to be consistent with the definition in measurements [2]. Parameters n and σ can be extracted from the experimental measurement data by least mean squares error fitting, and they are highly dependent on measurement environments and scenarios [14].

The indoor scenarios, with the line-of-sight (LoS) cases and the nonline-of-sight (NLoS) cases[2], are measured respectively, as shown in Fig. 1.2. The corresponding statistical parameters are listed in Table 1.1. It can be observed that the NLoS hall transmission experiences the worst path loss, compared with the other two scenarios, which are LoS hall and corridor

[2]The line-of-sight path is blocked and carrier can only reach the Rx by reflecting/scattering of the environment in this case.

transmission respectively. It is also worth mentioning by taking $n = 2$, the path-loss model can be equivalent to the model in equation (1.5) as the $n = 2$ indicates an exponent value of free space transmission. The measurement results also show the LoS hall transmission has a very close path loss exponent to the free space transmission. In addition, the corridor transmission has less path loss than that of LoS hall transmission, which indicates that path reflections can be helpful for reducing path loss in a reasonable distance range.

1.1.3 Beamforming

To compensate for channel path-loss and channel fading, beamforming techniques are typical approaches for the signal power loss over the wireless channel [3, 15]. As shown in Fig. 1.3, a commonly considered uniform linear array (ULA) is presented, where it has M antennas/elements[3] in total, and the spacing is d_0 between adjacent elements. The carrier wavefront is a plane wave due to the far-field transmission. Assume a narrowband signal $s(t)$ of frequency ω impinges on the ULA with unified frequency response. The angle of arrival (AOA), θ, is measured from the axis of the array. With the $m(t)$ to be the complex baseband signal, the input signal $s(t)$ can be written as

$$s(t) = m(t)e^{j\omega t}. \tag{1.7}$$

Since each element receive the same signal with a different delay, given the different AOA, for the received signal at i-th element, where $i = 1, 2, ..., M$, it can be written as

$$x_i(t) = m\left(t - \Sigma_{n=1}^{i}\tau_n(\theta)\right) e^{j\omega\left(t-\Sigma_{n=1}^{i}\tau_n(\theta)\right)}, \tag{1.8}$$

where $\tau_n(\theta)$ is the time delay between the n-th and $(n+1)$-th elements. For narrowband beamforming, it is assumed that the bandwidth of the signal is sufficiently narrow enough, and the array dimension is sufficiently small enough for the baseband signal $m(t)$ to stay almost constant during the overall considered coherence time across the array with

$$\tau(\theta) = \Sigma_{n=1}^{M-1}\tau_n(\theta), \tag{1.9}$$

then we can have the approximation

$$m(t) \approx m(t - \tau(\theta)). \tag{1.10}$$

[3] In the following thesis, the term "element" is used to indicate the basic unit for arbitrary array structure.

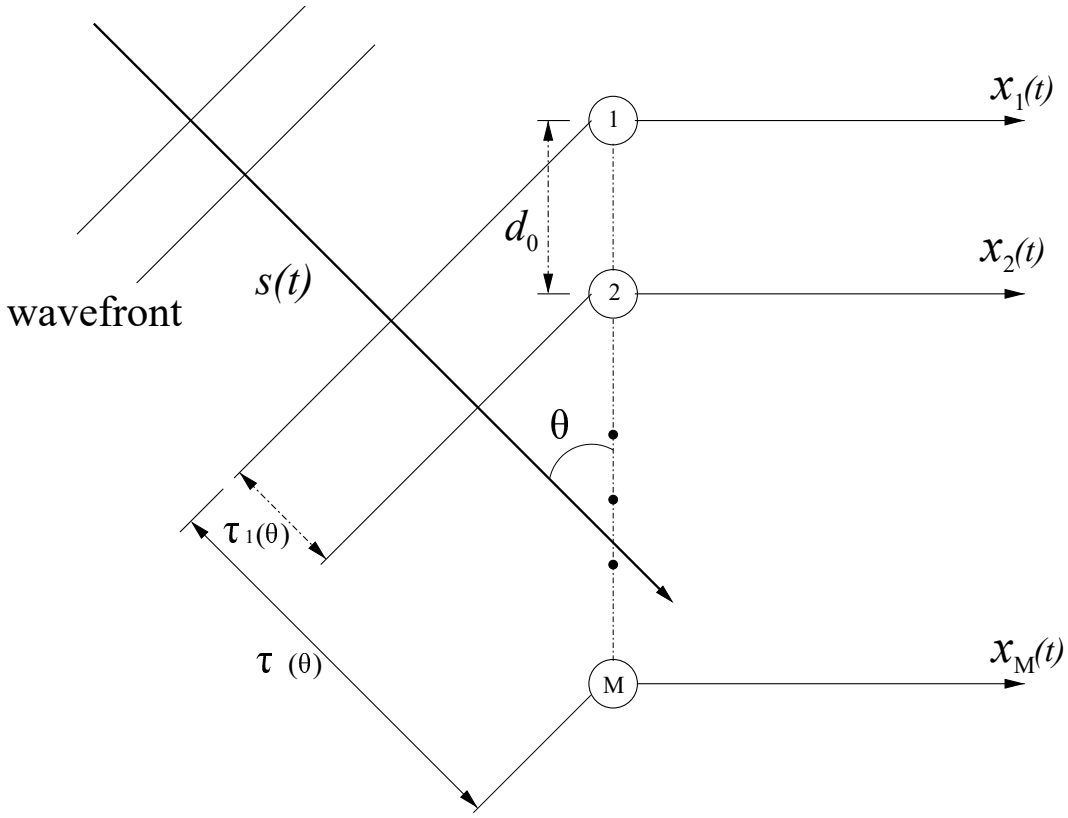


Figure 1.3: An ULA for receiving carrier signal [3].

Therefore, equation (1.8) can be simplified to

$$x_i(t) = e^{-j\omega\Sigma_{n=1}^{i}\tau_n(\theta)} m(t) e^{j\omega t} = a_i(\theta) s(t), \tag{1.11}$$

where $a_i(\theta) = e^{-j\omega\Sigma_{n=1}^{i}\tau_n(\theta)}$. Conventionally, the signal received by the array is written in a vector form as

$$\mathbf{x}(t) = \mathbf{a}(\theta) s(t) + \mathbf{n}(t), \tag{1.12}$$

where

$$\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_M(t)]^T \in \mathbb{C}^{M\times 1} \tag{1.13}$$

and

$$\mathbf{a}(\theta) = \left[1, e^{-j\omega\tau_1(\theta)}, \cdots, e^{-j\omega\Sigma_{n=1}^{M-1}\tau_n(\theta)}\right]^T \in \mathbb{C}^{M\times 1}. \tag{1.14}$$

The $\mathbf{n}(t)$ is the noise vector (usually Gaussian), and $\mathbf{a}(\theta)$ is the expression of the steering vector[4]. Note that the form of steering vector has been conventionally used in the literature, and we will keep the notation of steering vectors for the following thesis. The steering vector can be seen physically as the phase delay characteristic of a plane wave from one specific

[4]Steering vector here can also represent the same meaning for the transmitting array due to the transmitting/receiving symmetry of the model.

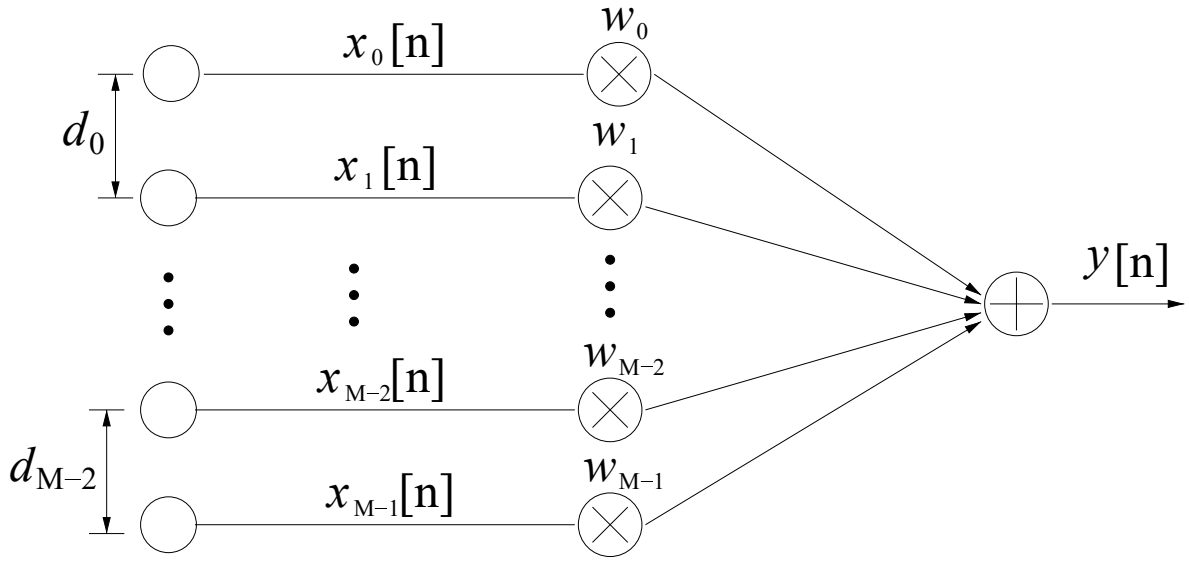


Figure 1.4: A general beamforming structure with arbitrary spacing [3].

direction impinging on the receiving array.

Note that equation (1.12) represents the received signal vector from only one direction. Involving the multi-path effect, assume the carrier experiences an ideal scattering environment with L paths. Then, the signal received directly by the array can be extended such that

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{1.15}$$

where

$$\mathbf{s}(t) = [s(t_1), s(t_2), \cdots, s(t_L)]^T \in \mathbb{C}^{L\times 1} \tag{1.16}$$

is the input signal vector impinging on the array from $\mathscr{L}$ directions. The input signal from each direction has its corresponding t_l path delay, for $l = 1, 2, ..., \mathscr{L}$. The steering vector from $\mathscr{L}$ directions can be written in the matrix form

$$\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_{\mathscr{L}})] = \begin{bmatrix} 1 & e^{-\omega\tau_1(\theta_1)} & \dots & e^{-j\omega\Sigma_{n=1}^{M-1}\tau_n(\theta_1)} \\ 1 & e^{-\omega\tau_1(\theta_2)} & \dots & e^{-j\omega\Sigma_{n=1}^{M-1}\tau_n(\theta_2)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-\omega\tau_1(\theta_{\mathscr{L}})} & \dots & e^{-j\omega\Sigma_{n=1}^{M-1}\tau_n(\theta_{\mathscr{L}})} \end{bmatrix}^T \in \mathbb{C}^{M\times L}. \tag{1.17}$$

Equivalently, the received signal on the array can be represented as the form by the sum of paths as

$$\mathbf{x}(t) = \Sigma_{l=1}^{\mathscr{L}} \mathbf{a}(\theta_l)\, s(t_l) + \mathbf{n}(t). \tag{1.18}$$

As now L paths can bring different delays, these signals with different phases can construct positively or negatively, resulting from the multi-path effect. Note that different from the

fading brought by the multi-path effect in the path-loss model, where it is usually modelled as a random variable, the multi-path effect can be characterized precisely by the exact knowledge of the incident directions of these multi-path signals. After the array receives the raw impinging signal vector $\mathbf{x}(t)$, the signal processing can be implemented to optimize the received signal with the beamforming structure. Similar to the finite impulse response (FIR) filter in the traditional signal processing of the frequency domain [16], the beamforming is a class of filtering in the space domain [17]. A more general discussion on the beamforming structure with arbitrary elements spacing is shown in Fig. 1.4. The w_i, $i = 1, 2, ...M$, is the designed weight parameter, which can adjust the phase delay and signal gain. The d_j, $j = 1, 2, ..., M-1$, is the generalized element spacing. The beamformer output signal $y(t)$ can be expressed as

$$y(t) = \sum_{i=1}^{M} w_i x_i(t). \tag{1.19}$$

Therefore, the noise and interference[5] can be respectively mitigated by the designed beamformer's weights value.

1.1.4 Multiple-input Multiple-output Systems

Thanks to the multiple-input multiple-output (MIMO) technique [4] and other advanced multi-antenna approaches such as massive multiple-input multiple-output (m-MIMO) [18–20] and millimeter wave (mmWave) communications [13], faster communication and better spectrum utilization can be achieved. As indicated by the name of MIMO, multiple antennas are required for both Rxs and Txs, as shown in Fig 1.5(a). Due to antenna arrays being employed, the MIMO channel model combined the physical channels of both the transmitting array and receiving array, $\mathbf{H}'$ and $\mathbf{H}''$. The transmission model of MIMO system can therefore be represented as

$$\mathbf{x}(t) = \mathbf{H}''\mathbf{H}'\mathbf{s}_{tr}(t) + \mathbf{n}(t), \tag{1.20}$$

where $\mathbf{s}_{tr}(t) = [s_{tr,1}(t)s_{tr,2}(t), \ldots, s_{tr,M}(t)]$ is the signal vector from the transmitting array. Considering L paths in the transmission environment, $\mathbf{H}'$ and $\mathbf{H}''$ can be written in an exact the same form with $\mathbf{A}$ in equation (1.17). By writing the MIMO channel with the sum of the L paths' signals, we have

$$\mathbf{x}(t) = \Sigma_{l=1}^{\mathscr{L}} \mathbf{a}\left(\theta_{r_l}\right) \mathbf{a}^H\left(\theta_{t_l}\right) \mathbf{s}_{tr}(t) + \mathbf{n}(t), \tag{1.21}$$

[5]The interference can be MU interference if other Txs' signals are received, or potential self-interference due to the multi-path effect.

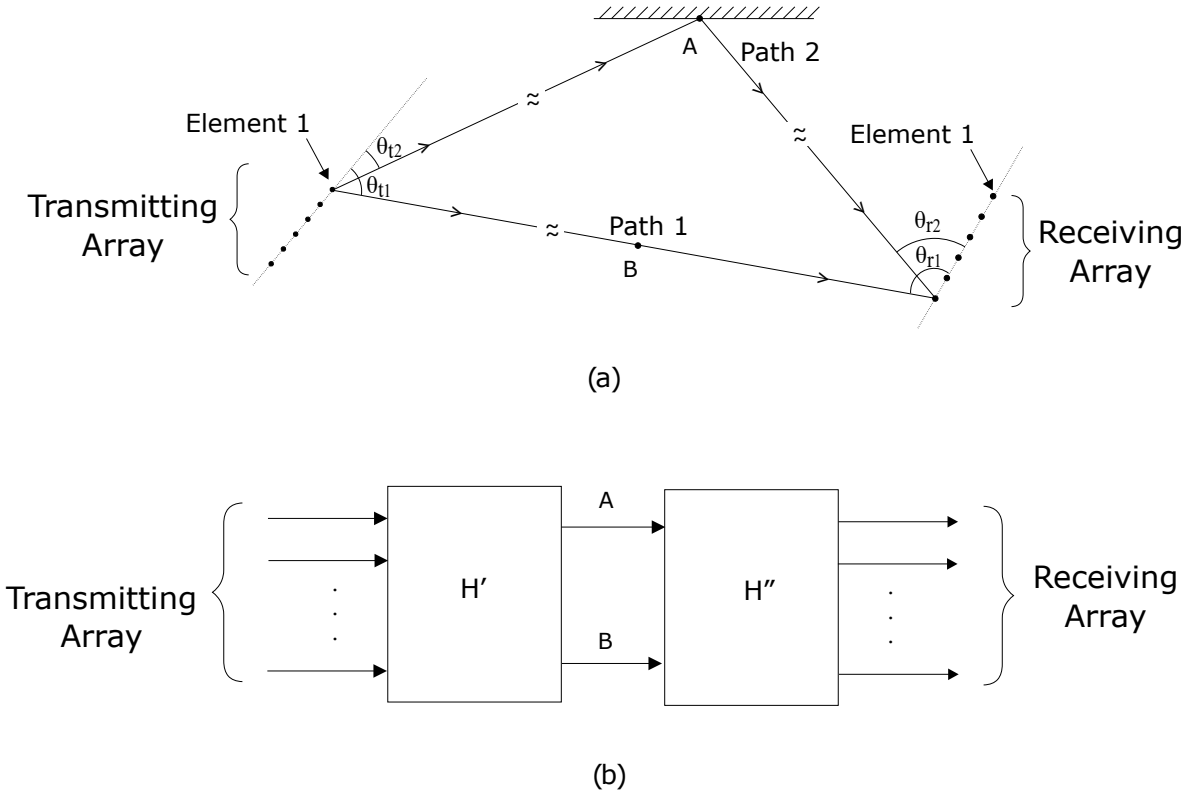


Figure 1.5: (a) Example of physical MIMO channel with 2 paths. (b) Equivalent MIMO channel by matrices cascading [4].

where θ_{t_l} and θ_{r_l} are the angle of departure (AOD) and AOA at l-th path, respectively.

Theoretically, the MIMO system has the ability to scale the capacity of a wireless connection by leveraging the multi-path effect of the wireless channel. This is due to multiple signal streams that can be transmitted in the wireless channel by leveraging the multi-path effect. In particular, we can denote the overall MIMO channel as $\mathbf{H} = \mathbf{H}''\mathbf{H}'$, where the block diagram of this model is shown in Fig. 1.5(b). Considering the beamforming techniques applied at Txs and Rxs, where the weights in the processing matrix at Txs/Rxs are usually called precoding/decoding matrix in the literature, the overall MIMO output signal vector can be expressed as

$$\mathbf{y} = \mathbf{P}_{deco}\mathbf{H}\mathbf{P}_{preco}\mathbf{s} + \mathbf{P}_{deco}\mathbf{n}. \tag{1.22}$$

Notice that compared with the beamforming in equation (1.19) where only one radio frequency (RF) chain is employed to have one output value, there are usually multiple RF chains in MIMO system. Hence there are multiple columns in precoding matrix, $\mathbf{P}_{preco} \in \mathbb{C}^{M \times M_{preco}}$, and multiple rows in decoding matrix, $\mathbf{P}_{deco} \in \mathbb{C}^{M_{deco} \times M}$, where M_{preco} and M_{deco} are the number RF chains for precoding and decoding. Moreover, $\mathbf{s} \in \mathbb{C}^{M_{preco} \times 1}$ is raw signal vector before precoding and $\mathbf{y} \in \mathbb{C}^{M_{deco} \times 1}$ is the output signal vector after decoding. By applying the

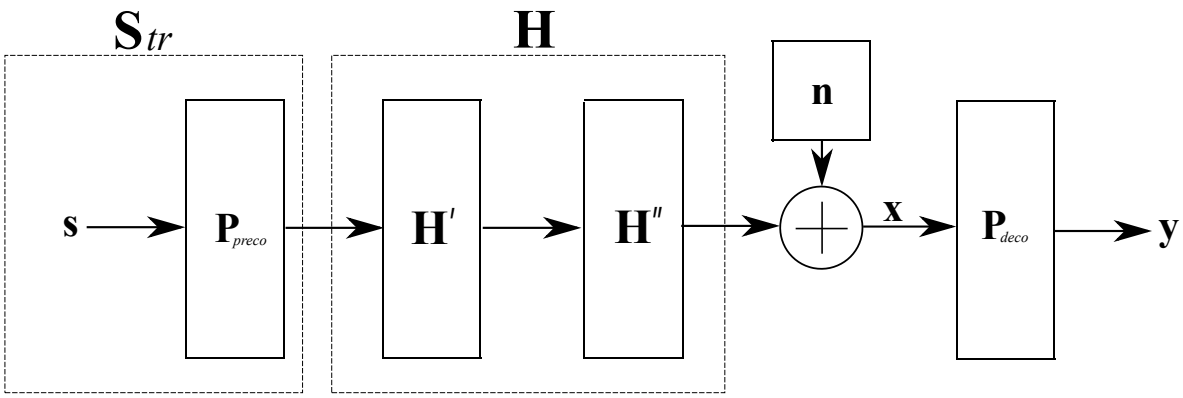


Figure 1.6: The block diagram of MIMO system.

singular value decomposition (SVD) to the channel matrix $\mathbf{H}$ above, we have

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H, \tag{1.23}$$

where $\mathbf{U} \in \mathbb{C}^{M\times M}$ and $\mathbf{V} \in \mathbb{C}^{M\times M}$ are (rotation) unitary matrices and $\mathbf{\Lambda} \in \mathbb{C}^{M\times M}$ is the matrix whose diagonal elements are nonnegative real numbers and whose off-diagonal elements are zero. The diagonal elements $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n_{\sin}}$ are the ordered singular values of the matrix $\mathbf{H}$, where n_{sin} is the number of non-zero singular number of the matrix $\mathbf{H}$. Since we have $\mathscr{L}$-paths and

$$\mathbf{H} = \sum_{l=1}^{\mathscr{L}} \mathbf{a}\left(\theta_{r_l}\right)\mathbf{a}^H\left(\theta_{t_l}\right) = \sum_{i=1}^{n_{sin}} \lambda_i \mathbf{u}_i \mathbf{v}_i^H \tag{1.24}$$

where singular values $\lambda_1 = \lambda_2 = \cdots = \lambda_{n_{\sin}} = 1$ and $n_{sin} = L$. It can be seen that the rank of $\mathbf{H}$ is precisely the number of paths in the wireless transmission environment. If the number of RF chains activated at both Txs and Rxs is larger than the number of paths in the transmission environment, let

$$\begin{aligned} \mathbf{P}_{preco} &= \mathbf{V}^H, \\ \mathbf{P}_{deco} &= \mathbf{U}^H, \\ \tilde{\mathbf{n}} &= \mathbf{U}^H\mathbf{n}, \end{aligned} \tag{1.25}$$

then the output signal vector can be written as

$$\mathbf{y} = \mathbf{\Lambda}\mathbf{s} + \tilde{\mathbf{n}} \tag{1.26}$$

where $\tilde{\mathbf{n}}$ has the same distribution as $\mathbf{n}$. Therefore, L streams are transmitted in parallel and

the capacity of MIMO system is

$$C_{MIMO} = \sum_{i=l}^{\mathscr{L}} \log\left(1 + \frac{P_l \lambda_i^2}{N_0}\right) \quad \text{bits /s/Hz} \tag{1.27}$$

where P_l is the power allocations at l-th path, satisfying the total power constraint $\sum_l^{\mathscr{L}} P_l = P$, and N_0 is the power of noise. As can be seen clearly, it is the number of paths that scale the capacity of MIMO system. To conclude, with more scattering/reflecting paths in the environment, more diversity in the space can be leveraged to improve the transmission rate.

1.1.5 Evolution of Beamforming/MIMO

The beamforming and MIMO have both been proposed for different application scenarios but the evolution of them are gradually merging with the development of wireless communication. In particular, the beamforming technique is initially proposed for directional transmission or reception, which requires the LoS condition. For MIMO technique, it is the multi-path effect from the NLoS that support the multi-stream transmission.

In terms of the origin of the beamforming technique, the German inventor and physicist Karl F. Braun demonstrated beamforming for the first time in 1905. At that time, Braun created a phased array by positioning three antennas to reinforce radiation in one direction and diminish radiation in other directions [21]. Nevertheless, only the directional radiation on the beam was addressed. The idea of using electromechanical devices for beam steering had not been clearly formed. After, the phased array with full electromechanical control is proposed, as most transmitting/receiving directional signals used at that time relied on the antennas' physical configuration and motion [22]. As beamforming is created much earlier than the MIMO technique to increase the signal scanning and receiving, it has been widely used in early warfare for eavesdropping purposes and radar detection.

The propose of MIMO technique originates from 1970s research papers, concerning multi-channel digital transmission systems and interference between wire pairs in a cable bundle [23, 24]. Some of the mathematical techniques for dealing with mutual interference in these works are found to be useful to MIMO development, though these are not examples of exploiting multi-path propagation to send multiple information streams. In the mid-1980s Jack Salz at Bell Laboratories took this research a step further, investigating multi-user systems such as time-division multiplexing and dually-polarized radio systems [**?**]. Later in 1991, Richard Roy and Björn Ottersten proposed space-division multiple access (SDMA), using directional or smart antennas to communicate on the same frequency with users in different locations within the range of the same base station and patented in 1996 [25]. Then,

Arogyaswami Paulraj and Thomas Kailath proposed an SDMA-based inverse multiplexing technique in 1993. Arogyaswami and Thomas described a method of broadcasting at high data rates by splitting a high-rate signal "into several low-rate signals" to be transmitted from "spatially separated transmitters" and recovered by the receive antenna array based on differences in "directions-of-arrival ", which is at the heart of the current high-speed WiFi and 4^{th} generation (4G) mobile systems.

Nowadays, in the 2020s, beamforming and MIMO have been discussed simultaneously quite often in most current research due to they both leverage the multi-antenna set-up. Specifically, the signal processing in beamforming was further extended from using analogue phase shifters to multiple RF chains for digital beamforming, which can realize multi-beam transmissions [26]. In this scenario, carriers can be transmitted and received from multi-path in the environment, which is equivalent to the MIMO transmission. Hence, the hardware set-up for digital beamforming can be equivalent to MIMO precoding and decoding, inducing similar systematic signal processing. Later, hybrid beamforming is considered to reduce the number of RF chains connected to each antenna by combining with more low-cost phase shifters to compress the cost overhead.

1.1.6 Challenges

As we can see, relative techniques have been proposed to combat various issues in the wireless channel. Meanwhile, improving the transmission performance to meet the ever-increasing quality of service (QoS) requirement is also urgent.

Due to limited spectrum resources in low-frequency bands and more spectrum resources can be harvested in high-frequency bands, mmWave and higher frequency bands have been considered in 5G and more advanced wireless networks [27]. For these frequency bands, some characteristics are intrinsic. For instance, the exponent measured in the corridor, where scattering and reflecting paths are able to form the so-called guided-wave effect [2] such that $n < 2$, which indicates a worthy transmission scenario of mmWave and can be seen from the measurements result in Fig. 1.2. However, it is conventionally acknowledged that mmWave carriers are not suitable for long-distance transmission due to huge path loss and vulnerability to blockages in the LoS transmission. The serious channel path-loss for high-frequency bands has to be compensated by a large-scale antenna array (or directional antenna) at the transceivers [13]. In this case, the cost of extra required space, power and complexity is also a non-trivial issue for multi-antenna beamforming. The relay technique is another typical approach to compensate for the channel path-loss but faces similar cost problems as multi-antenna beamforming [28].

In addition, the multi-path effect unavoidably produces channel fading and inter-symbol interference (ISI), which is detrimental to the communication quality. To overcome this effect, elaborate techniques have been proposed across all modules in the baseband procedures, e.g., multi-carrier systems such as orthogonal frequency-division multiplexing (OFDM) [29], cyclic prefix (CP), advanced channel estimation and equalization algorithms [30,31], adaptive modulation and coding schemes [32,33], and various detection algorithms [34,35]. Nevertheless, the signal processing algorithm, protocol, and system complexity within the transceiver are extremely high, which also causes overhead on the cost.

The theoretical MIMO have its limitation as MIMO requires channel state information (CSI) from channel estimation. Realistic wireless channels are tricky to be predicted precisely and there are always uncontrollability and randomness in the wireless channel [36,37]. In addition, the scalability of MIMO system is seriously limited due to the cost of RF chains [38]. Moreover, it is well known that the channel matrix of a MIMO system is fully determined by the communication environment. Since it has been invented, the system designing problems have focused on how to design a precoder at the Tx and/or decoder at the Rx, such that the dimension/capacity of the transmitted signal is maximized (i.e., optimizing the multiplexing gain), and meanwhile minimizes the difference between the estimation and the transmitted signal. However, in terms of multiplexing gain, the precoder and decoder cannot improve the channel rank even with full channel knowledge. On the other hand, enforcing orthogonality causes a power loss and thus a beamforming gain limitation (e.g. decomposition reduces the strength of eigenvalues in the eigenvectors). Therefore, though MIMO still improve the system performance based on the CSI, it does not actively change and improve the quality of the wireless channel.

To conclude, in order to improve spectral efficiency (SE)[6] and energy efficiency (EE)[7] over the wireless channels, previous works have been extensively devoted to combating the wireless channel's loss and fading. However, with the explosive ever-increasing capacity and QoS demand, it is difficult to realize the desired target with unavoidable system complexity, cost, and the uncontrollable wireless channel.

1.1.7 Motivations

Barriers for further improving the spectral efficiency and energy efficiency are tricky but the communication system have had to put more cost for responding. The m-MIMO will incur

[6]Spectral efficiency, spectrum efficiency or bandwidth efficiency refers to the information rate that can be transmitted over a given bandwidth in a specific communication system. It is usually measured in bit/s/Hz [39].

[7]The energy efficiency metric, measured in bit/Joule, which describes how much energy is consumed per correctly received information bit.

a significant system complexity while merging the OFDM transmission scheme, requiring effective implementation cost on the hardware. In addition, the cost for having massive antennas and RF is another critical issue. Moreover, base station (BS) will be deployed in a much denser way to mitigate the signal coverage issue. Therefore, the wireless communication industry can hardly further improve the overall transmission performance to the next generation without solving the overhead on scaling up the dimension of system complexity and deployment density.

As communication systems tend to be more complicated, and the cost of leveraging spectrum resources keeps increasing, urgent steps are needed to fundamentally improve wireless transmission performance with reduced system complexity and low cost. One effective way to lower the complexity and cost is to reduce the use of energy harvesting components such as RF chains. Thus, the structure of transceivers is usually the object to be engineered. For instance, the hybrid beamforming/MIMO structure on the device is proposed to reduce the number of RF chains required while maintaining a desired QoS [38, 40]. However, the intrinsic issues of wireless channels are not therefore overcome. In particular, all designs on the transceivers are and will always be subject to the realistic wireless channel. The channel fading and path loss are never tackled directly at all.

By transferring the designed object from the transceivers to the wireless channel, a natural intuition can be obtained to combat the uncontrollable and unknown wireless channel by changing it directly. For instance, the channel fading is mitigated or even cancelled by designing the wireless channel. The channel path-loss is carefully compensated by the reconstructed multi-paths in the transmission environment. With this conjecture, the wireless channel can even be designed into a filter such that multi-stream of MIMO transmission can be received orthogonally without precoding and decoding. Since recognition in previous works does not consider the possibility of controlling the wireless channel, there is no doubt how significant a revolution in wireless communication can be if we can firmly control the wireless channel.

1.2 The Intelligent Reflecting Surface

The recent advances on EM materials found the possibility to change the wireless channel from a new type of retrofitted low-cost material [41], usually called meta-material [42]. The artificial thin film of this material is usually referred to as reconfigurable intelligent surface (RIS) [43], intelligent reflecting surface (IRS) [44], or meta-surface [45]. It is worth mentioning that in both acoustics and optics, there can be some duplicated terminologies [46–48].

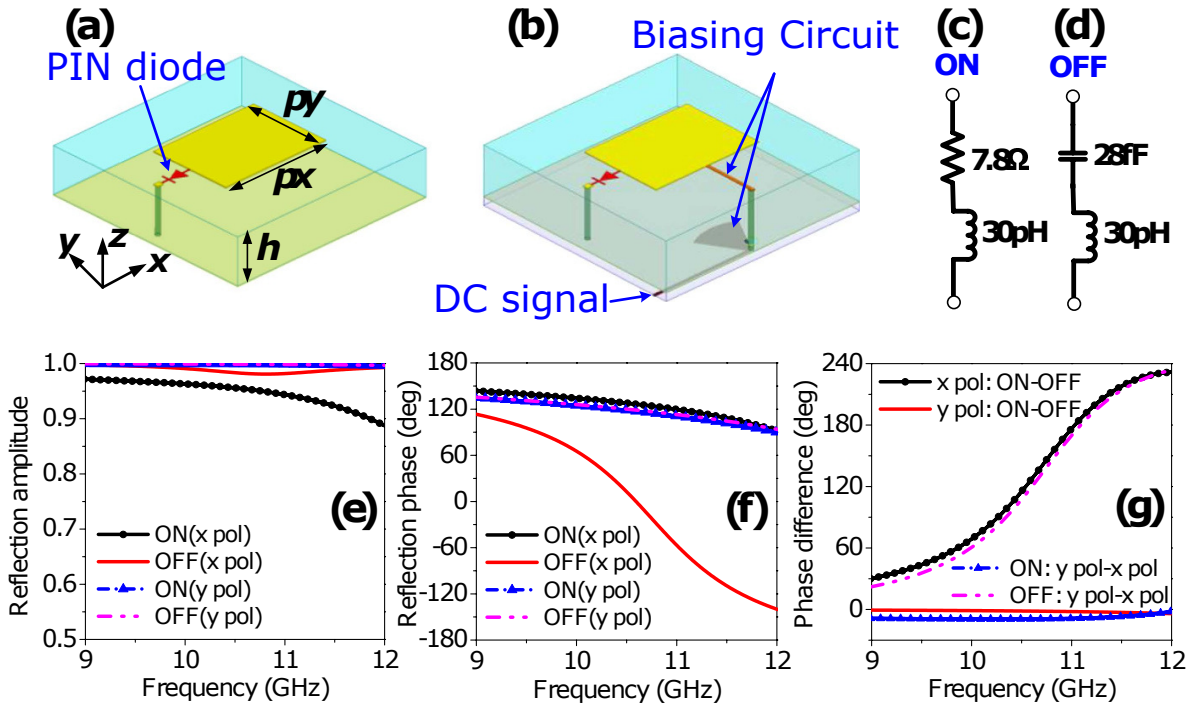


Figure 1.7: Proposed element structure and its coding property [5]. (a) Schematic of the element, with specific dimensions. (b) Proposed element with actual biasing architecture. The biasing circuit is elaborately designed to isolate the DC and RF signals. (c,d) The equivalent circuits of the PIN diode at ON and OFF states, respectively. (e) Reflection amplitudes. (f) Reflection phases. (g) Reflection phase differences.

For wireless communication, it is often used in the literature to use IRS[8] and RIS.

1.2.1 IRS Structure

An IRS can be a very simple 2D structure, comprising a set of conductive patches, diodes, and conductive power/signal lines. Specifically, the work in [5] presented a typical element structure, which is the basic unit on the thin film, made up by positive-intrinsic–negative (PIN) diode, and the corresponding biasing circuit is illustrated, as shown in Fig. 1.7.

The schematic of the proposed element for IRS is presented in Fig. 1.7(a), which has a sandwich structure composed of a simple rectangular patch and a metal-ground plane spaced by a substrate with a certain dielectric constant and loss tangent [49]. A PIN diode[9] is employed to connect one edge of the patch to the ground through a metal via. Thus an anisotropic element with binary coding reflective performance is obtained along the x direction. To facilitate the biasing in practical implementations, a direct-current (DC) circuit in the element topology is introduced, as shown in Fig. 1.7(b). The deliberately designed bias circuit

[8]Without loss of generality, IRS is the term that will be used in the thesis for the following discussion.

[9]Using PIN diode is only one particular setting, thus more switching techniques can be leveraged base on other hardware components to achieve higher bit controlling [50].

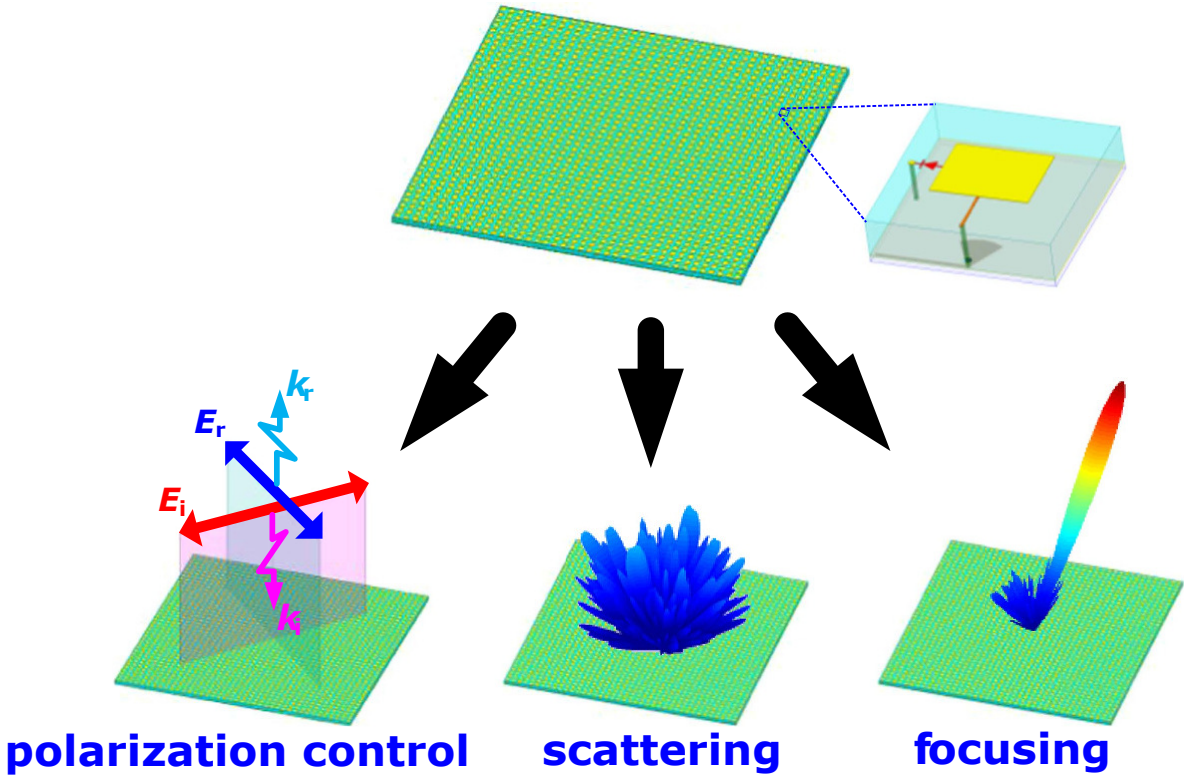


Figure 1.8: Exemplification of multiple functions of IRS [5].

includes the quarter-wavelength microstrip line, the open-ended radial stub and DC signal line. The quarter-wavelength microstrip line and the open-ended radial stub are employed to choke the RF signal and ensure a good isolation between the DC and RF performance. The PIN diode is modelled as an equivalent circuit shown in Fig. 1.7(c,d) when biasing circuit is switched ON or OFF, respectively. Due to the binary controlling of the biasing circuit, the element can manipulate the reflected EM wave with different characteristics to the impinging EM wave, as shown in Fig. 1.7(e,f). It is observed that almost total energy is reflected for both x- and y-polarized incidences. The slightly higher energy loss at the ON state under x-polarized incidence results from the large equivalent resistance shown in Fig. 1.7(c).

As seen in Fig. 1.7(f), due to the asymmetrical integration of the PIN diode, a distinct reflection phase is obtained for the x-polarized incidence when biasing the PIN diode ON or OFF. Moreover, the reflection phase for y-polarized incidence is very similar to that for x-polarized incidence at the ON state.

Fig. 1.7(g) further plots the reflection phase differences. Apparently, the effective phase difference is observed for x polarization with different PIN diode states and for PIN diode working at OFF state with different polarizations. For both cases, a perfect binary coding phase (180° phase difference[10]) is achieved at the design frequency of 11.1 GHz. Controlling

[10]The ON-OFF control on the biasing circuit to realize binary coding on the phase is also called 1-bit con-

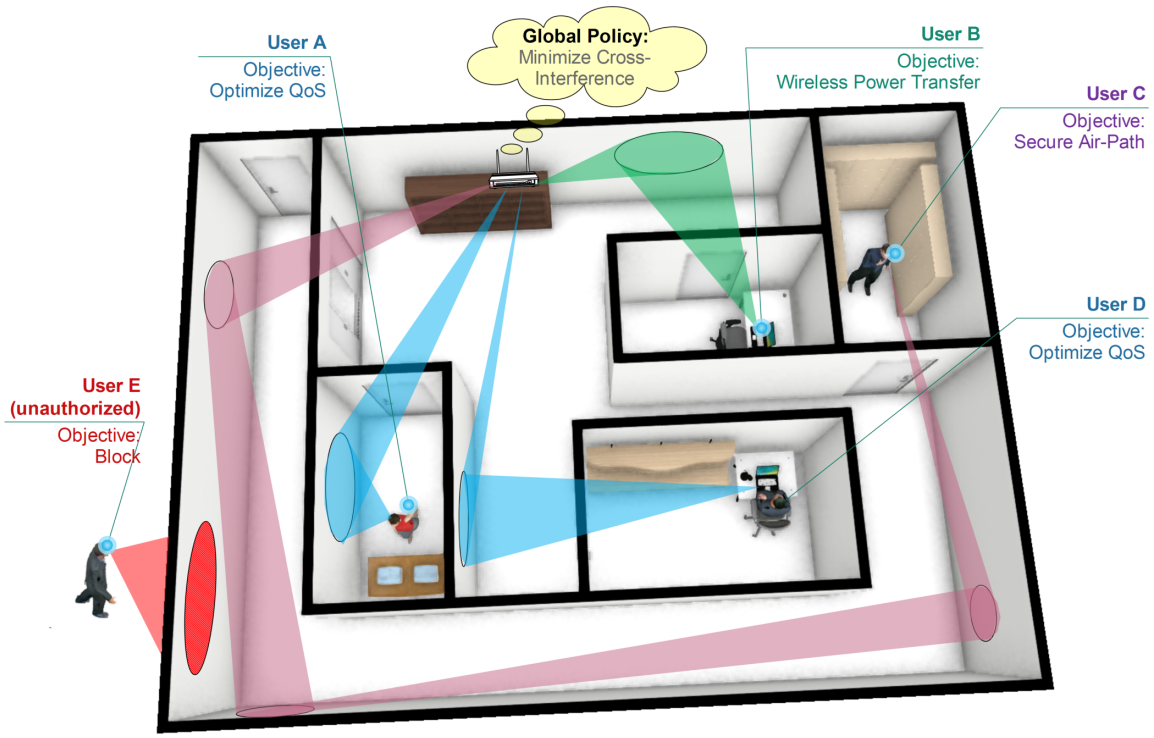


Figure 1.9: The indoor IRS application scenario [6].

the phase of the element can be arbitrary, and the amplitude of the reflection can also be adjusted, though this may require other designs for the hardware.

It is also observed that multiple elements are packed into an array-like structure to formulate a surface. By integrating multiple elements in the array-like surface of meta-material, IRS can realize various manipulating functions to the impinging EM wave, as shown in Fig. 1.8 [5]. In particular, the reflected EM wave can be manipulated with a specific polarization statement, power scattering and focusing in an arbitrary direction. To realize a corresponding function, the excitation and phase of each electronic element should be controlled by processors in a designed controlling manner, e.g., field programmable gate array (FPGA) [41, 51], such that the phase and amplitude of the EM wave impinging on the surface can be manipulated and reflected any desired directions, which lays a foundation of the programmable wireless channel.

1.2.2 Application Scenarios

From the perspectives mentioned in the sections above, it is the first time we can (at least partially) control/program the wireless channel. Due to its ability to manipulate the EM wave

trol/weight. Note that the element designed in [5] is one of the designed examples in many designs. Higher bits' weight can be realized with more advanced circuit design. The continuous weight is the general case of quantized bits control.

and the communication channel, IRS is expected to be a driving technology of 6^{th} generation (6G) and leading a disruptive evolution [52].

With the deployment of IRS, the wireless channels within the natural environment become controllable with such surface posting on the basic building block (e.g., can even replace the traditional surface of the ceiling and wall). In other words, those objects, e.g., buildings or indoor walls, which bring negative effects to the signal transmission, can turn into something helpful to produce more design degrees of optimization with the coating of such IRS on those objects [45]. Thus, IRS can be used in many scenarios of wireless communications. In particular, some interesting scenarios have been introduced by the work [6], with full IRS indoor deployment to all walls on one floor, as shown in Fig. 1.9. There are five applications mentioned in [6], based on different weights design on IRS. As introduced in Chapter 1.2.1, the beam focusing by the reflection of IRS can be used to increase signal-to-noise ratio (SNR) for obtaining better QoS, which corresponds to the case of user A and user D in Fig. 1.9. It is also worth mentioning that MU interference can also be mitigated to improve QoS as well [53]. In addition, the case of user B indicates that devices of internet of things (IoT) can also benefit from beam focusing for wireless power transfer [54]. For the case of user C, multiple pieces of IRS are leveraged to let the transmitted signal bypass any possible eavesdropping devices in the physical transmission path. To achieve this function, these IRSs require a corporation such that signals can propagate via multiple times of reflect beam focusing and steering. And the case of user E indicates that the IRS based programmable environment can block any unauthorized users from transmitting/receiving any signals to/from the access point (AP). To realize a blocking function to specific users, IRS can implement beam scattering to minimize the unauthorized user's transmit/receive power or rely on nullification algorithms, such as zero forcing (ZF) [55]. Thus, the blocking function from IRS also shows its sufficient potential in physical layer security [56]. Moreover, not limited to the above scenarios where IRS is used as a reconfigurable reflector, it is also interesting that IRS can be used as a passive information transmitter with modulation [57], which is similar to the backscatter communication [58, 59]. In particular, the feed signal from an active transmitter is required in this case and the information is encoded into the amplitude and phase of elements on the IRS such that a receiver can obtain the encoded information from the IRS passively [60].

From the architecture perspective, it is worth noting that the IRS is similar to a multiple antenna relay system [61]. However, their principles, challenges, and applications of them are different. The IRS reflects impinging signals (at some specific frequency bands) as an object rather than re-transmit the received signal as a Tx in the relay systems [28]. From this point of view, it is an EM mirror that can reflect the signal to arbitrary direction by controlling the phase/amplitude of the elements on the surface. Unlike the active relay system that relies

on multiple RF chains, the IRS is passive so that it can be made as thin as a wallpaper to be posted on any physical surface and, therefore it is low cost, low power, flexible, and scalable [6]. Besides, the active relay system typically works in a half-duplex manner to avoid interference while IRS has no such constraint (i.e., it can achieve full-duplex function without generating interference) [45, 62]. In [63], the IRS is shown to support a high achievable rate with better energy efficiency than the relay system. As such, passive surfaces are different from the active relay, which is more efficient in leveraging power and spectrum resource.

To conclude, due to the properties of the IRS that can be manipulated in real-time by software programming, it has been considered a quite promising study area for wireless communication since IRS fundamentally leverages, changes and improves the wireless channel. From this perspective, it is the first time for us to be able to (at least partially) control/program the wireless channel to achieve the desired communication environments.

1.3 Literature Review

As IRS has been proposed for supporting various applications in wireless transmission, previous IRS research focused on a wide range of topics, such as beamforming, channel estimation, resource allocation and hardware implementation, etc. Within these works, many subjective issues and objective issues are also considered as well. The subjective issues of IRS are those which can be controlled by designers. Specifically, the power constraint of IRS, the control-bit constraint and the number of IRS, which can be subjectively determined. Those objective issues can not be controlled subjectively but can be considered prior to the transmission design. In particular, the number of transceivers, the distance of transceivers between IRS and transceivers, and hardware impairment etc., which objectively existed with respect to the IRS.

To further classify, we review from following four transmission schemes, where relative subjective and objective issues are addressed, respectively.

- Single IRS assisted single user (SU) transmission.
- Single IRS assisted MU transmission.
- Multiple IRSs assisted SU transmission.
- Multiple IRSs assisted MU transmission.

1.3.1 SU Transmission Based on Single IRS

Many literature about the wireless transmission of IRS originated from the design for SU scenario and single BS. The work by Yan et al. [64] proposed to maximize the received power of SU under multi-input-single-output (MISO) scheme by designing the passive beamforming phase shift on the IRS while also considering the design the information transfer from IRS to BS by controlling the ON-OFF states of elements simultaneously. The work in [65] maximized the received power and propose the manifold optimization to efficiently deal with the constant modulus constraint. In the work of [66], the SU channel capacity is analyzed by leveraging the global co-phasing technique to design the phase of IRS while considering jointly optimization scenario with MIMO. In addition to that, the IRS based channel estimation is proposed [67], where the low complexity is the main focus. Dadari et al. [68] leverage the grouped tile to implement the localization of SU based on OFDM downlink system. The work in [69] analyze the scalability of IRS beam steering in terms of the directivity, target deviation, half-power beamwidth (HPBW), and side-lobe level (SLL) under different element spacing, size of IRS and controlling bits. Thus, in this case, the weights need to be specifically designed, which is similar to some considerations in traditional precoding/decoding design [70].

The experimental realization of IRS has also been achieved in different applications. Zhang et al. [51] implemented space-time modulation of quantized IRS to manipulate EM waves in both space and frequency domain. The work in [71] proposed to minimise the desired energy distribution with optimized energy distribution for Wi-Fi signals. With quantized weights, Gao et al. validated the broadband scattering with numerical simulations and experiments using 2-bit coding IRS in terahertz wave (THz) band [72].

A few objective factors, such as near-field effect [73], in the IRS aided wireless transmission are also considered. As done in many previous non-IRS works [74] of near-field, the second order Taylor expansion on the distance of the channel model is common. This approximation in higher order has its advantage in a closer range than the first order expansion in the far-field assumption. In [75], Kennedy et al. addressed an equivalence between the near and far field's beampattern, by solving the beampattern coefficients from wave equation analytically under spherical coordinate and realizing broadband beamforming with the traditional far-field array design. In [76], Degli-Esposti proposed a diffuse scattering model of a reflecting surface in the near-field as a ray-tracing tool. With previous knowledge, the work in [77] establish the general path loss model within the far-field and near field, respectively, where the results are validated with experiment measurement. In [78], Björnson et al. derive IRS's channel gain with near-field effect and the power scaling law with growing on the number of elements, which addressed the limitation of far-field approximation-based theorems in

this case, where the critical boundary to differentiate the far-field and near-field is revised in [79]. It is worth to mention the effective aperture of IRS has been discussed in [80].

Except for works dealing with the wave reflection on the IRS, some works considered that IRS could have a wave penetration feature. In particular, for the wave impinging on the element with penetration feature, the energy of the wave is divided into two parts, where one part is reflected out as usual, and another part is penetrated to the back of the IRS and keep transmitting forward, which is similar to a beam splitter in optics [81]. This specific feature is commonly called simultaneously transmitting and reflecting (STAR) and such a concept is firstly proposed in work [82], where Liu et al. made a preliminary review and proposed the signal models of STAR-IRS.

1.3.2 MU Transmission Based on Single IRS

For single IRS aided system, dealing with MU is more general but more challenging. A single IRS can either contribute to an improvement of QoS for users in some area covered by IRS while it can also cause worse QoS for some other users. The work in [44] applied semidefinite relaxation (SDR) technique to obtain a high-quality approximate solution to improve the MU downlink transmission with constraint modulus constraint. Meanwhile, the IRS assisted communication system has been considered for EE maximization and weighted sum rate maximization [83, 84].

Nadeem et al. [85] focuses on the downlink of a single-cell MU system, where the minimum signal-to-interference-plus-noise ratio (SINR) is maximized, subject to a power constraint of IRS. MIMO model of IRS also have been proposed and analyzed by using well-known minimal mean-squared error (MMSE) and ZF beamforming [44].

Di et al. [86] studied the hybrid beamforming assisted by IRS with discrete phase shifts. In particular, the sum rate maximized for MU where the continuous digital beamforming has been performed at the BS, and the discrete analog beamforming has been achieved inherently at the IRS. The formulated problem is decomposed into subproblems and solved iteratively to joint optimize the hybrid beamforming and passive beamforming.

For channel estimation and beamforming design, the work in [87] designed the MU downlink estimation protocol and maximized the sum rate, where an additional focus is put on the hardware impairments, e.g., phase noise, on transmitters and IRS. The constant modulus constraint is also considered and solved by the projected gradient ascent algorithm. The work in [88] proposed an effective channel estimation protocol to estimate the cascaded channels directly using the compressive sensing (CS) technique. Specifically, the channel estimation problem is formulated into a sparse channel matrix recovery problem in order to achieve

robust channel estimation with limited training overhead. He et al. [89] achieved cascaded channel estimation of the IRS-assisted MIMO systems.

For MU beam sweeping, the beam training is designed in [90] to implement single beam steering for SU and multi-beam steering for MU. In particular, the training is achieved by dividing IRS elements into multiple sub-arrays and properly designing sub-array beam directions over different training symbols with users' independent beam identification based on received power/SNR comparisons.

With MISO non-orthogonal multiple access (NOMA), the work in [91] maximized the downlink sum rate by jointly optimizing the passive beamforming, decoding order and power allocation coefficient vector. The formulated problem is solved by a three-step approach, which leverages three machine learning algorithms. The work in [92] investigated the MU capacity limits for NOMA and orthogonal multiple access (OMA) transmission schemes. The IRS reflection matrix and resource allocation were jointly optimized for characterizing the Pareto boundary of the capacity and rate regions under the constraints of discrete phase shifts and a finite number of IRS reconfiguration times.

Another work with MISO scheme considers optimizing the EE by gradient-based energy efficiency maximization algorithm and deep learning algorithm to design phase control [43]. In [93] Najafi et al. proposed the physics-based modelling to analyze the effect of elements-pack-to-tile's grouping for the large IRS surface and proposed the codebook-based joint optimization of BS and IRS to minimize the power consumption.

The work in [94] investigated MU passive beamforming via IRS. In this work, the sum transmitted power of the network is minimized by controlling the phase beamforming and transmit power of BS. This problem is formulated jointly and solved by the dual method. It is worth mentioning that this work proposes the use of IRS as the transmitter. In particular, the IRS is deployed close to the BS such that IRS and BS can be seen as a whole. This is different from most other works that consider IRS is deployed in the middle of the transmission channel to let the transmitted signal bypass the blockages.

For STAR-IRS, the work in [95] further studied the channel fading model for the desired signal and interference signal, respectively. In addition, the coverage probability and the ergodic rate for users to receive the reflected signal and the penetrated signal are derived, and STAR is evaluated to have a better performance than traditional IRS given a specific range of energy splitting coefficients. In the work [96], multi-cast downlink transmission is investigated, where the power consumption is minimized to achieve the required communication rate constraints with penalty-based algorithms. The work in [97] maximizes the sum coverage range of two users, where one user stand in front of IRS and another one stand in the back or IRS under and OMA and NOMA with QoS constraint constant modulus con-

straint, and power constraint. In terms of security, the work in [98] maximize the secrecy rate and minimize the secrecy outage probability with full eavesdropping CSI and statistical eavesdropping CSI. Further, the effective channel estimation is achieved by minimizing the mean-squared error (MSE) with weights designs in separate time intervals, where element grouping (adjacent elements share the same transmission/reflection coefficient) is leveraged to reduce the channel estimation overhead [99].

Though it is common to consider one BS supporting MU, transmission with multiple Txs and Rxs can also be a common scenario with massive devices having device-to-device (D2D) communication via the aid of a single IRS [100]. The work in [101] maximizes the sum rate of downlink Rxs while involving the sum rate from D2D communication devices assisted by a single IRS at the same time. The resource allocation design is formulated as a non-convex optimization problem while taking into account the QoS requirement, the power allocations, and the limited backhaul capacity.

1.3.3 SU Transmission Based on IRS Networks

An IRS network, which is defined as deploying multi-piece IRS in the transmission environment, has been studied to further enhance the EE and SE. The throughput of a SU has been maximized by the IRS network leveraging the supervised learning approach [102]. Considering the multi-order-reflection (MOR) [103–105], the authors of [106] analyzed the single-user multi-order-reflection (SUMOR) transmission in one path of the IRS network and then provided the beam routing solution.

1.3.4 MU Transmission Based on IRS Networks

MU transmission via IRS network is also investigated, considering minimizing the power consumption of transmit beamforming with constraints of the power supply, SINR of each Rx, and constant modulus [107]. The work in [108] maximized the weighted sum rate with IRS cooperatively implement passive beamforming. In [109], the statistical path loss model of a large-scale IRS network is derived, where user (UE) and BS density are considered to derive the area of blind-spots.

The authors of [110] derived the lower bound of the MU average SINR by considering the rayleigh fading channel in the IRS network. The wideband transmission of MU has further been designed to maximize the sum rate with limited power and constant modulus constraints in the IRS network [111]. To realize a decentralized IRS network, the authors of [112] proposed distributed scheme of IRS networks to maximize the MU weighted sum rate. Additionally, the IRS network has been proposed to realize robust, secure MU communication by

jointly designing the transmit beamforming, artificial noise and IRS network [56].

Under MISO scheme, the work in [113] characterized the angle estimation from imperfect location information, derived a closed-form expression for the achievable rate of MU downlink transmission and proposed the beamforming via minimizing the transmit power. Further, the authors gave a tutorial for optimizing the wireless channel of one reflection to multi-user multi-order-reflection (MUMOR) transmission [114].

1.3.5 Research Questions

As can be seen in the literature review, most works have achieved analysis and design of SU and MU transmission based on single IRS with various subjective and objective issues. Within those works, it is common that the transmission channel is assumed to be rich scattering, which is usually applied for low-frequency transmission. In this case, only a limited amount of multi-paths will be affected by IRS and traditional design without knowing the IRS can still work effectively. Hence, IRS may not play a dominant role in supporting low-frequency transmission for MU. Compared with rich scattering channels in low-frequency transmission, the channel ranks of high-frequency carriers are more limited due to the reduced multi-path effects. In this case, IRS is more critical for high-frequency band transmission, which has significant path loss and high sensitivity for blockages [115]. The transmission link can fail without IRS supporting the signal reflection. In addition, for supporting MU transmission, the design on IRS will be much more challenging on power maximizing and interference minimizing since the spatial diversity is more limited for multiplexing.

Note that fewer works focus on the passive beamforming techniques of IRS in a less scattering channel for high-frequency MU transmission. For single IRS, a well-considered and thorough analysis on realizing MU beamforming, as shown in Fig. 1.10 with multiple Txs and multiple Rxs in such a system, is unknown. For single IRS aided transmission, we mainly would like to know: how can IRS shape multiple beams passively for serving MU at the same time. Whether we can achieve interference-free transmission based on single IRS and how to design the weights in this case. What is the system limit for a single IRS, in particular, how many users can be served effectively?

Moreover, since the deployment cost of IRS is cheaper than deploying a BS, it is straightforward that we can increase the control of the wireless channel by deploying multiple IRS in the environment, as a few proposed in the literature. However, it is also noticed that the intrinsic nature of EM wave transmission in multiple IRS environments has been overlooked in the literature, i.e., the dual reflection of MOR signal between two reflectors. Dual reflection is a common phenomenon for reflectors having spatial correlation and has been

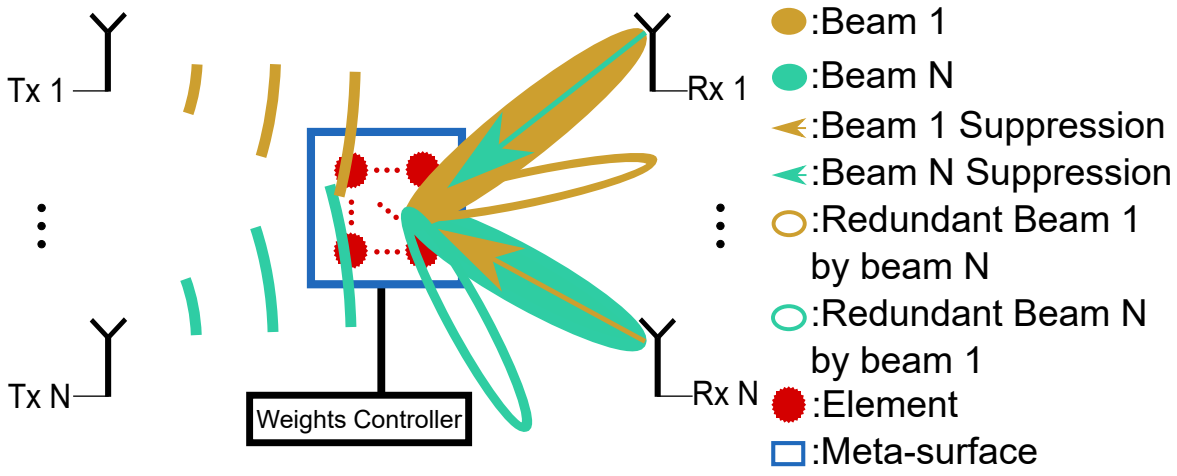


Figure 1.10: MU beamforming and transmission based on IRS, where peaks and zeros can be formed for desired signal and interference, respectively. Redundant beams will be generated for optimal beamforming to support MU transmission by a single IRS.

widely considered in radar systems [116–118]. In light of dual reflections, we notice most works about IRS only consider first-order-reflection (FOR). Though works [106, 114] further consider cascaded-nonline-of-sight (C-LoS) paths in MOR, the signal component via dual reflection is omitted in their models. To the best of the authors' knowledge, two main issues remained unsolved. First, no IRS works completely considered a channel model in the reflective environment with MOR. The above works ignore the presence of wave reflection within two IRSs in a much higher order, which is incomplete in establishing the transmission environment. Thus, establishing the complete model for the IRS network is necessary for analyzing generic and arbitrary reflecting scenarios. It is the most critical prerequisite to laying a foundation for a precise, robust, and reliable design for the IRS network. Further, no analytical works have indicated clear bounds to guide the deployment of IRS networks with MU interference, i.e., how much EE and SE can be improved, where the sum rate upper bound is and how to reach the upper bound.

While solving the questions raised above, it is also worth investigating subjective and objective influences on the system's performance in order to gain more insights and enhance the practical deployment of IRS. In particular, for subjective influences, we wonder how the constraints on weights of IRS, such as constant modulus constraint and multi-bit constraint affect the interference suppression performance from IRS. In addition, we wonder how objective influences, such as mutual coupling and near-field effects, affect the designed results.

By answering the questions above, it is believed that a straightforward improvement on MU transmission based on deploying IRS networks can be made to support robust links of high-frequency carriers. The current state-of-the-art EE and SE can be improved respectively

by minimizing the co-channel interference and maximizing the signal power received. Moreover, the insight can further lead to clear guidance on practical network deployment to further enhance wireless transmission effectively.

1.4 Original Contribution

In this thesis, the above questions are answered respectively. First, the design of the MU beamforming on the single IRS is given, which is the fundamental cornerstone. As IRS is limited in its physical size, the sum rate after MU beamforming is therefore limited. A clear limit of sum rate is critical for guiding the deployment of IRS and, more generally, the IRS networks with different topologies. With more practical effect, the bias for the MU beamforming design should also be evaluated. In particular, the contributions of the thesis are summarized as follows:

- The IRS assisted MU channel model based on a *single* IRS, which can support multi streams transmission, is first proposed. Additionally, fundamental limitations on the directions of the transceiver are derived for practical deployment.
- Based on the proposed model, IRS-assisted MU optimization problems are formulated. A closed-form amplitude-unconstrained phase-continuous (AUPC) solution is derived first, and then practical amplitude-constrained phase-continuous (ACPC) solutions are achieved by using sequential quadratic programming (SQP) method and regularized multi-user linearly constrained minimum variance (MU-LCMV) beamformer.
- The insights of the redundant beam have been exploited and discussed theoretically while considering the general uniform rectangular array (URA) scenario. This conclusion can provide a guideline for further designs on IRSs' deployment and resource management due to the redundant beams essentially splitting the energy.
- To incorporate the MOR effect with dual reflection, an index matrix is introduced to derive a complete model of the IRS network, which is applicable for arbitrary orders of reflections, an arbitrary number of IRS and arbitrary network topologies.
- Two critical conditions are mathematically derived : the optimal condition to reach the sum rate upper bound and the condition to realize interference-free transmission as insights for studying the EE and SE of the IRS network.
- Considering different topologies of the IRS network, the sum rate upper bound of MU-MOR transmission assisted by an IRS network is analyzed; by employing the optimal condition, the maximized EE and SE are derived, using graph decomposition.

- Involving the mutual coupling of IRS, the side lobes characteristics are analyzed and also discussed with the redundant beam influence. The weights equalization is proposed to mitigate the significant mutual coupling.

- The transmission with MU beamforming involving near-field effect is further designed. To alleviate the beamforming error with the near-field effect, we propose to implement beamforming algorithms with the segmented model.

- Lastly, with the near field, the spatial correlation of space in both angular and distance domains is analyzed, which is the essence of redundant beams in MU transmission based on a single IRS. The redundant beam is obtained numerically, where discriminant is derived to alleviate the complexity of an exhaustive search on positions of the redundant beam.

1.5 Thesis Outline

In Chapter 2, the basic model of IRS is introduced. Contents focus on modelling from its hardware prototype to its theoretical transmission model. Considering multi-piece IRSs can be deployed in the transmission environment, we further model the MU and multi-order-reflection (MUMOR) IRS networks. We derive a complete MUMOR IRS network model that is applicable for an arbitrary number of reflections, arbitrary size and number of IRSs/reflectors.

Chapter 3 is mainly written on top of "Multi-user Beamforming and Transmission Based on Intelligent Reflecting Surface" (the first journal publication in **List of Publications**). Contents focus on the MU beamforming based on a single IRS for interference-free transmission. We formulate an optimization problem called MU-LCMV beamformer under the criterion of minimizing the overall received signal power subject to a certain level of power response (e.g., unit power response) at desired signal directions and arbitrary low power response (e.g., zero power response) at the interference directions. A closed-form AUPC solution is derived first, then an ACPC solution is obtained by using SQP. Given the solutions, the IRS beam pattern shows that to achieve MU (N pairs of transceivers, $N > 1$) transmission through a single surface, up to $N-1$ redundant beams are generated, significantly affecting power efficiency. Therefore, we mathematically analyze the results of the redundant beam. The directions of the redundant beams are mathematically derived. The effect of mutual coupling on IRS is also analyzed to show the characteristic of side lobes. Simulation results verify the existence and accuracy of the redundant beam directions.

Chapter 4 is mainly written on top of "Intelligent Reflecting Surface Networks with Multi-Order-Reflection Effect: System Modelling and Critical Bounds," (the second journal publication in **List of Publications**), where the optimal condition for achieving sum-rate upper bound with one IRS in a closed-form function and the analytical condition to achieve interference-free transmission, are derived respectively. Leveraging this optimal condition, we obtain the MUMOR sum-rate upper bound of IRS network with different network topology, where the linear graph (LG), complete graph (CG) and null graph (NG) topologies are considered. Simulation results verify our theories and derivations on the sum-rate upper bounds of different network topologies.

Chapter 5 is mainly written on top of "Multi-user Beamforming of IRS with Near-Field Effect," (the third journal publication in **List of Publications**), where the analysis for considering objective effects on the IRS is developed. In particular, the coupling effect of the redundant beam results is analyzed. To enhance the beamforming performance in the near-field, the near-field effect is incorporated into the transmission model. The behavior of redundant beams in the near-field is also analyzed numerically. Simulation results verify the proposed theorems, and relative numerical results are given.

Finally, conclusions and an outlook on possible future work for IRS and IRS networks are given in Chapter 6.

Chapter 2

Fundamental Model of IRS and IRS Network

In this chapter, we first introduce the single IRS channel model first, i.e., the channel transmission model for transceivers. Then, we present the model for the basic element on IRS and last present the model of IRS networks. The ideal transmission assumptions are considered, i.e., the channel fading is considered to be quasi-static while the narrow-band signal is assumed to be transmitted in an ideal case of the far-field condition, as did in [44,64,67,85,89]. Nevertheless, these assumptions will be justified and specific analysis with practical concerns will be discussed in the following thesis. Following the single IRS model, the lumped circuit, to achieve the required amplitude-phase response of each IRS element, will be introduced as the insight of the hardware. Then, we will introduce the IRS network models and the corresponding MOR effect. We also assume all CSI of IRS networks above are known in the following chapters.

2.1 Basic Single IRS Model

As shown in Fig. 2.1, a single IRS model is considered to be embedded with M elements for N pairs of single antenna users' communication. The blockages between transceivers are considered, where the LoS paths are considered between the Tx to IRS and IRS to Rx, since NLoS condition is a long-lasting issue in mmWave communications [119–121].

Based on the transmission model given in Chapter 1, the signal impinges on the IRS can be expressed as

$$\mathbf{y}_s = \mathbf{A_{in}}\mathbf{s}\,, \tag{2.1}$$

where $\mathbf{s} = [s_1, s_2, \ldots, s_N]^T \in \mathbb{C}^{N\times 1}$ is the source signal vector and s_i $(i = 1, 2, \cdots, N)$ is the

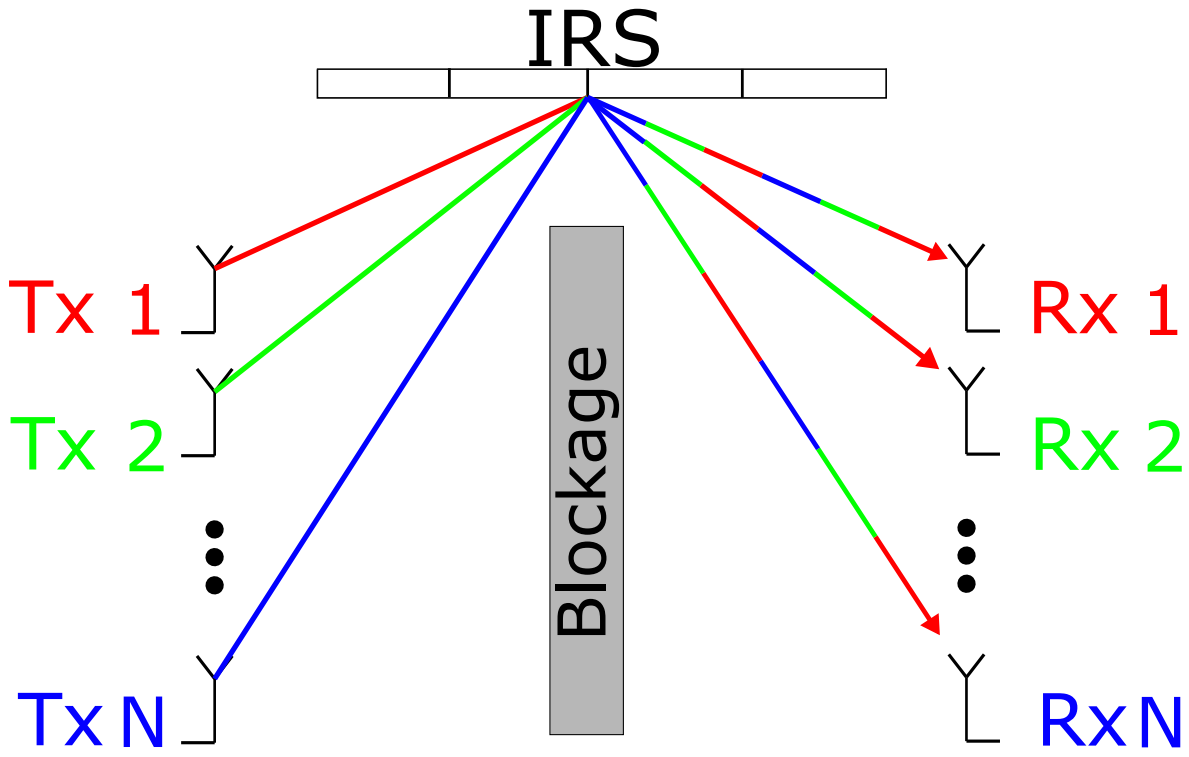


Figure 2.1: A single IRS model for MU transmission. Different colors mark the signal transmission path from different Txs.

signal for the i-th pair of transceiver. $\mathbf{A_{in}}$ is the channel matrix from the transmitters to the IRS, which can be represented by the matrix composed of steering vectors as

$$\mathbf{A_{in}} = [\mathbf{a}(\Omega_{in,1}), \mathbf{a}(\Omega_{in,2}), \ldots, \mathbf{a}(\Omega_{in,N})] \in \mathbb{C}^{M \times N}, \tag{2.2}$$

where $\mathbf{a}(\Omega_{in,i})$ is i-th user's steering vector of incident directions and $\Omega_{in,i}$ is the term containing the spatial information of incident directions from i-th transmitter Tx_i. According to the antenna array theorem, Ω is a function of azimuth and elevation angles for two-dimensional deployment (e.g., URA), or only contains one-dimensional information of azimuth angle such as ULA. Distance between transceivers and IRS can also be involved in Ω for representing a general location of transceivers.

Note that there is no noise in equation (2.1) since the IRS is a passive system, which is different from relay systems. Though this model is suitable for all frequency bands, however, it is particularly useful for mmWave communications, which suffer from severe coverage issues. On the IRS, each element will reflect the impinging signal from the Tx. However, by controlling the phases of the elements, the phase of the reflected signal at each element can be different. Optimal weights can be calculated to assure that the signals are coherently added at the direction of Rx. As one of the implementations, the elements' phases can be

controlled by a biasing circuit switch (e.g., as proposed in [5]). Thus, the weighted signal at the surface can be written as

$$\hat{\mathbf{y}}_s = \mathbf{W}\mathbf{y} = \mathbf{W}\mathbf{A}_{\mathbf{in}}\mathbf{s}\,, \tag{2.3}$$

where $\hat{\mathbf{y}}_s$ is the phased signal on IRS, the weights matrix $\mathbf{W} \in \mathbb{C}^{M\times M}$ is a diagonal matrix with each entity on the diagonal being the weight of each element.

At Rxs, the received signals reflected from the IRS are phased by another matrix of steering vectors $\mathbf{A}_{\mathbf{out}}$ which defined as

$$\mathbf{A}_{\mathbf{out}} = [\mathbf{a}(\Omega_{out,1}), \mathbf{a}(\Omega_{out,2}), \ldots, \mathbf{a}(\Omega_{out,N})] \in \mathbb{C}^{M\times N}\,. \tag{2.4}$$

$\mathbf{A}_{\mathbf{out}}$ has the same form as $\mathbf{A}_{in}$ in equation (2.2). $\Omega_{out,i}$ contains direction of i-th receiver Rx_i. $\mathbf{a}(\Omega_{out,i})$ represents the steering vector of exit angles of phased signals from the IRS to Rx_i. Thus, the received signal $\hat{\mathbf{y}}_r$ at all N receivers can be expressed as a vector form

$$\hat{\mathbf{y}}_r = \mathbf{A}_{out}^T\mathbf{W}\mathbf{A}_{in}\mathbf{s} + \mathbf{n}\,, \tag{2.5}$$

where the $\mathbf{n}$ is the noise vector at the receivers. The weights matrix, $\mathbf{W}$, needs to be solved and applied on the surface to make sure the reflected signals are optimally reflected towards the desired direction without cross-interference among the users. As the incident steering vectors and exit steering vectors with respect to the IRS are independent of each other and both controlled by the weights at the same time, to this end, we merge $\mathbf{a}(\Omega_{in,u})$ and $\mathbf{a}(\Omega_{out,v})$ as a compound steering vector

$$\mathbf{a}_C(\Omega_{out,v}, \Omega_{in,u}) = l_{IRS}\mathbf{a}(\Omega_{out,v}) \odot \mathbf{a}(\Omega_{in,u})\,, \text{for } u, v = 1, \ldots, N\,, \tag{2.6}$$

which means the equivalent steering vector or channel response from the transmitter at $\Omega_{in,u}$ to the receiver at $\Omega_{out,v}$. Here, l_{IRS} is the path loss factor for LoS path and its specific expression has been given in [122]. It should be noted that l_{IRS} includes the FSPL and effective area of IRS [123]. The effective area is the area impinging effectively by the EM wave and effective area can change with different incident angle in the near-field. For far-field transmission, without loss of generality, we normalize the effective area of IRS and consider the path loss factor of far field to be inversely proportional to $(d_{in}d_{out})^2$ [77], where d_{in} and d_{out} is the distance from Tx to IRS and IRS to Rx. Subsequently, equation (2.5) can be represented in an equivalent way as

$$\hat{y}_{r,i} = \mathbf{w}^H\mathbf{A}_{C,i}\mathbf{s} + n_i\,, i = 1, 2, \ldots, N\ . \tag{2.7}$$

$\hat{y}_{r,i}$ is the phased signal received by i-th receiver Rx_i, $\mathbf{w}$ is a column vector with its elements being the main diagonal elements of $\mathbf{W}$. It is worth mentioning that $\mathbf{w}$ is the *single* weight vector employed on the surface to achieve *all* transceiver pairs' desired signal response and mutual interference suppression simultaneously. n_i is the noise at Rx_i, and $\mathbf{A}_{C,i}$ is the combined steering matrix for i-th receiver which can be expressed as

$$\mathbf{A}_{C,i} = [\mathbf{a}_C(\Omega_{out,i}, \Omega_{in,1}), \ldots, \mathbf{a}_C(\Omega_{out,i}, \Omega_{in,N})] \in \mathbb{C}^{M \times N} . \tag{2.8}$$

With respect to the i-th receiver, according to equations (2.7) and (2.8), the received signal is a mixture of signals from all directions including both the desired signal and interference, where only $\mathbf{a}_C(\Omega_{out,i}, \Omega_{in,i})$ corresponds to desired signal direction and other terms in $\mathbf{A}_{C,i}$ correspond to interference for Rx_i. Base on equation (2.7), the SINR for the receiver of i-th pair can be represented as

$$SINR_i = \frac{|\mathbf{w}^H \mathbf{a}_C(\Omega_{out,i}, \Omega_{in,i})|^2}{\sigma^2 + \sum_{j=1, j\neq i}^{N} |\mathbf{w}^H \mathbf{a}_C(\Omega_{out,i}, \Omega_{in,j})|^2}, i = 1, 2, ..., N, \tag{2.9}$$

where σ^2 is the power of noise.

Note that the single IRS model can represent an arbitrary reflecting object with multiple impinging and reflected signals. In particular, the 'Txs' and 'Rxs' can represent multiple incident and reflected signals' direction in Fig. 2.1. The directions of 'Txs' are not necessarily from actual Txs as they can also be reflected signals from other objects. Similarly, reflected signals' directions can also be those of other reflecting objects. Meanwhile, if the weights of IRS are fixed, which means non-reconfigurable, the single IRS model can also represent other non-IRS reflecting objects, such as walls/floors/ceilings uncovered by the IRS. Thus, the single IRS model is fundamental for extending one-time reflection on one object toward MOR on multiple objects in the environment, which becomes a basic unit model in IRS networks.

2.1.1 Steering Vector of IRS

Though the shape of the surface and the arrangement of the elements on the surface can be arbitrary, however, we need to consider some regular IRSs in practical deployment and production. Again, it is mentioned that our model in this chapter and optimization in the following Chapters are not limited by these specific structures.

IRS with ULA Configuration

For ULA, both $\Omega_{in,u}$ and $\Omega_{out,v}$ contain only azimuth angle that $\Omega_{in,u} = \phi_{in,u}$, $\Omega_{out,v} = \phi_{out,v}$ and $\phi_{in,u}, \phi_{out,v} \in [0, \pi]$, thus we can rewrite (2.2) as

$$\mathbf{A_{in}} = \begin{bmatrix} 1 & \dots & 1 \\ e^{-jkd\cos\phi_{in,1}} & \dots & e^{-jkd\cos\phi_{in,N}} \\ \vdots & \ddots & \vdots \\ e^{-jkd\cos\phi_{in,1}(M-1)} & \dots & e^{-jkd\cos\phi_{in,N}(M-1)} \end{bmatrix} . \tag{2.10}$$

(2.6) can be rewritten as

$$\mathbf{a}_C(\phi_{out,v}, \phi_{in,u}) = [1, e^{-jkd(\cos\phi_{in,u}+\cos\phi_{out,v})}, \dots, e^{-jkd(\cos\phi_{in,u}+\cos\phi_{out,v})(M-1)}]^T , \tag{2.11}$$

where $k = 2\pi/\lambda$ is the angular wavenumber, with λ being the wavelength of the signal. d is the distance between the center of the adjacent electronic elements on the surface.

IRS with URA Configuration

Consider the elements deployment on IRS as shown in Fig 2.2. There are M elements in total. M_x is the number of elements along the X-axis and M_y is that of along the Y-axis, i.e., $M = M_x \cdot M_y$. d_x and d_y are elements distance along X-axis and Y-axis respectively. In this case, both $\Omega_{in,u}$ and $\Omega_{out,v}$ contain azimuth angle and elevation angle that $\Omega_{in,u} = (\phi_{in,u}, \theta_{in,u})$, $\Omega_{out,v} = (\phi_{out,v}, \theta_{out,v})$, $\phi_{in,u}, \phi_{out,v} \in [0, 2\pi]$ and $\theta_{in,u}, \theta_{out,v} \in [0, \pi/2]$. The corresponding form of compound steering vector for URA can be written as

$$\begin{aligned} \mathbf{a}_C(\Omega_{out,v}, \Omega_{in,u}) = [1, & e^{-jk(f_{cs}(\Omega_{out,v},\Omega_{in,u})d_x(0)+f_{ss}(\Omega_{out,v},\Omega_{in,u})d_y(1))} \dots, \\ & e^{-jk(f_{cs}(\Omega_{out,v},\Omega_{in,u})d_x m_x+f_{ss}(\Omega_{out,v},\Omega_{in,u})d_y m_y)}, \dots, \\ & e^{-jk(f_{cs}(\Omega_{out,v},\Omega_{in,u})d_x(M_x-1)+f_{ss}(\Omega_{out,v},\Omega_{in,u})d_y(M_y-1))}]^T , \end{aligned} \tag{2.12}$$

where $m_x \in (0, M_x - 1), m_y \in (0, M_y - 1)$. In addition

$$f_{cs}(\Omega_{out,v}, \Omega_{in,u}) = f_{cs}(\phi_{out,v}, \theta_{out,v}, \phi_{in,u}, \theta_{in,u}) = \cos\phi_{out,v}\sin\theta_{out,v} + \cos\phi_{in,u}\sin\theta_{in,u} , \tag{2.13}$$

$$f_{ss}(\Omega_{out,v}, \Omega_{in,u}) = f_{ss}(\phi_{out,v}, \theta_{out,v}, \phi_{in,u}, \theta_{in,u}) = \sin\phi_{out,v}\sin\theta_{out,v} + \sin\phi_{in,u}\sin\theta_{in,u} . \tag{2.14}$$

Compared with the form of ULA, it contains one more dimension of information of space, which is the elevation angle θ.

2.2 Modeling of the Basic IRS Element Characteristics

For the IRS, the most important designed factor would be the weights vector design, i.e., to design $\mathbf{w}$ in equation (2.7) for satisfying required performance. For the ideal weights vector,

$$\mathbf{w} = [w_1, \ldots, w_M]^H \tag{2.15}$$

where

$$w_m = A_m e^{j\Xi_m}, A_m \in \mathbb{R}^+, \Xi_m \in (0, 2\pi], \forall m = 1, \ldots, M. \tag{2.16}$$

The A_m is the amplitude response of each IRS element, and Ξ_m is the phase response. In this case, the weights vector is said to have the form of AUPC solution. Since the power supply on the IRS is limited, it is common to have a power constraint for $\mathbf{w}$, which is

$$||\mathbf{w}||_2 = M \ , \tag{2.17}$$

and the weights vector is satisfying the requirement of amplitude-relaxed phase-continuous (AXPC) solution in this case. However, the designed weights vector itself may have several other constraints from the hardware. In the following section, we present the basic circuit for realizing the prototype of IRS's element and its corresponding model concerning the impairment.

2.2.1 Impedance and Reflection coefficient of IRS element

It is conventionally acknowledged that the impedance of IRS elements can be derived from a class of lumped circuits [124–127]. For each IRS element, one way to control the reflecting coefficients is through leveraging varactors[1] to change the capacitance of the lumped circuit and thus the weight w_m. Specifically, the practical phase and amplitude response of the reflection coefficient can be studied with respect to different capacitance values and frequencies. In particular, the overall impedance of a lumped circuit for the m-th element of the IRS can be written as [124]

$$Z_m(C_m, f) = \frac{\mathrm{j}2\pi f L_1 \left(\mathrm{j}2\pi f L_2 + \frac{1}{\mathrm{j}2\pi f C_m} + R\right)}{\mathrm{j}2\pi f L_1 + \mathrm{j}2\pi f L_2 + \frac{1}{\mathrm{j}2\pi f C_m} + R}, \tag{2.18}$$

where L_1 and L_2 are inductances in different layers, and R is the effective resistance. The effective capacitance C_m can be controlled using a diode or a varactor. With a specific

[1] Varactors are used as voltage-controlled capacitors.

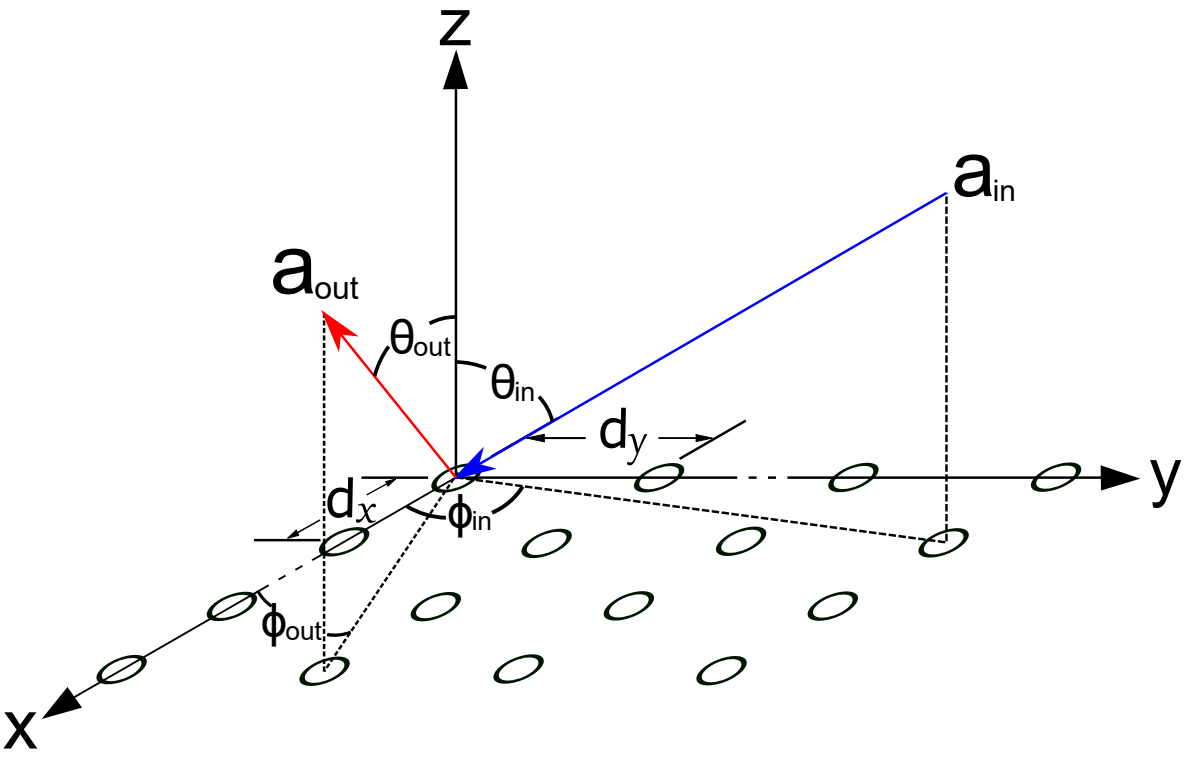


Figure 2.2: 2D IRS under URA deployment in 3D geometry axis

impedance of each element, the reflection coefficient of this IRS element can be computed as [124]

$$w_m(C_m, f) = \frac{Z_m(C_m, f) - Z_0}{Z_m(C_m, f) + Z_0}, Z_0 = 377\Omega. \tag{2.19}$$

To show the characteristics of the lumped circuit which makes up an IRS element, simulations are provided here for illustration. The amplitude and phase function of the reflection coefficient are shown in Fig. 2.3 and Fig. 2.4 with respect to different frequency, f, and different values of capacitance, C_m. For the simulation set up, we have $L_1 = 2.5$ nH and $L_2 = 0.7$ nH and $R = 1\Omega$.

By referring to the plots of Fig. 2.3 and Fig. 2.4 below, for frequency around 30 GHz, the amplitude can be nearly equal to 1 while the phase can not cover over 360 degrees or 2π in radians. When the frequency is around 3 GHz, the phase is nearly ideal for covering 360 degrees of phase, while the amplitude can experience an apparent notch. Due to the fact that the amplitude response of weights is commonly fixed, a class of tricky constraints called the constant modulus is introduced. Most formulated problems involving the constant modulus constraint are generally non-convex and NP-hard. For constant modulus constraint on weights, we can write

$$w_m = A_m e^{j\Xi_m}, A_m = A_c, \Xi_m \in (0, 2\pi], \forall m = 1, \ldots, M, \tag{2.20}$$

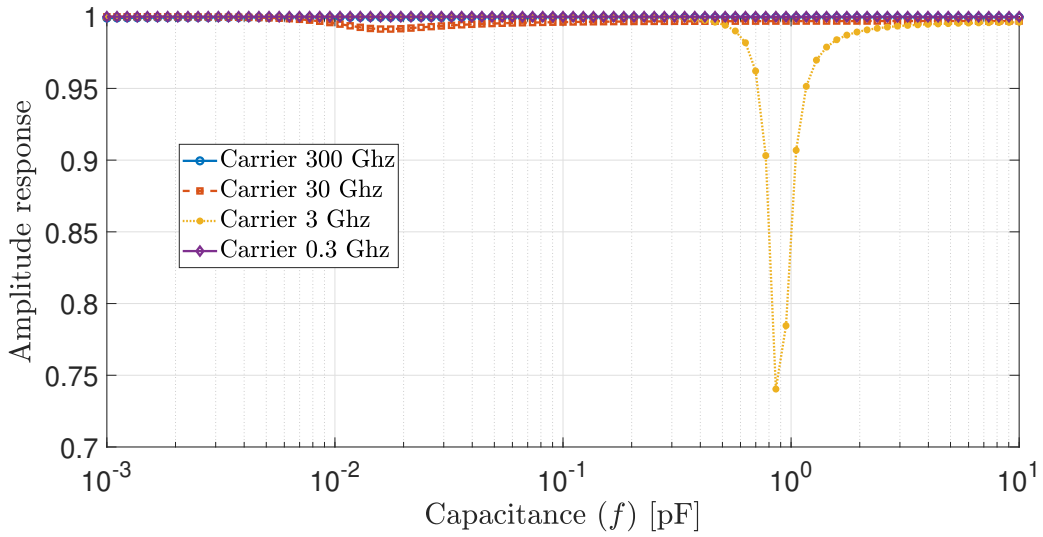


Figure 2.3: Amplitude vs Capacitance

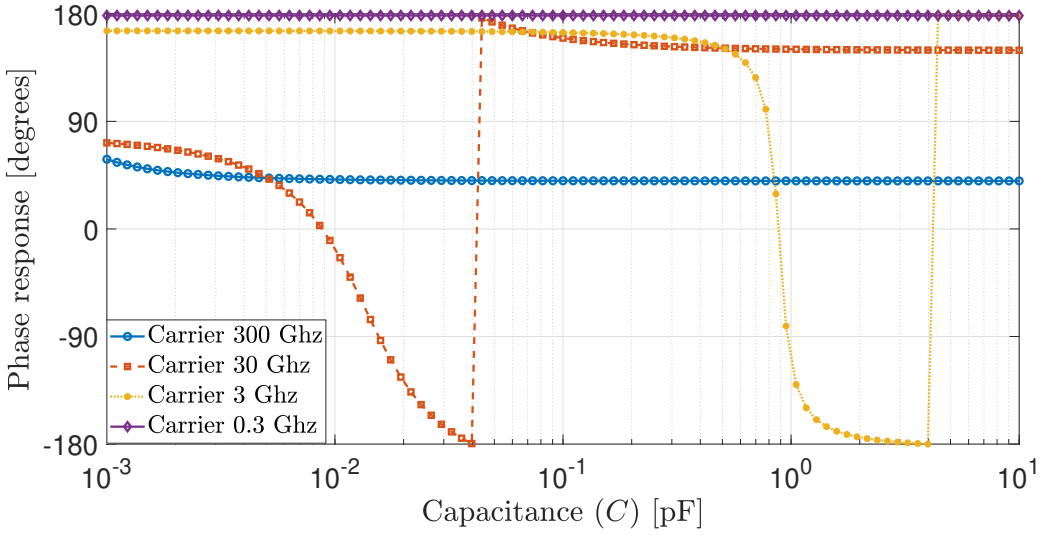


Figure 2.4: Phase vs Capacitance

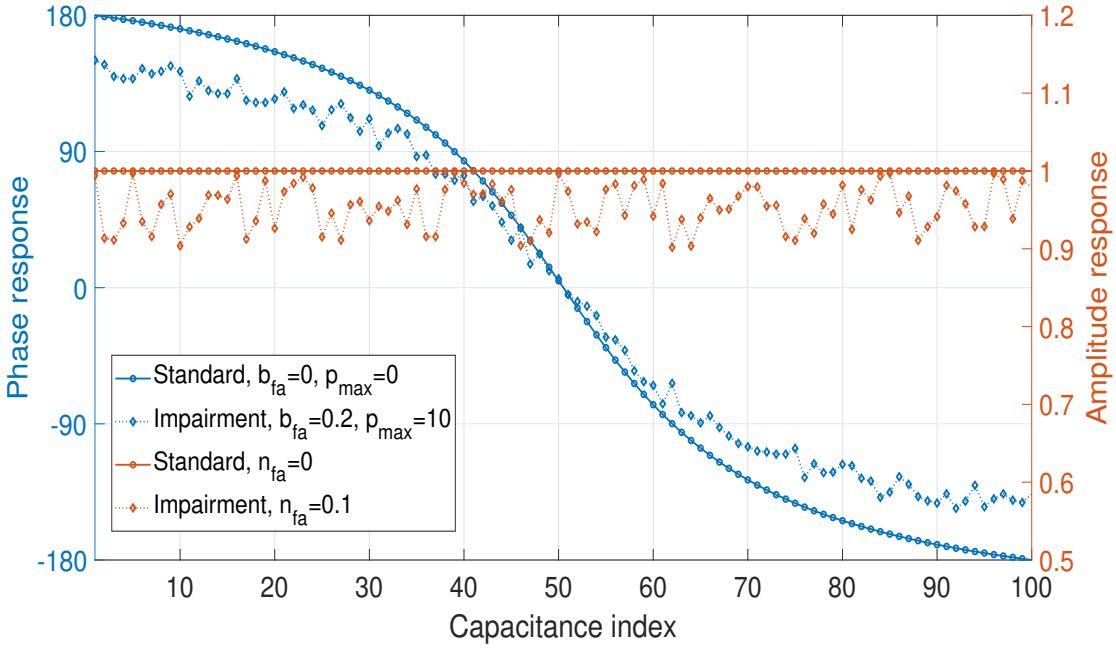


Figure 2.5: Characteristic of IRS.

where A_c is a real constant. Since each element is passive that does not have any power amplifying function, A_c should be less than 1, as shown in Fig. 2.3. In order to be consistent with the power supply constraint, we can have a unified value that $A_c = 1$. In this case, the weights vector satisfies the requirement of ACPC solution. However, the constant modulus constraint is still hard to practically satisfy by hardware design. This is due to the current controlling to IRS being digital. With digital controls for the capacitance switching, the capacitance value change is not a continuous process but a discrete process. Thus, only a few values on the phase response and amplitude response can be obtained.

It is also conventionally acknowledged that a higher bit of control over the number of states for the amplitude/phase responses of elements is more difficult [128]. This is due to higher bits control requiring more active devices (e.g., PIN diodes) to achieve higher digital states. Hence, to design an efficient solution for practically constrained IRS, the constraints with lower costs on the hardware design of the basic element should be considered. Specifically, phases can only have a limited number of states to be adjusted. As a result, the constraints bring the quantization to the continuous state space of the weights vector, which motivates a shift from theoretical ACPC to a more practical amplitude-constrained phase-quantized (ACPQ) solution. Nevertheless, the theoretical ACPC solution holds its importance as it can be regarded as a class of performance upper bound.

2.2.2 Impairment of IRS element

The impairment of IRS element is from the bias between the desired amplitude and phase response of IRS and the actual amplitude and phase response of the IRS affected by the phase/amplitude noise. To present the model with impairment, we consider some error parameters on the phase and amplitude of IRS weights, respectively. First, without loss of generality, we define $C \in \{1, 2, ..., N_i\}$ as the set of index numbers to replace the realistic value of capacitance for theoretical analysis. The integer N_i is the maximum number of circuit states to which the controller can tune. Further, we define b_{fa} as the blind phase factor because we notice the phase may not reach a full range of 360 degrees. We call the phase range at which the practical weights (reflection coefficients) are unable to reach the blind phase range. Thus, the ratio of the blind phase range that existed in such an IRS element can be represented by b_{fa}, e.g., if $b_{\text{fa}} = 0$, then the phase range of the practical weights is able to span across the full 360 degrees. Next, to consider the impairment of the phase response, we introduce the maximal phase bias, $p_{\max}$, which represents the maximum absolute value of error added to the ideal phase response. As the amplitude response may not be constant and it also has impairment issues, we define the notch factor, n_{fa}, which indicates the maximum percentage that amplitude negatively deviates away from the designed amplitude response $f_a(C)$, representing the extent of impairment on the amplitude. In particular, the fitting function of m-th element for the phase response, $\Xi_m(C)$, and the amplitude response, $A_m(C)$, can be written as

$$x(C) = \frac{2\pi}{N_i - 1} C + \frac{1 + N_i}{1 - N_i} \pi \,, \tag{2.21}$$

$$\Xi_m(C) = -\frac{\arctan(x(C))\pi}{\arctan(\pi)} (1 - b_{\text{fa}}) + \epsilon_1 \,, \tag{2.22}$$

and

$$A_m(C) = f_a(C) - \epsilon_2 \,, \tag{2.23}$$

where $\epsilon_1 \sim U(-p_{\max}, p_{\max})$ and $\epsilon_2 \sim U(0, n_{\text{fa}})$ are considered as errors on the phase and amplitude respectively due to the electronic impairment and modelled as continuous uniform distribution. The value of function $f_a(C)$ in equation (2.23) is always less equal to one due to each element working passively and $f_a(C)$ can be designed correspondingly by optimizing the architecture of the hardware. As shown in Fig. 2.5, where $N_i = 100$ and $f_a(C) = 1$, if impairment factors are all zero value, i.e., $b_{\text{fa}} = 0$, $n_{\text{fa}} = 0$ and $p_{\max} = 0$, then the amplitude and phase function is ideal. An example of practical amplitude and phase functions is shown in Fig. 2.5 with a non-zero value of impairment factors. We can observe the amplitude response can be unstable due to the introduced error and the blind phase factor can show the imperfect phase range.

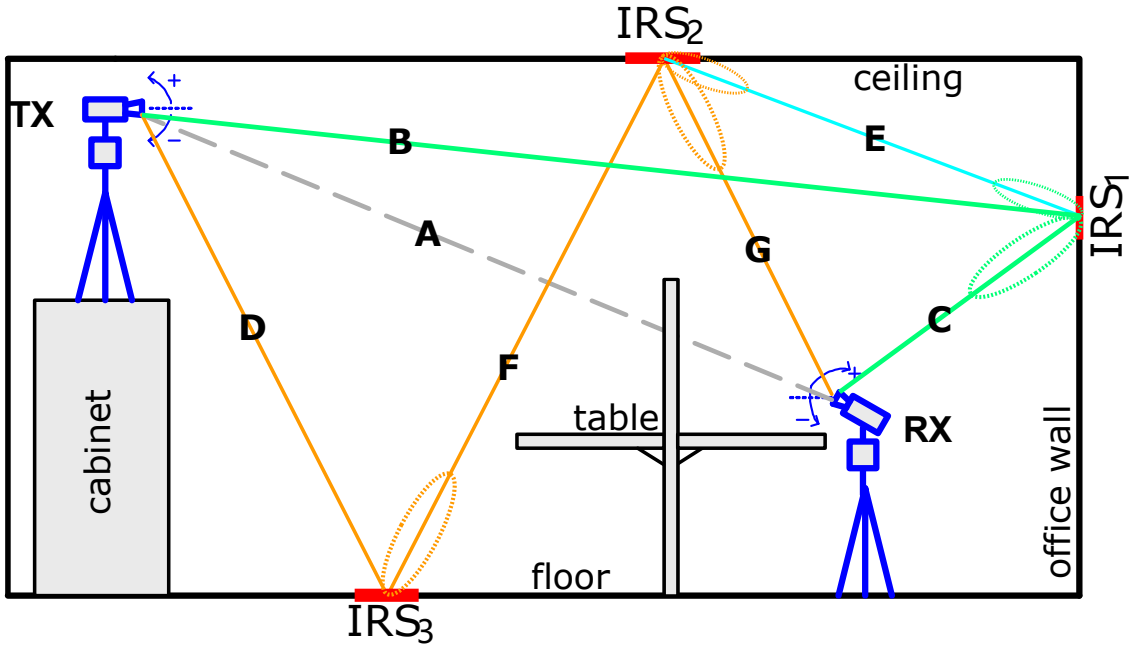


Figure 2.6: An example of indoor transmission assisted by IRS network with the same furniture setting.

2.3 Model of Multi-IRS network

Note that a single IRS deployed in the environment have a limited amount of ability to improve the wireless channel. This is due to the size of IRS being limited and a single IRS can not be fabricated in an infinitely large size. In addition, the reflected wave can also face further blockage and hence still bear a significant power loss. Therefore, multiple-IRS networks are also critical to enhance the control of the wireless channel.

In IRS networks, the dual reflection plays a critical effect in constructing the programmable channel. Specifically, the dual reflection happens between two reflectors in placement with a dihedral angle such that the beam lobes of two reflectors can point towards each other. Here, we exemplify the dual reflection between a pair of IRS in an indoor transmission scheme, where the same setup of furniture is referred from work in [104], as shown in Fig. 2.6. Path **A** is the blocked path between a transceiver pair thus leveraging IRS is necessary. Paths **B** and **D** are LoS paths between Tx to an IRS, paths **E** and **F** are that of from one IRS to another IRS, and paths **C** and **G** are that of from one IRS to Rx. Then Rx can receive FOR signal from a C-LoS path **B**-**C** and second-order-reflection (SOR) signal from a C-LoS path **D**-**F**-**G**. In addition, there are paths caused by dual reflection and without loss of generality, we introduce the dual reflection between IRS_1 and IRS_2. As an LoS path **E** exist between IRS_1 and IRS_2, the signal components impinging on IRS_1 and IRS_2 can be reflected towards each other due

to side lobes. In particular, IRS_2 can receive the FOR signal from a C-LoS path **B**-**E**. Meanwhile, the SOR signal via a C-LoS path **D**-**F**-**E** can be received by IRS_1 as well. In this case, the dual reflections between IRS_1 and IRS_2 are introduced. Immediately, the Rx further receive the SOR signal passing through a C-LoS path **B**-**E**-**G** and third-order-refection signal along another C-LoS path **D**-**F**-**E**-**C**. Due to the dual reflection will still occur, higher-order reflection signals are successively produced by repetitive signal reflections between IRS_1 and IRS_2 (for example, C-LoS paths **B**-**E**-**E**-**C** and **B**-**E**-**E**-**E**-**G**). As a result, some signal components keep continuous reflecting between the dual IRS pair, while other parts can either reach Rx or dissipate in trivial directions. Note that, signal components from higher-order reflections should not be neglected as long as they are not overwhelmed by Rx's noise power, or potential destruction of signal amplitude, fatal phase distortion and inter-symbol interference can significantly undermine the overall system performance. Thus, the dual reflection should be well considered in a complete signal model of IRS networks.

In the following parts, two fundamental models in IRS networks are presented. We consider LoS channels obey the quasi-optical transmission nature of EM carrier following works [66,77,104,105,129,130]. Meanwhile, the NLoS channel between transceivers is considered, as no LoS paths between transceivers could be a common and pressing issue [131], as shown in Fig. 2.6. In addition, we also assume each transceiver and IRS is located in a far-field as did in the literature [56, 83, 84, 102, 106, 107, 109–112, 114]. The above relevant state-of-art works involving these condition are often valid for most wireless transmission scenarios since users are assumpted to be far away from the BS. Also, it has been considered a classic assumption in books [4, 22].

2.3.1 Channel Model Between Two IRSs

Here, we derive the LoS channel model between one IRS to another as it is fundamental to make up a part of the complete model of the IRS network.

Lemma 1. *The channel matrix between any two IRSs is **rank-one** and can be written as*

$$\mathbf{E} = \mathbf{a}(\phi_{in})\mathbf{a}(\phi_{out})^T, \tag{2.24}$$

where ϕ_{out} is the AOD of the signal leaving from the first IRS towards the next IRS, and ϕ_{in} is the AOA of the signal arriving at the next IRS.

Proof. We consider IRS_A and IRS_B have M_A, M_B elements with element spacing d_A and d_B respectively. We denote A_i and B_j are the i-th element and j-th element on IRS_A and IRS_B, $i \in [1, M_A], j \in [1, M_B]$. The relative distance from the i-th element on IRS_A to the

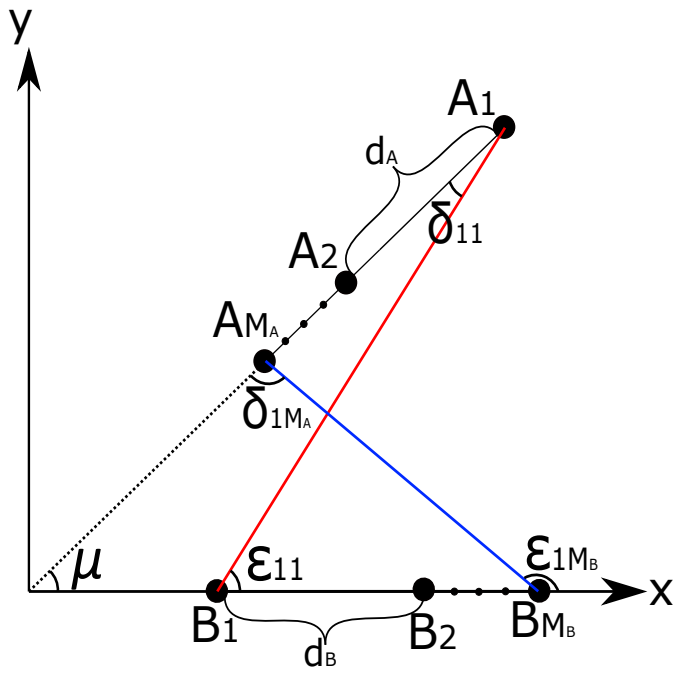


Figure 2.7: The illustration of the channel model between two IRSs.

first element A_1 is $d_{A,i}$ and for that of IRS$_B$ is $d_{B,j}$ between the j-th element on IRS$_B$ and B_1. Since now we have two pieces of IRS, to distinguish, we denote the azimuth AOD of IRS$_A$ between elements A_i and B_j as δ_{ij}, and denote the azimuth AOA of IRS$_B$ as ε_{ij}. Then, we denote the distance between element A_i on IRS$_A$ and element B_j on IRS$_B$ as D_{ij}, and μ is the angle between IRS$_A$ and IRS$_B$, as shown in Fig. 2.7. We assume D_{11}, δ_{11} and ε_{11} is known, and there is $\mu = \varepsilon_{11} - \delta_{11}$. From the trigonometric relationship, we have

$$\varepsilon_{ij} = \tan^{-1}\left(\frac{D_{11}\sin\varepsilon_{11} - d_{A,i}\sin\mu}{D_{11}\cos\varepsilon_{11} - d_{A,i}\cos\mu - d_{B,j}}\right). \tag{2.25}$$

The distance D_{ij} between elements A_i and B_j can be calculated as

$$D_{ij} = \frac{D_{11}\sin\varepsilon_{11} - d_{A,i}\sin\mu}{\sin\varepsilon_{ij}}. \tag{2.26}$$

Since the far-field condition holds where distance is much greater than the aperture of IRS such that $D_{11} >> Md$, we have $\varepsilon_{ij} \approx \varepsilon_{11}$ and $\delta_{ij} \approx \delta_{11}$ correspondingly. Thus, by substituting ε_{ij} with ε_{11} in equation (2.25), we have

$$d_{A,i}\sin\delta_{11} = -d_{B,i}\sin\varepsilon_{11}. \tag{2.27}$$

Then, we substitute equation (2.27) into equation (2.26), we have

$$D_{ij} = D_{11} - d_{A,i}\cos\delta_{11} - d_{B,j}\cos\varepsilon_{11}. \tag{2.28}$$

Since the LoS channel between IRS_A and IRS_B can be represented by

$$\mathbf{E}_{AB} = e^{-jk\mathbf{D}}, \tag{2.29}$$

where $\mathbf{D} \in \mathbb{C}^{M_A \times M_B}$ is the distance matrix derived from D_{ij}, $\mathbf{E}_{AB}$ can be rewritten as

$$\mathbf{E}_{AB} = e^{-jkD_{11}}(\mathbf{a}(\varepsilon_{11})\mathbf{a}(\delta_{11})^T)^*. \tag{2.30}$$

Note that the path delay D_{11} is a constant between any two fixed IRSs. Since the path delay is known, it can be regarded as constant. In addition, by taking the inverse element order of IRS_A and IRS_B, which is equal to taking conjugate to the steering vectors of AOA and AOD, equation (2.30) can be rewritten as

$$\mathbf{E}_{AB} = \mathbf{a}(\varepsilon_{11})\mathbf{a}(\delta_{11})^T. \tag{2.31}$$

As $\delta_{11} = \phi_{out}$, $\varepsilon_{11} = \phi_{in}$, we can observe that the LoS channel between arbitrary two IRSs can be considered as the out product of two steering vectors, which is a rank-one matrix given in equation (2.24). □

From the view of Lemma 1, each IRS can regard another IRS as a point source located in the far-field, but both of them can shape a pencil beam towards each other. Additionally, it means that only a single data stream can be supported by an LoS channel between two pieces of IRS. In real applications, ranks can be greater than one due to the diffraction and refraction effects of EM waves. However, as the carrier frequency increases for more spectrum resources, the diffraction and refraction effects become weak and vulnerable. Hence multiple streams' transmission between two IRSs becomes impractical, where the traditional rayleigh fading model is inconsistent in this case [130]. Therefore, in this work, we consider the rank-one channel between two IRSs.

2.3.2 MUFOR IRS Network

Denote the channel of FOR IRS network as $\mathbf{H}_{I,1}$. Based on equation (2.5), the received signal with K pieces IRS can be expressed as

$$\mathbf{y} = \mathbf{H}_{I,1}\mathbf{s} + \mathbf{n}\,, \tag{2.32}$$

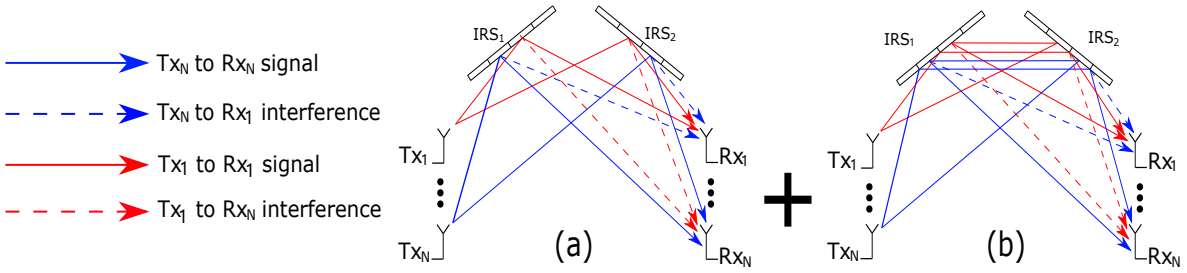


Figure 2.8: MUMOR transmission within the IRS network given $\Gamma = 2$, $K = 2$. (a) MU signals passing along the FOR IRS network channel, $\mathbf{H}_{I,1}$. (b) MU signals passing along the SOR IRS network channel, $\mathbf{H}_{I,2}$.

where

$$\mathbf{H}_{I,1} = \sum_{k=1}^{K} \mathbf{A}_{out,k}^{T} \mathbf{W}_k \mathbf{A}_{in,k} \tag{2.33}$$

and $\mathbf{A}_{in,k}$ and $\mathbf{A}_{out,k}$ are two steering vector matrices of AOA and AOD with respect to k-th IRS. Note that, the k-th C-LoS path component is made up via multiplexing only one weights matrix $\mathbf{W}_k$ one time with the other two steering vector matrices, $\mathbf{A}_{in,k}$ and $\mathbf{A}_{out,k}$. Thus, the IRS network channel $\mathbf{H}_{I,1}$ embodies K different paths and all of these paths only experience one-time reflection.

2.3.3 MUMOR IRS Network

To model the MOR effect analytically, we define the maximum order of reflections within the IRS network as Γ. Essentially, Γ plays the effective cut-off parameter on the MOR effect. However, it is possible to consider $\Gamma \to \infty$[2], we need to cut off using Γ as a finite value because we have path loss in practical scenarios. In this case, the reflection order of signal components less than or equal to Γ is considered. In contrast, the signal components with orders higher than Γ are assumed to be overwhelmed by the noise power and can be neglected. To involve an arbitrary number of reflections in the IRS network, we denote the IRS network channel in γ-th order as $\mathbf{H}_{I,\gamma}$, where $\gamma \in [1, \Gamma]$. Then, we extend the IRS network channel in equation (2.33) from FOR to MOR via superposition as

$$\mathbf{H}_I = \sum_{\gamma=1}^{\Gamma} \mathbf{H}_{I,\gamma} \,, \tag{2.34}$$

where totally Γ orders of IRS network channels are added up. The γ-th order MU channel component $\mathbf{H}_{I,\gamma}$ includes all C-LoS path components that experience γ orders in the network, which also means each C-LoS path component in $\mathbf{H}_{I,\gamma}$ is exactly weighted for γ times.

[2]It is similar to the LoS path of visible light reflected within two mirrors or more mirrors for infinite times.

Note that FOR only involves one-time reflection on a single IRS while FOR is a more common case that signal is reflected multiple times from multiple IRS in different places. To differentiate MOR from FOR in the IRS network and illustrate the dual reflection, we exemplify by considering two pieces of IRS, where $K = 2$ and $\Gamma = 2$, as shown in Fig. 2.8. In this case, each C-LoS path passes maximal 2 pieces IRS. The C-LoS paths with a reflection order of more than three are ignored. Thus, the received signal of all `Rxs` should consist of MU signals passing along the FOR IRS network channel $\mathbf{H}_{I,1}$ and the SOR IRS network channel $\mathbf{H}_{I,2}$ which has been respectively shown in Fig. 2.8(a) and (b). Considering the FOR IRS network channel, we can have

$$\mathbf{H}_{I,1} = \mathbf{A}_{out,1}^{T}\mathbf{W}_1\mathbf{A}_{in,1} + \mathbf{A}_{out,2}^{T}\mathbf{W}_2\mathbf{A}_{in,2} \ . \tag{2.35}$$

For the SOR IRS network channel, we have

$$\mathbf{H}_{I,2} = \mathbf{A}_{out,2}^{T}\mathbf{W}_2\mathbf{E}_{12}\mathbf{W}_1\mathbf{A}_{in,1} + \mathbf{A}_{out,1}^{T}\mathbf{W}_1\mathbf{E}_{21}\mathbf{W}_2\mathbf{A}_{in,2} \ . \tag{2.36}$$

Note that $\mathbf{E}_{12}$ and $\mathbf{E}_{21}$ are the LoS channels between between IRS_1 and IRS_2, as we derived in equation (2.31), where $\mathbf{E}_{12} = \mathbf{E}_{21}^{T}$. We can observe the number of C-LoS path components in $\mathbf{H}_{I,\gamma}$ with different K and different γ variates and still follow the permutation's rule. For example, for $K = 2$, we have the IRS candidate set $\kappa = \{1, 2\}$ which means there are only IRS_1 and IRS_2 in the environment. Since $\gamma = 1$, the number of FOR paths is equal to 2 where one FOR path passes through IRS_1, and another one passes through IRS_2. Using the permutation rule, we can denote the number of FOR paths as ${}^{2}P_1 = 2$ (${}^{X}P_Y$ means Y-permutations of a set with X elements, where $X, Y \in \mathbb{N}^{+}$). Similarly, for $K = 2$ and $\Gamma = 2$, the number of SOR paths equal to 2 since ${}^{2}P_2 = 2$. Specific order sequences of these two SOR paths can be enumerated here, i.e., we have $[1 \ \ 2]$, meaning a SOR path first passes through IRS_1 and then IRS_2, and $[2 \ \ 1]$, meaning another SOR path passes through IRS_2 and then IRS_1. Consequently, the total number of C-LoS paths of γ orders in the K-piece IRS network is equal to ${}^{K}P_{\gamma}$. Note that although ${}^{K}P_{\gamma}$ only includes the number of C-LoS paths which pass each IRS only once in IRS networks, we will discuss and consider C-LoS paths which repetitively visit the same IRS later. To expand $\mathbf{H}_{I,\gamma}$ in general expression, we define an index matrix $\mathbf{X}_{\gamma}$ to denote the order sequences for all C-LoS paths in γ orders. In particular, all rows of index matrix $\mathbf{X}_{\gamma}$ are used to hold specific order sequences of all C-LoS paths of γ orders. For example, given $K = 2, \gamma = 2$, by leveraging the index matrix $\mathbf{X}_2$, equation (2.36) can now be written as

$$\mathbf{H}_{I,2} = \sum_{u=1}^{{}^{2}P_2} \mathbf{A}_{out,X_{\gamma,u2}}\mathbf{W}_{X_{\gamma,u2}}\mathbf{E}_{X_{\gamma,u1}X_{\gamma,u2}}\mathbf{W}_{X_{\gamma,u1}}\mathbf{A}_{in,X_{\gamma,u1}} \ , \tag{2.37}$$

where the index matrix $\mathbf{X}_2$ for $\mathbf{H}_{I,2}$ is

$$\mathbf{X}_2 = \begin{bmatrix} X_{2,11} & X_{2,12} \\ X_{2,21} & X_{2,22} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}. \tag{2.38}$$

We can observe $[X_{2,11}\ \ X_{2,12}]$=[1 2] and $[X_{2,21}\ \ X_{2,22}]$ = [2 1] are exactly two sequences we enumerate. For arbitrary value of γ and K, we define the index matrix as $\mathbf{X}_\gamma \in \mathbb{N}^{+^{KP_\gamma \times \gamma}}$. The term $X_{\gamma,uv}$ at the u-th row and the v-th column of $\mathbf{X}_\gamma$ is a positive integer representing an index of a specific IRS in the network. Each row of $\mathbf{X}_\gamma$ holds a specific and non-repetitive sequence with γ columns. The index matrix $\mathbf{X}_\gamma$ has ${}^KP_\gamma$ rows in total, which means all order sequences under a partial permutation ${}^KP_\gamma$ are included. To generate the index matrix, one can enumerate the permutation sequences of γ terms from the IRS candidate set $\kappa = \{1, 2, ..., K\}$ respectively as rows of $\mathbf{X}_\gamma$ [132].

Theorem 1. *The general expression of γ-order IRS network channel $\mathbf{H}_{I,\gamma}$ can be written as*

$$\mathbf{H}_{I,\gamma} = \sum_{u=1}^{{}^KP_\gamma} \mathbf{A}_{out,X_{\gamma,u\gamma}} [\prod_{v=1}^{\gamma-1} \mathbf{W}_{X_{\gamma,u(v+1)}} \mathbf{E}_{X_{\gamma,uv}X_{\gamma,u(v+1)}}] \mathbf{W}_{X_{\gamma,u1}} \mathbf{A}_{in,X_{\gamma,u1}} \tag{2.39}$$

where $\mathbf{X}_\gamma \in \mathbb{N}^{+^{KP_\gamma \times \gamma}}$ is the index matrix of γ orders.

Note that the dual reflection is common, and we should consider other C-LoS paths whose order sequences are with repetitive indices. These C-LoS paths should at least visit a single IRS of all pieces twice. For the order sequences of these C-LoS paths with repetition, the adjacent two terms in rows of the index matrix should be different as we consider that there is no LoS path between one IRS and itself. Thus the EM wave would not impinge on the same IRS twice immediately, i.e., $X_{\gamma,uv} \neq X_{\gamma,u(v+1)}, u \in [1, {}^KP_\gamma], v \in [1, \gamma - 1]$. To complete the IRS network model, we include order sequences, whose two interleaved indices can be equal to one another, into the index matrix $\mathbf{X}_\gamma$ with extra rows. Therefore, the row dimension of $\mathbf{X}_\gamma$ extends from ${}^KP_\gamma$ to $K(K-1)^{(\gamma-1)}$. By far, if we replace $\mathbf{H}_{I,1}$ with equation (2.34) and equation (2.39) in equation (2.32), then the complete model of IRS network is established.

Chapter 3

Passive Beamforming of Single IRS

In this chapter, the beamforming approaches for a single IRS are investigated given different constraints. As MU beamforming is the focus, the proposed technique aims to provide significant interference mitigation. With respect to the different users who can require different QoS, the signal power reflected from IRS aims to reach the corresponding QoS requirement for each transceiver. In addition, given different hardware constraints, for both ideal continuous response controlling or low-cost multi-bit controlling, the corresponding solutions are obtained respectively. Specifically, we study the AUPC solution, AXPC solution, ACPC solution and ACPQ solution for the designed weights vector. It is worth mentioning that one of the critical phenomena of single IRS for MU transmission, the redundant beam has been respectively analyzed for ULA IRS and URA IRS.

3.1 Proposed Beamforming Weights

3.1.1 Optimal Weights of IRS For Single-user

For single-user transmission aided by a single IRS, it is well known that max ratio combining (MRC) beamformer can realize the maximal power gain, as long as we know the direction of transceivers in the far-field[1]. In particular, given $N = 1$, we have

$$\hat{y}_{r,1} = \mathbf{w}^H \mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1})\mathbf{s} + n_1 . \tag{3.1}$$

In order to receive the maximal power, the phases of all elements on the IRS should be aligned with the phases of the cascaded channel to obtain the maximal power response. The

[1]Specific locations or perfect CSI are required if near-field and multi-path scenarios are considered.

closed-form solution of MRC in this case can be written as

$$\mathbf{w}_{MRC} = \frac{\mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1})}{||\mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1})||_2}. \tag{3.2}$$

Since for each element on the IRS, the amplitude is assumed to be unified, the power for elements IRS would be proportional to the total elements' number M. The power constraint, $||\mathbf{w}_{MRC}||_2 = M$, of IRS, is satisfied and it is a finite value.

3.1.2 Optimal Weights of IRS For Multi-user

As managing the interference is critical for MU transmission, we propose to obtain the optimal solution to minimize the interference for IRS MU system through the MU-LCMV beamformer. Note that linearly constrained minimum variance (LCMV) is used since it is a class of common beamformers that can obtain the optimal result as a benchmark [17]. As a result, not only the known interference is nullified but that from unknown directions is also minimized. In particular, the optimization problem of minimizing interference can be formulated as

$$(\text{P1}): \quad \min_{\mathbf{w}} J(\mathbf{w}) = E\{\Sigma_{i=1}^{N} |\hat{y}_{r,i}|^2\} \triangleq \mathbf{w}^H \mathbf{R} \mathbf{w} + \mathbf{I} N \sigma_n^2 \tag{3.3}$$

$$s.t. \quad \mathbf{C^H w} = \mathbf{f}\ . \tag{3.4}$$

Equation (3.3) means to minimize overall power at all receivers and (3.4) refers to the constraint equations for all transceiver pairs. σ_n^2 is the power of noise and we assume it is the same at all receivers. $\mathbf{R}$ in (3.3) is the covariance matrix and can be written as

$$\mathbf{R} = \mathbf{A}_C \mathbf{A}_C^H\ , \tag{3.5}$$

where $\mathbf{A}_C$ that containing all critical steering vectors are written, via concatenation of each steering matrix in equation (2.8) for N receivers, as

$$\begin{aligned} \mathbf{A}_C = \Bigg[& \underbrace{\mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1}), \ldots \mathbf{a}_C(\Omega_{out,1}, \Omega_{in,N})}_{\text{N}}, \underbrace{\ldots}_{\text{N(N-2)}} \\ & \underbrace{\mathbf{a}_C(\Omega_{out,N}, \Omega_{in,1}), \ldots, \mathbf{a}_C(\Omega_{out,N}, \Omega_{in,N})}_{\text{N}} \Bigg] \in \mathbb{C}^{M \times N^2}\ , \end{aligned} \tag{3.6}$$

$\mathbf{C}$ in (3.4) is the constraint matrix and can be expressed as

$$\mathbf{C} = [\mathbf{C}_1, \mathbf{C}_2] \,, \tag{3.7}$$

$$\mathbf{C}_1 = [\mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1}), \ldots, \mathbf{a}_C(\Omega_{out,N}, \Omega_{in,N})] \,, \tag{3.8}$$

$$\mathbf{C}_2 = [\ldots, \mathbf{a}_C(\Omega_{out,q}, \Omega_{in,p}), \ldots] \,, p \neq q, p = 1, \ldots, N, \, q = 1, \ldots, N \;\; , \tag{3.9}$$

where $\mathbf{C}_1$ means the channel matrix which contains the steering vectors of desired signals to remain, and $\mathbf{C}_2$ is the interference channel matrix composed of steering vectors of the unwanted directions where the signal will be suppressed. In addition, $\mathbf{f}$ is the beam pattern response constraint vector which contains two parts defined as

$$\mathbf{f} = [\mathbf{f}_1, \mathbf{f}_2]^T \,, \tag{3.10}$$

$$\mathbf{f}_1 = \left[\underbrace{\delta_{11}, \ldots, \delta_{NN}}_{N} \right]^T \,, \tag{3.11}$$

$$\mathbf{f}_2 = \left[\underbrace{\ldots, \delta_{pq}, \ldots}_{N(N-1)} \right]^T \,. \tag{3.12}$$

$\mathbf{f}_1$ corresponds to the response values for desired signals, while $\mathbf{f}_2$ contains the maximum tolerant response value for interference. For simplicity but without loss of generality, the value can be set to unit power response, i.e., $\delta_{ii} = \delta_{lower} = 1$, for all desired signals (i.e., the minimum achieved power response at the desired directions) as a lower bound. To suppress the interference, we can set $\delta_{pq} = \delta_{upper} << \delta_{lower}$, which means δ_{upper} is an upper bound of the interference response.

The optimal solution of (P1) can be derived based on the mindset of the LCMV beamformer with the Lagrange method [133], which can be written as

$$\mathbf{w}_{MU-LCMV} = \mathbf{R}^{-\mathbf{1}}\mathbf{C}(\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C})^{-1}\mathbf{f} \,. \tag{3.13}$$

Note that the optimization problem (P1) minimizes the sum power of interference at all receivers. Specifically, (P1) is to minimize the sum power of signals received by all receivers. However, the desired signal power component is kept constant through the constraint equation (3.4). Thus, it is equivalently that the interference is nullified by IRS. Although the weight is in the form of a single vector, the IRS can support multi-stream transmission since it is fundamentally different from the traditional MIMO system. We can prove mathemat-

ically that the IRS model with a single weight vector can be transferred to an equivalent traditional MIMO model with multiple weight vectors, subject to two conditions as follows. Firstly, the sufficient number of elements on the IRS should be equal to or greater than the number of pairs, i.e., $M \geq N$. This condition guarantees there is a sufficient degree of freedom for achieving zero-forcing at the interference locations. It is worth noting that to make the system better performed, we typically require $M \geq N^2$. While this requirement can be easily achieved in practical deployment since each IRS can be with hundreds or thousands of elements. Secondly, the directions of receivers and transmitters should be different to each other, i.e., $\mathbf{A}_{C,i} \neq \mathbf{A}_{C,j}$, for $i \neq j$, and $i, j \in [1, N]$. This condition can assure the coefficient matrix of the weights vector for zero-forcing is full rank. Under these two conditions, the manipulation by the common weight to different steering matrices would produce equivalent effects that different weight vectors manipulate the same steering matrix. Thus, multi-stream data transmission can be supported by a single IRS. The closed-form optimal solution in equation (3.13) achieves the required responses at the desired signal and interference directions. This AUPC solution provides an analytical upper bound of performance and can guide the real system design. However, such a solution is ideal and impractical for IRS since the calculated weights in $\mathbf{w}$ can have an arbitrarily large (or small) amplitude, which is impossible for a passive IRS where the amplitude is normally pre-designed and non-tunable. For this practical concern, we add the constant modulus constraint in the following problem. Though amplitude variation is an issue in designing the weights, relative methods can be applied to equalize the variation, which will be studied in future work.

3.1.3 Suboptimal Weights With Constraints

As the passive characteristics of weights, the amplitude of each element typically equals to or less than 1, depending on the design materials and circuits. Besides, in a practical IRS, all elements are periodically deployed on the surface with the same reflecting coefficients (i.e., they have the same amplitude). For simplicity and without losing the generality, the amplitude of weights is fixed to 1 in this subsection.

Here, we propose to solve the problem with the new constant modulus constraint by letting $\mathbf{w} = [w_1, \ldots, w_M]^H$ where $w_m = e^{j\Xi_m}, \Xi_m \in (0, 2\pi], \forall m = 1, \ldots, M$. We propose using

SQP [134] to get ACPC solution by solving following problem

$$\text{(P2)}: \min_{\mathbf{w}} J(\mathbf{w}) = E\{\Sigma_{i=1}^{N}|\hat{y}_{r,i}|^2\} \triangleq \mathbf{w}^H\mathbf{R}\mathbf{w} \tag{3.14}$$

$$s.t.\ \ \mathbf{C}_1^H\mathbf{w} \geq \mathbf{f}_1\ , \tag{3.15}$$

$$\mathbf{C}_2^H\mathbf{w} \leq \mathbf{f}_2\ , \tag{3.16}$$

$$w_m = e^{j\Xi_m}, \Xi_m \in (0, 2\pi], m = 1, \ldots, M\ . \tag{3.17}$$

To be noted, (P2) with nonlinear constraints is not convex due to the constraint of constant modulus being introduced. Therefore, (P2) is NP-hard [135]. We propose to use the SQP method, which is a class of popular methods for solving nonlinear and non-convex optimization, to provide a locally optimal result for our work. The biggest advantage is that it can search the infeasible points, which can be useful when constraints become strict such that other algorithms can not obtain a solution normally. As equation (3.17) should be satisfied to avoid the unbounded minimization of the original problem via SQP, we substitute $\mathbf{w}$ with $\boldsymbol{\Xi}$, where $\boldsymbol{\Xi} = [\Xi_1, \ldots, \Xi_M]^T$, in (P2) to optimize the phase directly using the mapping relationship of equation (3.17). Then, SQP solves (P2) by solving a quadratic programming subproblem first in each iteration to compute a search direction for (P2). Each subproblem is obtained by linearizing the constraints and approximating the Lagrangian function as

$$L(\boldsymbol{\Xi}, \boldsymbol{\lambda}) = J(\boldsymbol{\Xi}) - \sum_{i=1}^{N^2} \lambda_i G_i(\boldsymbol{\Xi})\ , \tag{3.18}$$

where $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_l]^T \in \mathbb{R}^{N^2}$ is the vector of the Lagrange multiplier. G_i represent i-th constraint in (3.15) and (3.16). The subproblem to be solved to find a search direction for the original problem (P2) can be expressed as

$$\text{(P3)}: \min_{\mathbf{d}(k)} \frac{1}{2}\mathbf{d}(k)^T\mathbf{B}(k)\mathbf{d}(k) + \mathbf{g}^T\mathbf{d}(k) \tag{3.19}$$

$$s.t.\ \ G_i + \nabla G_i^T\mathbf{d}(k) \geq 0 \quad i = 1, 2, \cdots N^2\ , \tag{3.20}$$

where $\mathbf{g} = \nabla J(\boldsymbol{\Xi}), G_i = G_i(\boldsymbol{\Xi})$ and $\mathbf{B}(k)$ is usually a positive definite approximation to the Hessian matrix of the Lagrangian function with respect to $\boldsymbol{\Xi}$. $\mathbf{d}(k)$ is the solution to search a $(k+1)$-th iterated solution of (P2) at k-th iteration, which can be used as

$$\boldsymbol{\Xi}(k+1) = \boldsymbol{\Xi}(k) + a_k\mathbf{d}(k)\ , \tag{3.21}$$

where $a_k \in (0, 1]$ is the step length parameter. Next, to guarantee the global convergence of

(P2), a merit function is constructed as

$$\eta(\boldsymbol{\Xi}) = J(\boldsymbol{\Xi}) + p||\mathbf{G}^{+}(\boldsymbol{\Xi})||_1 \, , p > 0 \tag{3.22}$$

where $J(\boldsymbol{\Xi})$ is the objective function of (P2) after substitution of $\mathbf{w}$ with $\boldsymbol{\Xi}$ and p is a positive constant to be chosen. The i-the term of $\mathbf{G}^{+}(\boldsymbol{\Xi})$ is

$$G_i^{+}(\boldsymbol{\Xi}) = \begin{cases} 0 & \text{if} \quad G_i(\boldsymbol{\Xi}) \leq 0 \\ G_i(\boldsymbol{\Xi}) & \text{if} \quad G_i(\boldsymbol{\Xi}) > 0 \end{cases}, i = 1, 2, ..., N^2 \, . \tag{3.23}$$

Thus, a_k can be computed by using the line search method in the direction provided by the solution of (P3) to minimize the value of merit function $\eta(\boldsymbol{\Xi})$ at k-th iteration. Once the minimum of the merit function is reached, we obtain $\boldsymbol{\Xi}^*$ and $\mathbf{w}^*$ to (P2).

SQP algorithm is defined as: Step 1, initialize $\boldsymbol{\Xi}$, and k. Step 2, solve the subproblem (P3) to determine $\mathbf{d}(k)$ and let $\boldsymbol{\lambda}_{k+1}$ be the vector of the Lagrange multiplier of the linear constraints obtained from the subproblem (P3). Step 3, compute the length a_k such that $\eta(\boldsymbol{\Xi}(k) + a_k\mathbf{d}(k)) < \eta(\boldsymbol{\Xi}(k))$ then update the solution by equation (3.21). Step 4, calculate $\mathbf{B}(k+1)$ from $\mathbf{B}(k)$ using a quasi-Newton formula. Step 5, stop until $||\eta(\boldsymbol{\Xi}(k+1)) - \eta(\boldsymbol{\Xi}(k))||_1 < \epsilon$ is achieved, where $\epsilon > 0$. Otherwise, go back to Step 2. The pseudo-code for implementing SQP can be seen as follows.

Algorithm 1 The SQP algorithm for solving (P2)

Input: $\mathbf{R}$, $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{C}_1$ and $\mathbf{C}_2$
Output: $\boldsymbol{\Xi}^*$

Initialize $\boldsymbol{\Xi}_0$, ϵ, k, a_k
Let $\boldsymbol{\lambda}_k$ be Lagrange multiplier of equation (3.18)
while (1) **do**
    Solve (P3) to get $\mathbf{d}(k)$
    for $\eta(\boldsymbol{\Xi}(k) + a_k\mathbf{d}(k)) \geq \eta(\boldsymbol{\Xi}(k))$ **do**
        Compute a_k by line searching
    end for
    $\boldsymbol{\Xi}(k+1) = \boldsymbol{\Xi}(k) + a_k\mathbf{d}(k)$
    Let $\boldsymbol{\lambda}_{k+1}$ be the vector of the Lagrange multiplier of (P3)
    Compute $\mathbf{B}(k+1)$ from $\mathbf{B}(k)$ using a quasi-Newton formula
    if $||\eta(\boldsymbol{\Xi}(k+1)) - \eta(\boldsymbol{\Xi}(k))||_1 < \epsilon$ **then**
        $Break$
    end if
end while
$\boldsymbol{\Xi}^* \leftarrow \boldsymbol{\Xi}(k+1)$
$\mathbf{w}^* \leftarrow e^{j\boldsymbol{\Xi}(k+1)}$

3.1.4 Phase quantization

In addition to the amplitude, the phase of the weights can be constrained to lower the cost, size, and complexity of the IRS. In literature, the phase of weights is usually realized by switching on and off of a simple inductor-capacitor (LC) circuit, which corresponds to 0 and π phase responses respectively on the elements. This is so-called 1-bit coding. Advanced IRS can achieve higher phase resolution through 2-bit coding with more choices of phase responses of 0, $\pi/2$, π and $3\pi/2$, or even higher bits coding which has a greater controlling degree of freedom but also brings higher system cost [5, 136, 137].

To this end, we consider constraints of quantized phase shifter to the IRS. The ACPQ solution is obtained by solving (P2) first and then quantizing the calculated optimal constant modulus solution by B bits. Such an idea has been proposed and widely used in mmWave communications [70]. The quantized step will be $1/2^B$ and available weights for each element will be selected from

$$w_m = e^{j2\pi n_B/2^B}, n_B = 1, \ldots, 2^B. \tag{3.24}$$

More specifically, after obtaining weights from equation (3.21), weights will be classified to the nearest discretized phases based on B, so the distance between the ACPC solution and the ACPQ solution is minimized on the complex plane.

3.2 Limitation of Single IRS

3.2.1 On Transceiver's Directions

Since the MU transmission is achieved by a single IRS, fundamental limits on the transceiver locations can be defined by the following corollary.

Corollary 1. *Directions of different transmitters should not be the same, and so as to directions of receivers, which can be represented as*

$$\Omega_{in,1} \neq \Omega_{in,2} \neq \cdots \neq \Omega_{in,N} \ \textit{and} \ \Omega_{out,1} \neq \Omega_{out,2} \neq \cdots \neq \Omega_{out,N} \, . \tag{3.25}$$

This condition can be verified easily through the constraints of the original optimization problem. Taking $N = 2$ as an example, Tx1 sends signal to Rx1 and Tx2 sends signal to Rx2. According to the formulation (3.14), the solution should satisfy constraints (3.16) and (3.17) for specific responses and then we have relationships by rewritten the constraints as

$$\begin{cases} \mathbf{w}^H \mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1}) \geq \delta_{lower} \\ \mathbf{w}^H \mathbf{a}_C(\Omega_{out,2}, \Omega_{in,1}) \leq \delta_{upper} \\ \mathbf{w}^H \mathbf{a}_C(\Omega_{out,2}, \Omega_{in,2}) \geq \delta_{lower} \\ \mathbf{w}^H \mathbf{a}_C(\Omega_{out,1}, \Omega_{in,2}) \leq \delta_{upper} \end{cases}, \tag{3.26}$$

where $\delta_{lower} >> \delta_{upper}$. If $\Omega_{in,1} = \Omega_{in,2}$, by substituting $\Omega_{in,2}$ with $\Omega_{in,1}$, equation (3.26) becomes

$$\begin{cases} \mathbf{w}^H \mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1}) \geq \delta_{lower} \\ \mathbf{w}^H \mathbf{a}_C(\Omega_{out,1}, \Omega_{in,1}) \leq \delta_{upper} \\ \mathbf{w}^H \mathbf{a}_C(\Omega_{out,2}, \Omega_{in,1}) \leq \delta_{upper} \\ \mathbf{w}^H \mathbf{a}_C(\Omega_{out,2}, \Omega_{in,1}) \geq \delta_{lower} \end{cases}. \tag{3.27}$$

Considering that the minimum response of desired signal δ_{lower} is normally much larger than the maximum response of interference δ_{upper}, i.e., $\delta_{lower} >> \delta_{upper}$, the first and the second equations (also the third and fourth) of (3.27) are self contradicted. The condition of transmitters in the same direction makes the original optimization problem unsolvable. The verification is the same if the receivers are in the same direction.

According to **Corollary 1**, it is not difficult to conclude that if incident angles or exit angles are too close to each other, low channel SINR will be achieved. However, these limitations can be overcome by using more than one IRS to make transceiver pairs spatially distinguishable. In a nutshell, for a single meta-surface system to realize multi-user transmission, no overlapping in terms of both transmitters' and receivers' directions is compulsory.

3.2.2 On Half-power Beamwidth

For an IRS with fixed element distance, the HPBW is also limited. Also, since the transceiver can be located in arbitrary directions with respect to the IRS, the HPBW is also subjected to the transceiver's direction.

In work [138], based on the antenna array theory, Han derived the HPBW of a single IRS with weights assumption on the IRS. In particular, the weights vector $\mathbf{w}$ are assumed to be $w_m = 1$, for $m = 1, 2, ..., M$. The author further contributes a class of generalization to the derived result. In essence, the derived HPBW in this case is still consistent with the HPBW under MRC beamformer. Specifically, for an Rx at a given location, no matter where the Tx is, we can always find an optimal solution by using MRC, so that the beampattern and HPBW under incident signals from different directions are the same. This can be proven by the characteristics of MRC beamformer on the specific arrays.

For $w_m = 1$, for $m = 1, 2, ..., M$, we write array factor (AF) as

$$AF = \mathbf{a}^T(\Omega_{out}) \cdot \mathbf{a}(\Omega_{in}) \ . \tag{3.28}$$

According to (3.28), the AF of IRS could be defined as

$$AF = e^{-j\xi_1} + e^{-j\xi_2} + \cdots + e^{-j\xi_m} + \cdots + e^{-j\xi_M} \ , \tag{3.29}$$

where ξ_m are the merged phases of incident and reflection plane wave at the element locations $m = 1, 2, \cdots, M$. For an IRS in ULA, we have

$$\xi_m = kd(m-1)(\cos\phi_{in} + \cos\phi_{out}) = (m-1)\hat{\xi} \tag{3.30}$$

where $\hat{\xi}$ is a constant, once we fixed the directions of transceivers, carrier frequency and element spacing. Since a large element spacing d (e.g., $d = \frac{2\lambda}{3}$) may lead to grating lobes causing interference and energy loss, a smaller d is preferred and the value of $\frac{\hat{\xi}}{2}$ will be small, thus the equivalence of $\sin(\frac{\hat{\xi}}{2}) \approx \frac{\hat{\xi}}{2}$ holds. According to (3.29) and (3.30), the AF could be normalized and then simplified to

$$AF = \frac{1}{M}\sum_{m=1}^{M} e^{-j(m-1)\hat{\xi}} = \frac{1}{M}[\frac{\sin(\frac{M}{2}\hat{\xi})}{\sin(\frac{1}{2}\hat{\xi})}] \approx \frac{\sin(\frac{M}{2}\hat{\xi})}{\frac{M}{2}\hat{\xi}}. \tag{3.31}$$

Since for weights employing the $w_m = 1$, for $m = 1, 2, ..., M$, there are no further phase terms introduced from the weights themselves, the weights from other algorithms can be combined with phases for other analysis. The MRC weights can ideally combine the phase from arbitrary paths delay for co-phasing add up. Therefore, consider applying weights with MRC on IRS, we can write the AF as

$$AF = \mathbf{w}_{MRC}^H \mathbf{a}_C(\Omega_{out}, \Omega_{in}) = \sum_{m=1}^{M} e^{-jkd(\cos\phi_{in} + \cos\phi_{out} + \zeta_m)(m-1)} \ , \tag{3.32}$$

where ζ_m is the weights term from MRC. In this case, $\zeta_m = -\cos\phi_{in} - \cos\phi_{in} + \delta_c$. The δ_c is also a designed constant that MRC can be introduced after combining the phase in steering vectors. By letting $kd\delta_c = \hat{\xi}$, the generalization is complete. The further derivation for HPBW is the same as we did in [138]. To conclude, for ULA, we have

$$HPBW = \begin{cases} |\pi - \phi_{in} - \cos^{-1}(\frac{-2.782}{kMd} - \cos\phi_{in})| + \phi_{in} \ , 0 < \phi_{in} \leq b_l \\ |\cos^{-1}(\frac{2.782}{kMd} - \cos\phi_{in}) - \cos^{-1}(\frac{-2.782}{kMd} - \cos\phi_{in})|, b_l < \phi_{in} < b_r \\ |\pi - \phi_{in} - \cos^{-1}(\frac{2.782}{kMd} - \cos\phi_{in})| + \pi - \phi_{in} \ , b_r \leq \phi_{in} < \pi \end{cases} . \tag{3.33}$$

where b_l and b_r are the boundary values of whether the HPBW is affected by the specific value transceiver's direction, and

$$b_l = \cos^{-1}(1 - \frac{2.782}{kMd}), \quad b_r = \cos^{-1}(\frac{2.782}{kMd} - 1) \ , \ . \tag{3.34}$$

3.3 Redundant Beams of IRS

Based on the model given by Section 2.1 and the optimization problem formulation in Section 3.1.2, it is not difficult to imagine IRS works as a special "mirror" in which each Tx can "see" its targeted Rx through such a mirror. In addition, since there is only one set of weights for all transceivers, each transmitter will be able to "see" other $N-1$ redundant beams. Given the fixed optimal weights, moving the observation location from one transmitter to another (e.g., the 2^{nd} transmitter) will "see" its beam but not the newly generated beam (and we cannot since only has one set of weights). Equivalently, it shifts the 1^{st} pair's one of the redundant beams as its main beam for another receiver. In other words, the beam pattern of each pair should have in total N peaks with one main beam at the desired direction and other $N-1$ redundant beams in other directions if redundant beams are judged to exist. In addition, each direction of redundant beams is only determined by the direction of such a pair and another pair (i.e., one of redundant beams for q-th Tx and Rx is determined by the directions of p-th Tx and Rx). The essence of the redundant beam is actually the channel equivalence of two pairs. As the physical characteristics of IRS are determined by itself, redundant beams should be a common phenomenon for multi-user communications. Naturally, we can find such property in IRS brings about the tradeoff design between the multiplexing and beamforming gains.

We mathematically derive the directions of redundant beams under different element distances in the next. However, to make the discussion simple and easy to follow, we first specify our discussion with an ULA IRS and then extend it to URA IRS.

To investigate the relationship between the direction of the redundant beam and the direction of transceiver pairs, we divide our discussion into three cases according to the range of element distance. It should be mentioned that in this Chapter, we discuss the cases that $d \in (0, \lambda/2]$. The scenario of $d \in (\lambda/2, \infty)$ is not considered here for both ULA and URA, since it may cause signal spatial aliasing. We denote the direction of a specific redundant beam as $\Omega_{RB,pq}$ which means that the q-th receiver can receive a signal from q-th transmitter at $\Omega_{RB,pq}$, and such a direction is essentially caused by the p-th pair's constraint. Appendix A.1 and Appendix A.2 list the proofs of the following corollaries for both ULA and URA.

3.3.1 Beam Pattern of IRS for Multi-user

Beam pattern is an ideal metric to measure the performance of the interference cancellation and observe redundant beams. Based on the combined steering vector of $\Omega_{in,u}$ and $\Omega_{out,v}$, the beam pattern of the IRS-based beamformer can be obtained and written as

$$B(\Omega_{out,v}, \Omega_{in,u}) = |\mathbf{w}^H \mathbf{a}_C(\Omega_{out,v}, \Omega_{in,u})| \,. \tag{3.35}$$

Note that the weight vector $\mathbf{w}$ can be either ideal or constrained weights calculated from different optimizations formulated in Section 3.1. The beam pattern $B(\Omega_{out,v}, \Omega_{in,u})$ can be regarded as the performance metric to describe the sensitivity of the meta-surface with respect to the signal transmitted and received from different AOA.

3.3.2 ULA's Redundant Beams

To carefully investigate the relationship between the position of the redundant beam and the user location, we will divide our discussion into three sectors according to the range of d. The derivation of the following corollary can be viewed in Appendix A.1 and A.2. Firstly, when d equals to $\lambda/2$. If we define $\phi_{pq} = \cos\theta_{in,p} + \cos\theta_{out,p} - \cos\theta_{in,q}$ with $p \neq q$ and $\{p, q\} \in [1, N]$, it is easy to find that $\phi_{pq} \in (-3, 3)$ since $\theta_{dir} \in (0, \pi)$. Then, we can have the following corollary:

Corollary 2. *When $d = \lambda/2$, the redundant beam of the q-th pair is produced by the pole constraint of the p-th pair. The redundant beam caused by the p-th pole is denoted as θ_{pq}, which can be calculated as follows:*

$$\theta_{pq} = \begin{cases} \arccos(\phi_{pq} + 2), \phi_{pq} \in (-3, -1) \\ \arccos(\phi_{pq} - 2), \phi_{pq} \in (1, 3) \\ \arccos(\phi_{pq}), \phi_{pq} \in (-1, 1) \\ 0 \quad and \quad \pi, \phi_{pq} = -1 \, or \, 1 \end{cases} \tag{3.36}$$

Next, the following corollaries are proposed with the elements distance $d \in (0, \lambda/4)$ and $d \in [\lambda/4, \lambda/2)$, respectively:

Corollary 3. *When $d \in (0, \lambda/4)$, the reflected angle of the redundant beam θ_{pq} can be calculated by:*

$$\theta_{pq} = \begin{cases} \arccos(\phi_{pq}), \phi_{pq} \in [-1, 1] \\ \emptyset, \ else \end{cases} \tag{3.37}$$

Corollary 4. *When $d \in [\lambda/4, \lambda/2)$, the reflect angle of the redundant beam θ_{pq} can be achieved as:*

$$\theta_{pq} = \begin{cases} \arccos(\phi_{pq}), \phi_{pq} \in [-1,1] \\ \arccos(\phi_{pq} + \lambda/d), \phi_{pq} \in (-3,-1), \lambda/d \leq 1 - \phi_{pq} \\ \emptyset, \phi_{pq} \in (-3,-1), \ \ else \\ \arccos(\phi_{pq} - \lambda/d), \phi_{pq} \in (1,3), \lambda/d \leq 1 + \phi_{pq} \\ \emptyset, \phi_{pq} \in (1,3), \ \ else \end{cases} \tag{3.38}$$

It should be mentioned that the scenario when $d \in (\lambda/2, \infty)$ is not considered here, since it may produce more than one redundant beams which would cause signal spatial aliasing. As such, this case is rarely applied in the multi-antenna design theory. Based on the three corollaries, we can find that under some specific conditions, the redundant beams do not exist, and therefore, the power of the signal can be focused in the desired direction.

Remark 1. *With N pairs of transceivers, up to $N-1$ redundant beams will be generated and thus may cause significant power loss. However, this property can be used for broadcasting/multicast. In addition, shortening the element distance can be helpful for decreasing the power separation of the signal since the number of redundant beams is also reduced.*

3.3.3 URA's Redundant Beams

Table 3.1: $\Omega_{RB,pq}$ of URA with $d = \lambda/2$, Corollary 5

$\phi_{pq} \in [0, 2\pi]$ \ β_{pq} / α_{12}	[-3,-1]	(-1,1)	[1,3]
[-3,-1]	$\arctan(\frac{\beta_{pq}+2}{\alpha_{pq}+2}) + n_3\pi$	$\arctan(\frac{\beta_{pq}}{\alpha_{pq}+2}) + n_3\pi$	$\arctan(\frac{\beta_{pq}-2}{\alpha_{pq}+2}) + n_3\pi$
(-1,1)	$\arctan(\frac{\beta_{pq}+2}{\alpha_{pq}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}}{\alpha_{pq}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}-2}{\alpha_{pq}}) + n_3\pi$
[1,3]	$\arctan(\frac{\beta_{pq}+2}{\alpha_{pq}-2}) + n_3\pi$	$\arctan(\frac{\beta_{pq}}{\alpha_{pq}-2}) + n_3\pi$	$\arctan(\frac{\beta_{pq}-2}{\alpha_{pq}-2}) + n_3\pi$

β_{pq}	[-3,-1]	(-1,1)	[1,3]
$\theta_{pq} \in [0, \frac{\pi}{2}]$	$\arcsin(\frac{\beta_{pq}+2}{sin\phi_{pq}})$	$\arcsin(\frac{\beta_{pq}}{sin\phi_{pq}})$	$\arcsin(\frac{\beta_{pq}-2}{sin\phi_{pq}})$

Consider a URA IRS with $d = \lambda/2$ and we define the other two intermediate variables as

$$\alpha_{pq} = \cos\phi_{in,p}\sin\theta_{in,p} + \cos\phi_{out,p}\sin\theta_{out,p} - \cos\phi_{in,q}\sin\theta_{in,q}, \tag{3.39}$$

$$\beta_{pq} = \sin\phi_{in,p}\sin\theta_{in,p} + \sin\phi_{out,p}\sin\theta_{out,p} - \sin\phi_{in,q}\sin\theta_{in,q}\ . \tag{3.40}$$

Table 3.2: $\Omega_{RB,pq}$ of URA with $d \in [\lambda/4, \lambda/2)$, Corollary 7

$\phi_{pq} \in [0,2\pi]$ \ β_{pq} / α_{pq}	[-3,-1], $\beta_{pq} + \frac{2\pi}{kd} \leq 1$	(-1,1)	[1,3], $\beta_{pq} - \frac{2\pi}{kd} \geq -1$
[-3,-1], $\alpha_{pq} + \frac{2\pi}{kd} \leq 1$	$\arctan(\frac{\beta_{pq}+\frac{\lambda}{d}}{\alpha_{pq}+\frac{\lambda}{d}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}}{\alpha_{pq}+\frac{\lambda}{d}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}-\frac{\lambda}{d}}{\alpha_{pq}+\frac{\lambda}{d}}) + n_3\pi$
(-1,1)	$\arctan(\frac{\beta_{pq}+\frac{\lambda}{d}}{\alpha_{pq}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}}{\alpha_{pq}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}-\frac{\lambda}{d}}{\alpha_{pq}}) + n_3\pi$
[1,3], $\alpha_{pq} - \frac{2\pi}{kd} \geq -1$	$\arctan(\frac{\beta_{pq}+\frac{\lambda}{d}}{\alpha_{pq}-\frac{\lambda}{d}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}}{\alpha_{pq}-\frac{\lambda}{d}}) + n_3\pi$	$\arctan(\frac{\beta_{pq}-\frac{\lambda}{d}}{\alpha_{pq}-\frac{\lambda}{d}}) + n_3\pi$

β_{pq}	[-3,-1], $\beta_{pq} + \frac{2\pi}{kd} \leq 1$	(-1,1)	[1,3], $\beta_{pq} - \frac{2\pi}{kd} \geq -1$
$\theta_{pq} \in [0, \frac{\pi}{2}]$	$\arcsin(\frac{\beta_{pq}+\frac{\lambda}{d}}{\sin\phi_{pq}})$	$\arcsin(\frac{\beta_{pq}}{\sin\phi_{pq}})$	$\arcsin(\frac{\beta_{pq}-\frac{\lambda}{d}}{\sin\phi_{pq}})$

Similar to ULA case, it is easy to find that both α_{pq} and $\beta_{pq} \in [-3, 3]$. Then, we can have the following corollaries.

Corollary 5. *Consider a URA IRS with $d = \lambda/2$ for N pairs' transceivers. The redundant beam direction $\Omega_{RB,pq} = (\phi_{pq}, \theta_{pq})$ in this case, and it can be calculated from* Table 3.1. Here, $n_3 \in Z$.

For $d \in (0, \lambda/4)$ and $d \in [\lambda/4, \lambda/2)$, we have the following corollaries respectively.

Corollary 6. *For URA IRS with $d \in (0, \lambda/4)$, the reflection angle of the redundant beam ϕ_{pq} and θ_{pq} can be expressed as*

$$\phi_{pq} = \begin{cases} \arctan(\frac{\beta_{pq}}{\alpha_{pq}}), \alpha_{pq}, \beta_{pq} \in [-1, 1] \\ \emptyset, \ else \end{cases}, \tag{3.41}$$

$$\theta_{pq} = \begin{cases} \arcsin(\frac{\beta_{pq}}{\sin\phi_{pq}}), \alpha_{pq}, \beta_{pq} \in [-1, 1] \\ \emptyset, \ else \end{cases}. \tag{3.42}$$

Corollary 7. *When $d \in [\lambda/4, \lambda/2)$, the reflect angle of the redundant beam ϕ_{pq} and θ_{pq} can be expressed in* Table II.

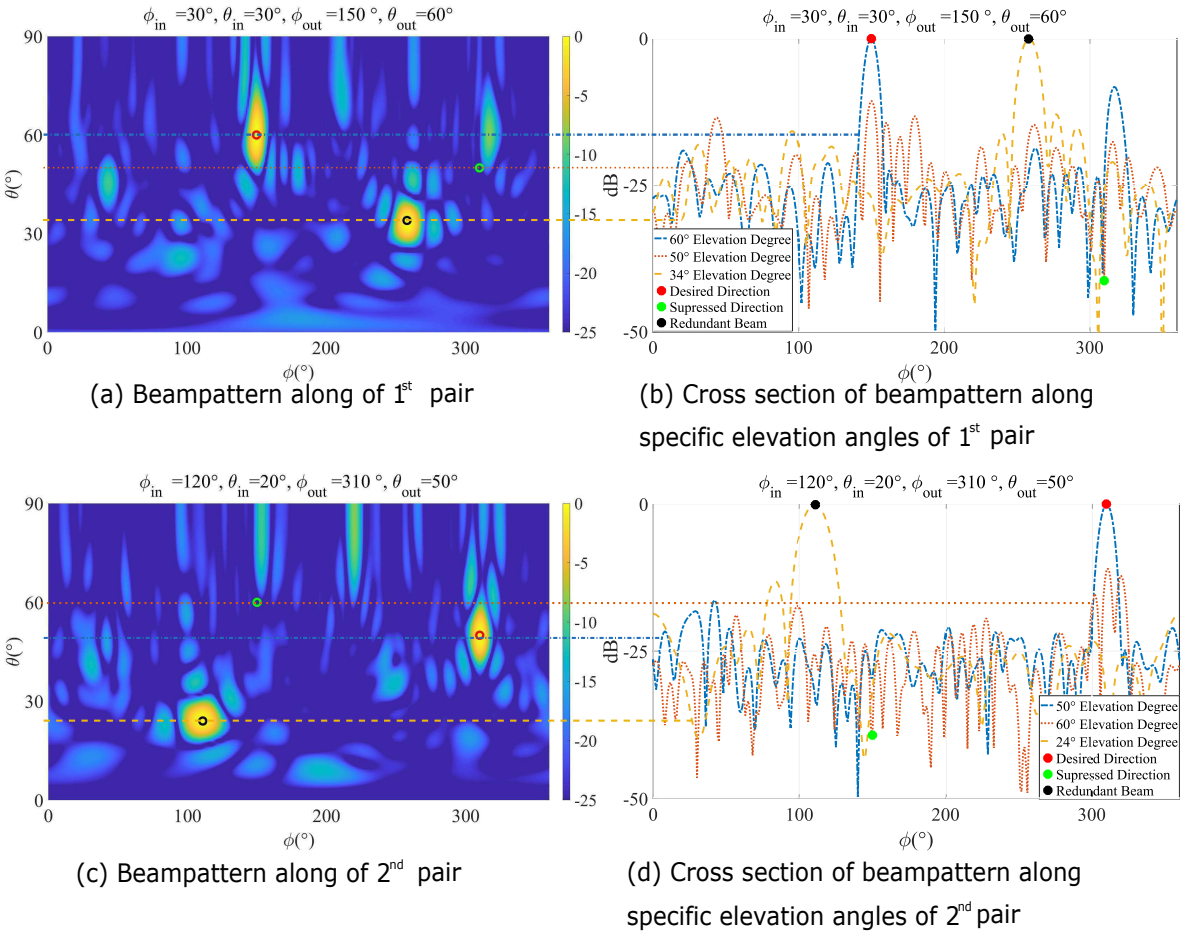
Note that the redundant beam is a side effect of the IRS when multiple pairs of signals are reflected, and we have the following remarks.

Remark 2. *With N pairs of transceivers, up to $N-1$ redundant beams can be generated by the optimal weights and thus may cause significant power loss due to redundant beams splitting the power from a typical direction to other trivial directions. In worst-case scenarios, a multiplexing gain of N can be achieved while the power efficiency will be $1/N$.*

Remark 3. *Under some specific transceivers' directions with relatively small d, there are no redundant beams and thus the power of the signal can be focused on the desired direction. However, with a small d, the beamforming resolution is not ideal as a large one.*

3.4 Simulation results

In the simulations, we use an IRS composed of 16×16 elements with an element distance of $d = \lambda/2$, unless otherwise specified. All elements are isotropic and assumed to have the perfect reflective coefficient, i.e., $|w_m| = 1$. For all simulations, we assume that signal power is normalized and the additive noise at each receiver is set as -10 dB unless otherwise specified.



(a) Beampattern along of 1st pair

(b) Cross section of beampattern along specific elevation angles of 1st pair

(c) Beampattern along of 2nd pair

(d) Cross section of beampattern along specific elevation angles of 2nd pair

Figure 3.1: Beampattern of ACPC solution. (a) Beampattern of 1st pair. (b) Cross section of 1st pair along elevation directions of the signal of interest. (c) Beampattern of 2nd pair. (d) Cross section of 2nd pair along elevation directions of the signal of interest. For the 1st pair, the three curves are about $\theta = 60°$, $50°$ and $34°$ respectively. For the 2nd pair, the three lines correspond to $\theta = 50°$, $60°$ and $24°$ respectively.

Two pairs of transceivers are considered in the simulations and their directions are $\Omega_{in,1} =$

$(30°, 30°), \Omega_{out,1} = (150°, 60°), \Omega_{in,2} = (120°, 20°), \Omega_{out,2} = (310°, 50°)$ unless other specified.

3.4.1 Redundant beams of URA based on ACPC

Firstly, we consider the ACPC solution of the problem (3.14) calculated by the SQP algorithm. According to the corollaries proposed in this Chapter, we can analytically calculate that the redundant beam for 1st pair of the transceiver is at $\Omega_{RB,21} = (258.3°, 33.5°)$ and the redundant beam for 2nd pair is at $\Omega_{RB,12} = (110.6°, 24.4°)$. Fig. 3.1 shows simulation results. The main beam and redundant beam are presented in each sub-figure at left, respectively. As marked in the figure, each main beam is marked by a red circle, each redundant beam is marked by a black circle, and the interference position that needs to be suppressed to make interference-free is marked by a green circle. Through Fig. 3.1, more details for particular θ value can be observed from the beampattern. These figures are the cross-section along the elevation angle of the figures on the left-hand side, and the cut positions of interest have been marked by the dashed lines from the left figures, connected directly to the figures on the right-hand side. For figures on the right-hand side, the red dot marks the response in the desired direction, the black dot marks the response of the redundant beam, and the green dot marks the response in the suppressed direction. Therefore, it is not hard to find, by comparing the two figures on the right-hand side, where the main beam is generated in the direction where the 1st Rx is, the signal of 1st Tx received by the 2nd Rx is suppressed to around -41 dB, while at where that of the main beam is generated for 2nd Rx, the signal from 2nd Tx received by 1st Rx is suppressed to about -39 dB, which is consistent to the desired characteristics of IRS system. It is also observed that a larger number of elements can have more control over the interference suppressing in this case, while a narrower main beam plus redundant beams will also be achieved.

3.4.2 Multi-user scenario of URA

In Fig. 3.2, Fig. 3.3, and Fig. 3.4, we showcase the scenario of three pairs of transceivers. The new added 3rd pair transceiver is located at $\Omega_{in,3} = (190°, 50°)$ and $\Omega_{out,3} = (240°, 10°)$. For lowering side lobes on the beampattern, the optimal AUPC solution is implemented. We observe that by considering an extra 3rd pair transceiver, one more redundant beam is generated, and one more nullified direction is produced on the beampattern of Tx_1 and Tx_2 respectively. At the same time, the redundant beams caused by the constraints of the first two pairs remained at the same direction for Tx_1 and Tx_2. As for positions of redundant beams of 3rd pair transceiver, they can be calculated accurately using our corollary in Section 3.3.3.

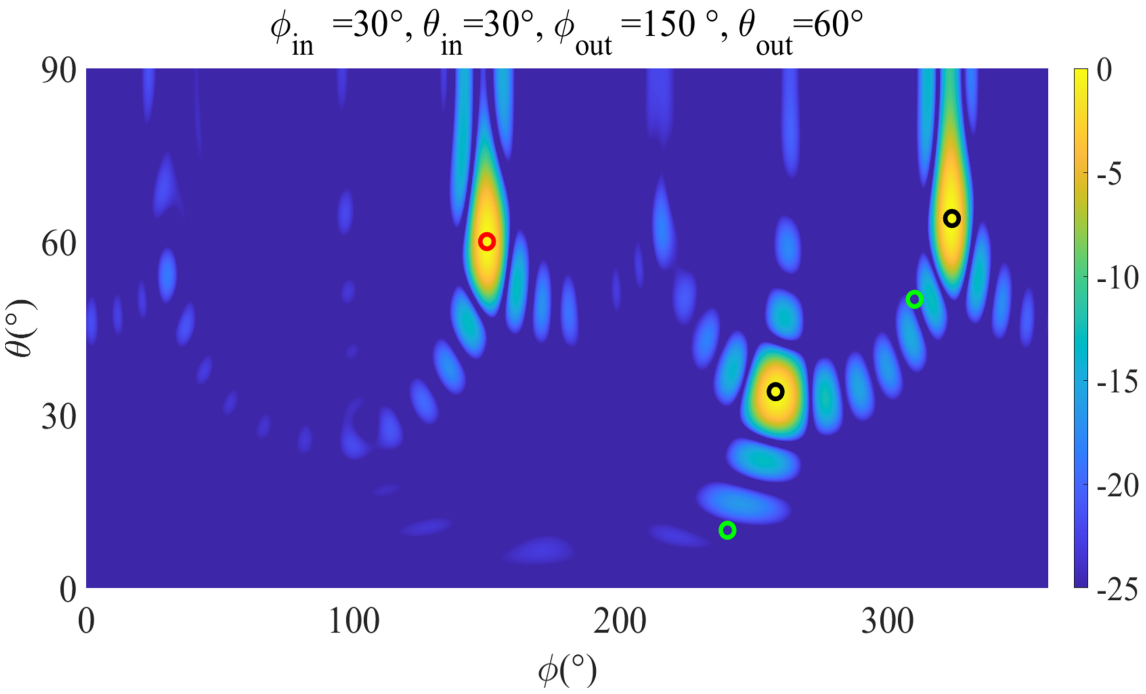


Figure 3.2: Beampatterns of three pairs under AUPC solution, for 1st pair.

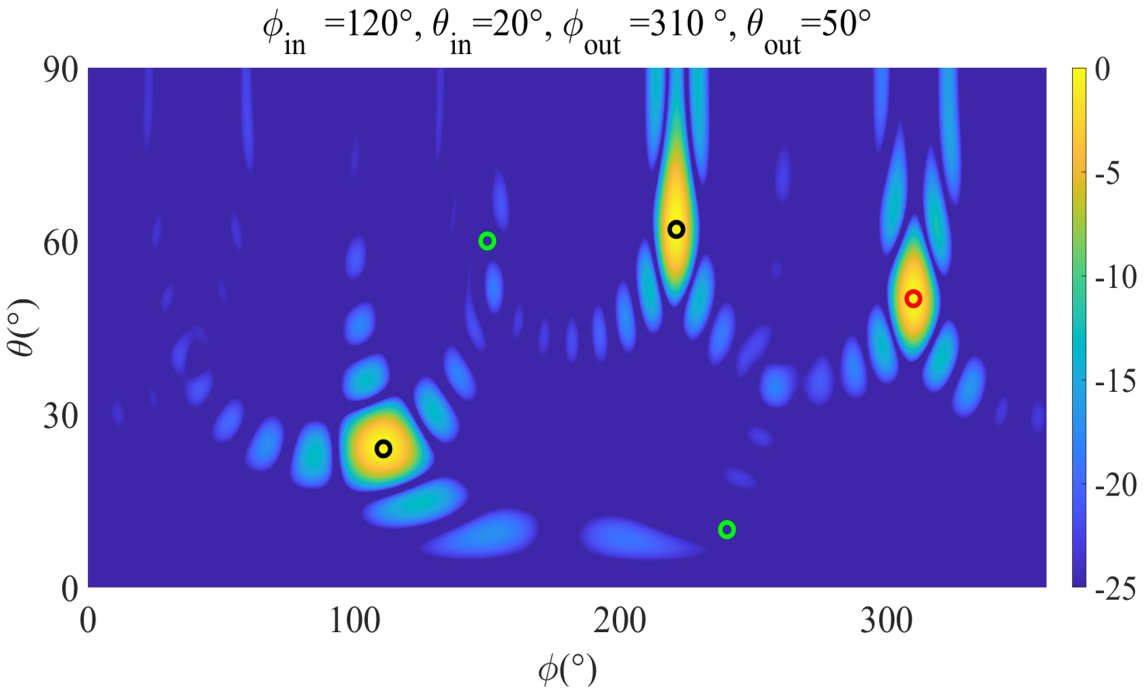


Figure 3.3: Beampatterns of three pairs under AUPC solution, for 2nd pair. (c) Beampattern of 3rd pair.

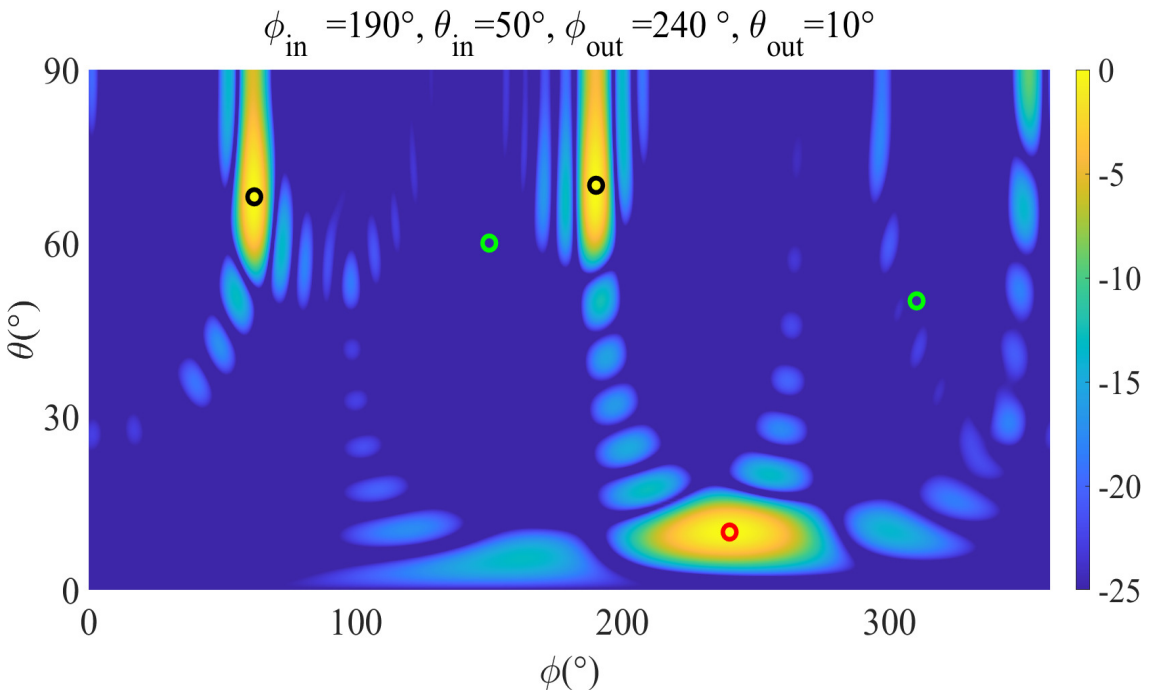


Figure 3.4: Beampatterns of three pairs under AUPC solution, for 3rd pair.

Table 3.3: The performance comparison between different algorithms

Noise power=10dB	Capacity (bps)		Signal Power(dB)		Interference(dB)	
Receiver	Rx1	Rx2	Rx1	Rx2	Rx1	Rx2
GA (ACPC) [139]	11.3	11.3	43.1	43.9	-12.3	-23.9
SQP (ACPC)	10.3	10.6	43.1	43.2	7.6	5.1
MU-LCMV-Phase (ACPC)	9.2	9.2	44.2	44.2	15.3	15.3
MU-LCMV-Amplitude (AXPC)	10.6	10.6	41.9	41.9	-246.3	-262.8

3.4.3 The performance analysis of proposed algorithms

To analyze the performance of the proposed algorithms in Section III, the performances on the capacity, received power and interference are illustrated. In Table 3.3, four algorithms are compared, and three algorithms are considered for solving the problem of ACPC for (P2). They are genetic algorithm (GA), SQP, and MU-LCMV-Phase which directly takes the phase of AUPC solution of (P1). MU-LCMV-Amplitude is shown for reference, which is obtained by normalizing the amplitude of the AUPC solution to provide the AXPC. Due to (P2) being NP-hard, the optimal solution can not be obtained with the polynomial-time algorithms. And no efficient algorithms can guarantee the optimal solution. Since GA is a class of heuristic algorithms that can provide a nearly optimal solution where the local optima can be avoided by sufficient times of exhaustive searching, we use GA as a near-global optimal benchmark. We can find the performance of capacity achieved by SQP is slightly lower than the benchmark GA, where the difference for capacity is less than 10% from the benchmark. However, GA's searching cost is much higher than other algorithms, which cannot be acceptable in real applications. To be noted, MU-LCMV-Phase and MU-LCMV-Amplitude achieve the highest received power and minimal interference power at receivers, respectively. Although MU-LCMV-Phase only achieves 80% percent capacity of GA, it is obtained analytically. MU-LCMV-Amplitude achieves interference-free transmission. It is worth mentioning that both MU-LCMV based solutions can be obtained with low-complexity algorithms since they are based on the analytical solution.

Fig. 3.5 shows the error of redundant beam between the analytical calculation through **Corollary 5** and the simulation of AUPC and ACPC solutions. In this figure, lines drawn in the same color are under conditions of the same element distance of IRS but different solutions of the algorithm. As mentioned above, systems with small element distances have a worse resolution, so they are more likely to get a larger error on the direction of redundant beams. The error reduces with the increase of elements, reaching a plateau at a range of large element numbers. In addition, AUPC outperforms ACPC in accuracy performance.

3.4.4 Redundant beams of URA based on ACPQ

In this simulation, we verify the performance of the ACPQ solution. The performance loss can be conceived by comparing Fig. 3.7, Fig. 3.8, Fig. 3.9, and Fig. 3.10 to Fig. 3.1. Due to the limitation of phase resolution which is 1 bit in this simulation, there are more side lobes occurred, which implies worse power efficiency. These side lobes are called quantized side lobes and have been verified by real experiments in work [140]. Nevertheless, we claim again that critical differences lie between redundant beams and sidelobes, where the former is not

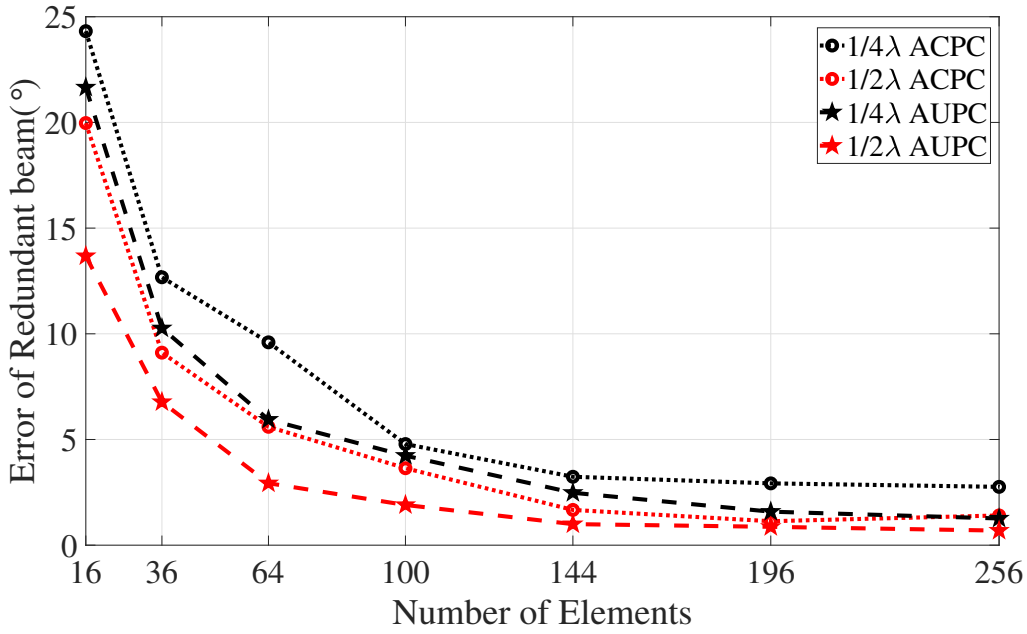


Figure 3.5: Redundant beam error between analytical calculation and simulation for different solutions.

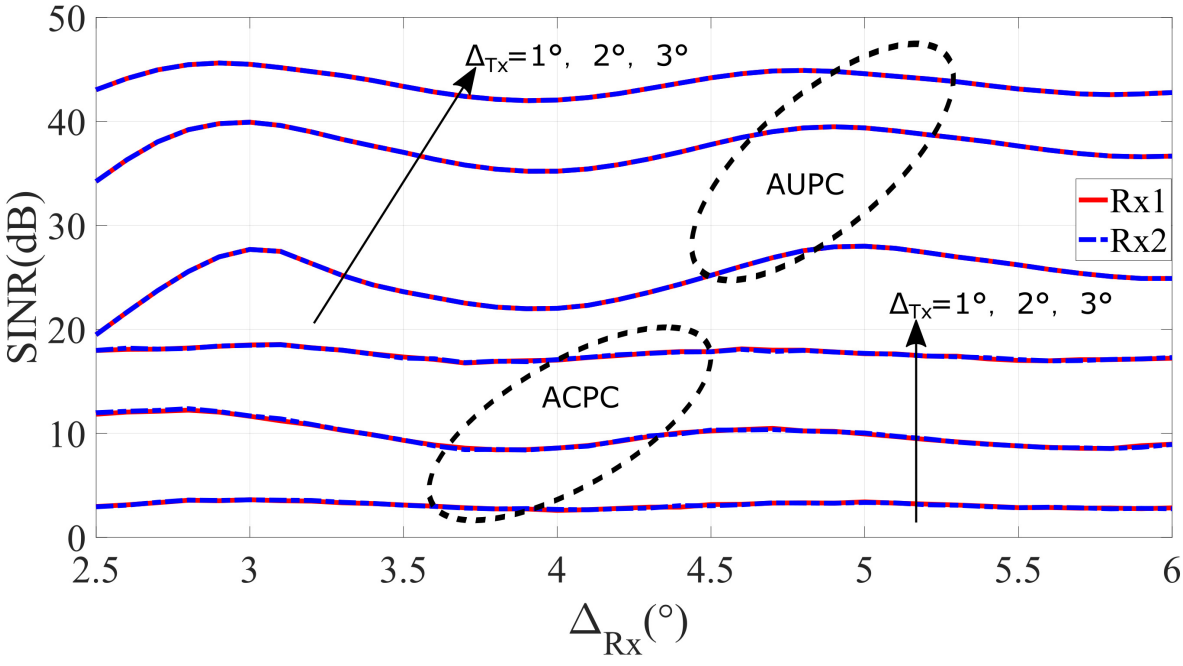


Figure 3.6: SINR vs the angle difference of transmitters (and receivers).

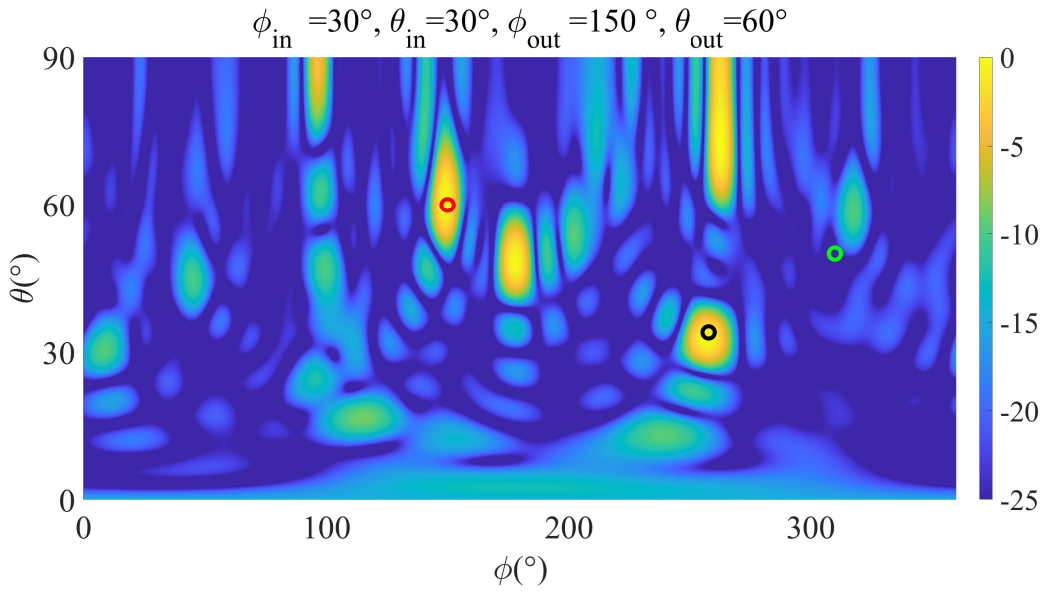


Figure 3.7: Beampattern of ACPQ solution for 1st pair.

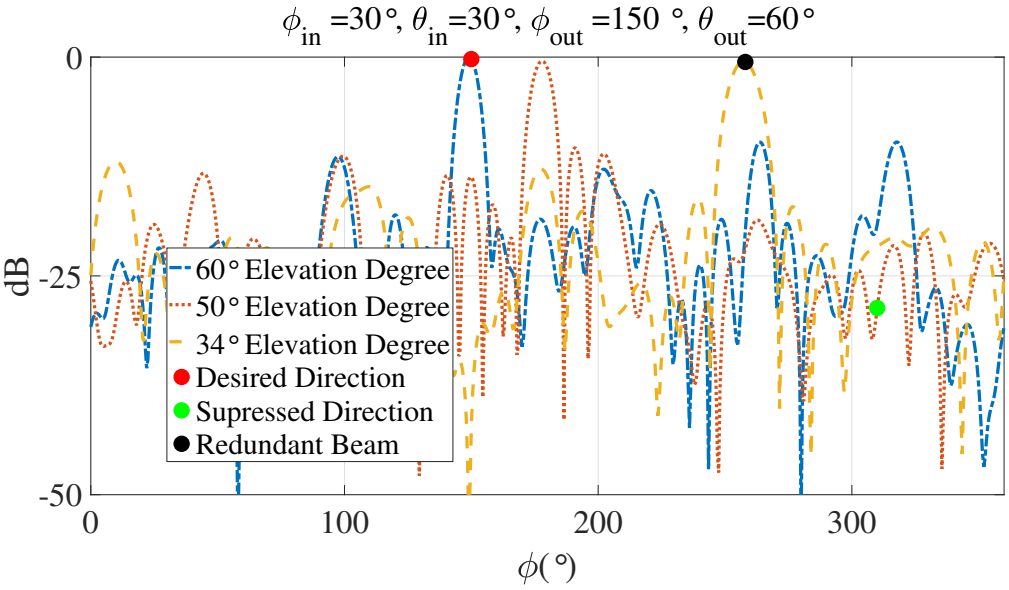


Figure 3.8: Beampattern cross-section of ACPQ solution for 1st pair.

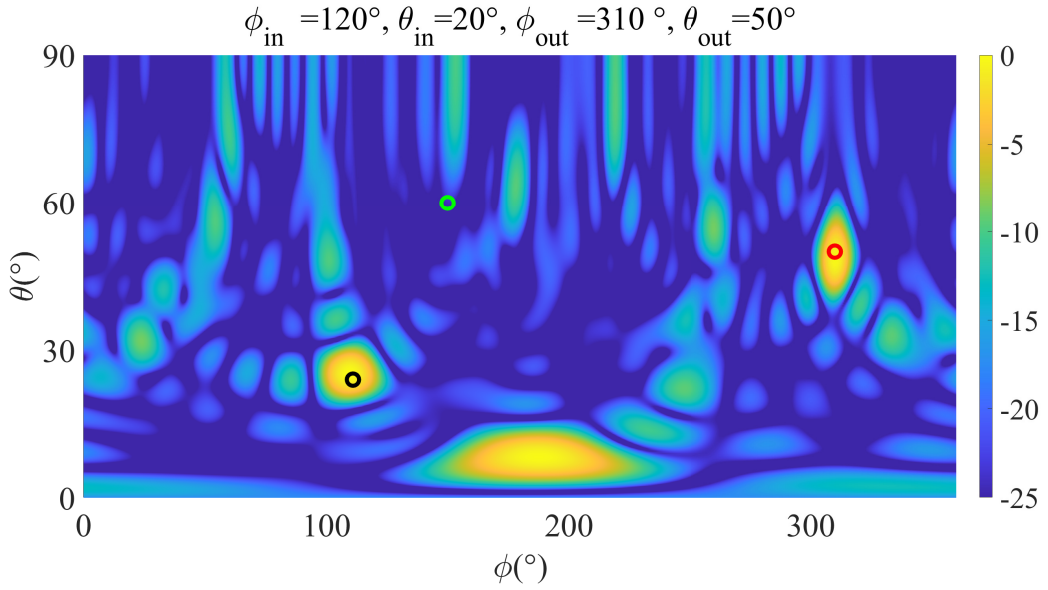


Figure 3.9: Beampattern of ACPQ solution for 2nd pair.

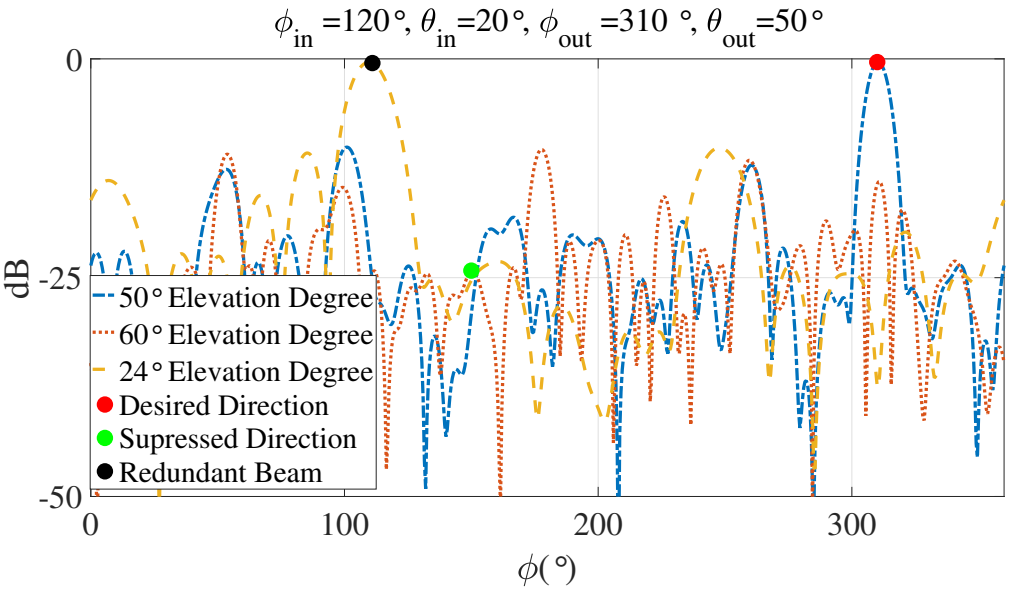


Figure 3.10: Beampattern cross-section of ACPQ solution for 2nd pair.

able to be eliminated while sidelobes are able to be mitigated.

Other solutions can also be numerically obtained from other approximations and discrete optimization [141, 142] and the optimal solution can be obtained by exhaustive searching. Nevertheless, in those critical directions where marked by circles, the characteristics of redundant beams are still preserved, which can also validate our theorem about redundant beams. As we can observe from Fig. 3.7 and Fig. 3.9, the beams marked by red circles keep an almost ideal maximum response. That marked by green circles indicates an acceptable interference suppressed, which is around −25 dB, and the redundant beams still keep the same level response as the main beam.

3.4.5 SINR performance

Since the algorithm aims to provide equal SINR for different users located in different directions with different channel correlations, to show the robustness of the algorithm, this simulation evaluates the SINR performance versus the angle difference between two users by calculating equation (2.9). For simplicity, we use ULA IRS with 64 elements in this simulation. $\phi_{in,1} = 30^\circ$ and $\phi_{out,1} = 100^\circ$. The distance between each element is set to $\lambda/2$. We define $\Delta_{\mathrm{Tx}} = \phi_{in,2} - \phi_{in,1}$ and $\Delta_{\mathrm{Rx}} = \phi_{out,2} - \phi_{out,1}$ as the angle differences of transmitters and receivers respectively. As AUPC is unconstrained on amplitude, we define the norm of AUPC solution to be consistent with ACPC so they are on the same scale. As shown in Fig. 3.6, with the angle difference of the transmitters Δ_{Tx} and receivers Δ_{Rx} increasing, the performance of SINR of both Rx1 and Rx2 are plotted. Their SINR improves and reaches a peak value when the angle difference is sufficiently large. The lines marked by the same color are under the same transceivers' position, and lines marked by the same symbol are obtained by the same way of solution. Apparently, AUPC apparently outperforms ACPC on SINR. The fluctuation after the peak value is due to the spatial correlation between the merged steering vectors of the desired direction and the interference direction varies. For results from AUPC weights, it realizes the optimal upper bound of SINR when the direction difference is large. While for ACPC, the upper bound is lower due to the amplitude of weights is constrained and thus beamforming angle resolution is limited. Therefore we can conclude that close angles between each transmitter or receiver achieve low SINR. The error still exists as long as the number of elements is limited.

3.4.6 Half-power Beamwidth

The analytical expressions of HPBW in equation (3.33) conform to the simulation result. As shown in Fig. 3.11, HPBW increases as the incident angle moves away from 90° and reaches

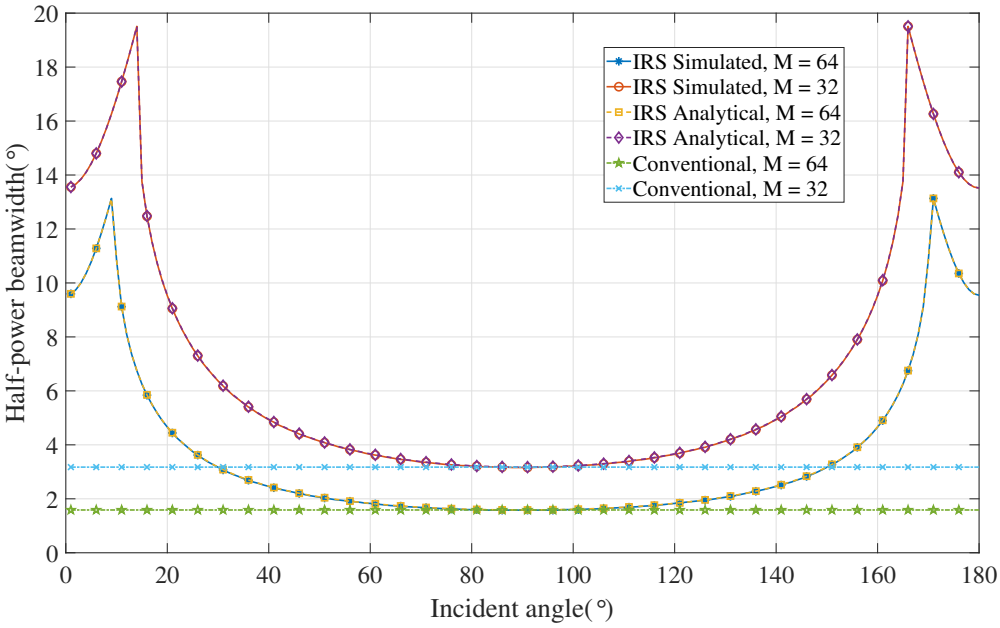


Figure 3.11: HPBW of IRS with different incident angle

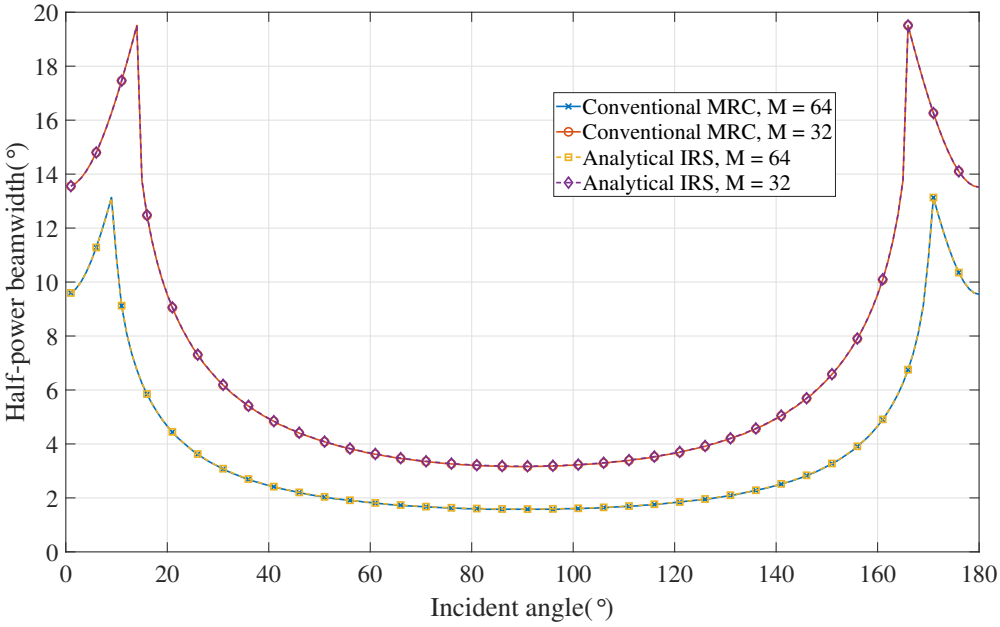


Figure 3.12: HPBW of conventional array applying MRC and IRS

the maximum at $\theta_{in} = a, b$ determined by equation (3.33). As the incident signal gets closer to the surface of IRS, more energy of the signal will be dissipated, resulting in a decrease in HPBW. Besides, HPBW decreases while M increases, which indicates that the greater the number of elements, the better the directivity will be. The HPBW of a conventional array, with the same size and element spacing to IRS, is also plotted for reference [22]. Although the two systems are different, their models are equivalent at vertical incidence, which explains why the two systems' HPBW are equal at $\theta_{in} = 90^\circ$ while the HPBW is greater than that of a conventional array in the other incident conditions. We can also find that when the reflect and transmit angles are the same, the antenna array's HPBW with MRC applied will be equal to HPBW in equation (3.33), as shown in Fig. 3.12. Besides, unlike the HPBW of conventional active antenna array which is the overall response to the incident signals with different incident angles, the HPBW of IRS will change with incident angles. Also, larger d causes smaller HPBW and vice versa.

3.5 Summary

Multi-user beamforming and transmission based on IRS are investigated in this chapter. Optimization problems are formulated to obtain the optimal weight vector for different scenarios. A closed-form solution is first derived by using LCMV beamformer based on AUPC configuration and the AXPC solution can be correspondingly obtained via linear scaling of AUPC. By constraining the amplitude and phase of the weight vector, ACPC solution is obtained for practical IRS implementation with SQP algorithm. A critical observation of the thesis is that redundant beams exist in the system and brings significant power loss. We have mathematically analyzed the redundant beams under different IRS configurations to provide useful design guidance for in-depth analysis. In addition, the HPBW of IRS is analyzed and compared with that of the conventional array. Results verified the finding and algorithms and suggested using a large angle difference between the Txs (and Rxs).

Chapter 4

Sum Rate Limit Analysis of IRS Networks

As a MU beamforming design has been given in previous chapters, with in-depth analysis on redundant beams, beamwidth and sum rate performance. In this chapter, we aim to investigate the reachable sum rate limit for IRS networks. Note that the IRS networks can form different topologies and result in different sum rate limits. Hence, we will first analyze the single IRS model which is the basic unit of the IRS networks model. To guarantee a reachable sum rate for single IRS and networks, the conditions for realizing maximal power and minimal interference are necessary. Therefore, for single IRS, two critical conditions are derived mathematically: the optimal condition to reach the sum rate upper bound and the condition to realize interference-free transmission. Leveraging these conditions, the sum rate upper bound of MUMOR transmission assisted by an IRS network is analyzed, with different topologies of the IRS network. Specifically, graph decomposition is implemented to the graphs for analyzing the maximized SE and EE. The required channel estimation strategy for the IRS network channel model will also be introduced.

4.1 Limit of Single IRS System

4.1.1 Optimal MU Sum-rate

In case of single IRS, where $\Gamma = 1$, $K = 1$ in equation (2.34), then the received signal for all Rxs becomes:

$$\mathbf{y} = \mathbf{H}_{I,1}\mathbf{s} + \mathbf{n} = \mathbf{A}_{out,1}^{T}\mathbf{W}\mathbf{A}_{in,1}\mathbf{s} + \mathbf{n}\,, \tag{4.1}$$

where

$$\mathbf{H}_{I,1} = \begin{bmatrix} \mathbf{w}^H \mathbf{a}_C(\phi_{out,1}, \phi_{in,1}) & \dots & \mathbf{w}^H \mathbf{a}_C(\phi_{out,1}, \phi_{in,N}) \\ \mathbf{w}^H \mathbf{a}_C(\phi_{out,2}, \phi_{in,1}) & \dots & \mathbf{w}^H \mathbf{a}_C(\phi_{out,2}, \phi_{in,N}) \\ \vdots & \ddots & \vdots \\ \mathbf{w}^H \mathbf{a}_C(\phi_{out,N}, \phi_{in,1}) & \dots & \mathbf{w}^H \mathbf{a}_C(\phi_{out,N}, \phi_{in,N}) \end{bmatrix}. \tag{4.2}$$

With equal power P_T from all transmitters, the channel capacity of MU transmission on a single IRS can be expressed as

$$C = log \det(\mathbf{I}_N + \frac{P_T}{N_0} \mathbf{H}_{I,1} \mathbf{H}_{I,1}^H) \, . \tag{4.3}$$

Then, the optimization on weights $\mathbf{w}$ is equivalent to maximizing the diagonal terms and minimizing the off-diagonal terms in the equation (4.2). However, it is hard to decide whether the main diagonal terms and the off-diagonal terms of $\mathbf{H}_{I,1}$ can be simultaneously maximized and minimized, i.e, $|\mathbf{w}^H \mathbf{a}_C(\phi_{out,i}, \phi_{in,i})| = M$ and $|\mathbf{w}^H \mathbf{a}_C(\phi_{out,i}, \phi_{in,j})| = 0$, $i \neq j$. This is due to the transceivers being located in different directions with respect to IRS, the IRS aided channels for transceivers are different. There might be a tradeoff on whether the signal power should be maximized or interference should be minimized to reach the sum rate limit by relying on a specific algorithm for designing the weights vector.

Note that once the spatial correlation between each transceiver pairs $\mathbf{a}_C(\phi_{out,u}, \phi_{in,v})$, $u, v = 1, ..., N$ is fixed, we can calculate $\mathbf{w}$. As the IRS channel is deterministic with fixed $\mathbf{w}$, the spatial correlation between all transceiver pairs is another dominating factor in deciding the sum rate upper bound. For example, the higher the spatial channel between Tx_i and Tx_j or between Rx_i and Rx_j, the lower the channel ranks and singular values of $\mathbf{H}_{I,1}$ are, which further lower the upper bound of the overall sum rate in a specific spatial realization. To find an optimal upper bound of the sum rate, we derive the optimal condition in the spatial correlation between each transceiver pair. In this case, every pair can leverage the optimal gain brought by the single IRS. Besides, the interference between each pair can be nullified simultaneously.

4.1.2 Conditions on Directions of Transceivers and Spacings

Let the i-th pair user locate at $\phi_{in,i} = \alpha_i, \phi_{out,i} = \beta_i$ and the j-th pair locate at $\phi_{in,j} = \alpha_j, \phi_{out,j} = \beta_j$ where $i \neq j, i, j = 1, 2, ..., N$. Denote $\Delta r = \frac{d}{\lambda}$ as the normalized spacing between each element since d is the distance between each element and λ is the carrier wavelength. Additionally, denote $L = M\Delta r$ as the relative length with respect to normalized spacing. Then we have

$$w_m = e^{-j\zeta_m kdm} \, , m = 1, \dots, M \ , \tag{4.4}$$

where

$$\zeta_m = -\cos\alpha_i - \cos\beta_i + \frac{K}{\Delta r} \tag{4.5}$$

is the optimal factor given by MRC algorithm[1] to realize the power gain of the i-th pair user, i.e., $|\mathbf{w}^H \mathbf{a}_C(\phi_{out,i}, \phi_{in,i})| = M$.

Lemma 2. *Given N pairs of transceivers assisted by a single piece IRS, the optimal sum rate upper bound for N transceivers aided by first order reflection can be obtained when each transceiver pair's position for Tx and Rx are at*

$$\alpha_j = \cos^{-1}\left(\frac{j}{L} - \zeta_m - \cos\beta_i \pm \frac{1}{\Delta r}\right) \tag{4.6}$$

and

$$\beta_j = \cos^{-1}\left(\frac{j}{L} - \zeta_m - \cos\alpha_i \pm \frac{1}{\Delta r}\right) \tag{4.7}$$

respectively.

Lemma 2 reveals that if the position of each transceiver can be coordinated correspondingly, $\mathbf{H}_{I,1}$ can be optimized such that diagonal terms can be maximized and off-diagonal terms can be nullified respectively at the same time. And it also observed that once one of the transceiver pairs is fixed in its position, other optimal positions for other pairs are fixed to satisfy the optimal spatial correlation. Physically, each reflected beam towards each Rx is orthogonal to each other. The whole IRS channel can be orthogonal space-division multiplexed (OSDM) by M pairs of the transceiver. Thus, we call links with correlation obeying Lemma 2 as the optimal link. Proof of Lemma 2 is shown in Appendix A.3. Additionally, without changing L, no matter how the element number and spacing vary, the optimal condition in Lemma 2 will not change since the nullifying point in the visible angular domain and the periodicity of channel response is now determined by L, which can also be viewed from equation (A.40) in Appendix A.3. In fact, the characteristic of channel rank is essentially proportional to L [4]. Therefore, we propose to use $d = \frac{\lambda}{2}$ since this is the maximal spacing for a fixed L to secure the narrowest reflected beam, which causes no grating lobe of the reflected beam. Note that there is a trade-off between energy efficiency and spacing as well. This is due to smaller spacing resulting in fewer channel ranks and larger beamwidth. Still, the redundant beam, causing energy waste in trivial directions, will less likely occur [143]. Thus, the actual spacing can be less than this value based on different design criteria. Some discussions about the spacing of elements and the beamwidth of IRSs can be referred to in [122,138,144]. With fixed L, M, and half-wavelength spacing, we have

[1] Since constant modulus constraint is still existed, the MRC now is equivalent to the co-phasing combining.

Theorem 2. *Given all transceivers are optimally positioned, the upper bound of the sum rate with N pair transceivers for a single IRS of M elements is*

$$C_{Max} = \sum_{n=1}^{N} log(1 + \frac{P_T M^2}{(d_{in,n} d_{out,n})^2 N_0}), \text{ if } N \leq M \,, \tag{4.8}$$

where P_T is the power of Txs, N is the spatial multiplexing gain and N_0 is the noise power at the Rxs, and $d_{in,n}$ and $d_{out,n}$ represent the incident distance and exit distance on the single IRS for the n-th pair transceiver.

Based on Theorem 2, the sum rate upper bound is reached when $N = M$ and each pair receives the power gain of M^2. If normalized power from the Tx is considered without path loss, then the power gain should be 1 since the whole IRS network is a passive system. When $N > M^2$, user interference is unavoidable, and now the sum rate users should be determined specifically by the spatial correlation of transceivers and the ratio between N and M. Just for the completeness of this work, other multiplexing schemes are proposed to avoid inter-user interference if $N > M$. In addition, we notice that the path loss is only related to the distance between transceivers and IRS. For simplicity, we can omit $(d_{in,n} d_{out,n})^2$ and consider the path loss as different constant values for specific analysis.

Note that the upper bound in Theorem 2 is hard to achieve as transceivers can not always stay at the optimal position in Lemma 2. Moreover, the element spacing may be less than half wavelength, and the mutual coupling effect can be an issue [145]. Nevertheless, Theorem 2 is meaningful as it analytically provides a sum rate upper bound for each IRS and can only be obtained by satisfying the optimal condition in Lemma 2.

4.1.3 Interference-free Condition Based on A Single IRS

When the spatial correlation between transceiver pairs is not orthogonal, with one IRS, we can nullify the interference to achieve interference-free transmission.

Lemma 3. *To achieve interference-free transmission without orthogonal spatial correlations between each transceiver pair, the element number on a single IRS should satisfy $M \geq N^2$.*

The interference-free condition can be easily achieved in practical deployment since each IRS can have a sufficient amount of elements. The proof is given in Appendix A.4, where we also show a single IRS is able to support multiple stream transmission with only one vector. As there is a similarity and equivalence in the function between IRS and MIMO precoding/decoding, with the deployment of the IRS, transceivers can transfer some workloads

[2]However, this case is rare in the real situation since $N \leq M$ can be guaranteed as there can be hundreds of thousands of IRS elements while keeping the far-field condition [123].

to the IRS. Thus the structure of transceivers can be simplified. Nevertheless, IRSs still have unique advantages over the traditional MU scheme compared with the traditional scheme. In particular, the IRS can suppress the inter-user interference before Rxs are jammed, while the traditional scheme can not conveniently suppress the inter-user interference at Rxs due to joint decoding is usually not available.

Remark 4. *The number of transceivers that access the IRS network from a single IRS should be significantly below the number of elements of that a single IRS. In addition, it is better if transceivers can locate differently in directions with respect to each other and the single IRS for lowering the spatial correlation of channels between each transceiver pair. Extra pieces of IRS nearby should be involved to solve high spatial correlation (spatial aliasing) since extra pieces of IRS can distinguish these transceivers from a much different location.*

4.2 Limit of MU Sum-rate Within Multi-hops IRS Network

For simplicity, we use the terminology of graph theory for the following discussion [146]. We call an LoS channel an edge, a single IRS/transceiver as a node, nodes connected to one node by an edge as adjacent nodes, the number of edges that are incident to a node as the degree, a C-LoS path as a path, the number of IRS nodes that the path passes through as the path length or simply length, and the IRS network as the network.

In the network, the sum rate is affected by the network's topology and geometry, the number of IRS/transceiver nodes, and IRS nodes' weights design. The topology is the connection statement of nodes by edges existing within the network, while the geometry is determined by the relative AOA and AOD between arbitrary two nodes. Thus all nodes and edges have a specific topological and geometrical relationship with each other, as shown in Theorem 1. However, it is difficult to derive the exact sum rate upper bound without prior determining the network topology, geometry, and weights design.

Note that, Theorem 2 indicates each IRS node can fulfill the criteria of power maximization and interference nullification given the optimal condition in Lemma 2. To maximize the EE and SE performance, we leverage Lemma 2 to determine the network's geometry, topology and weights design. In particular, each IRS node can optimally serve other adjacent IRS/transceiver nodes, where maximally M pair of adjacent nodes can be supported, or $2M$ degrees can be possessed by one IRS node. As the topology is versatile given K nodes of IRS to form a network, we derive the sum rate upper bound for two kinds of common

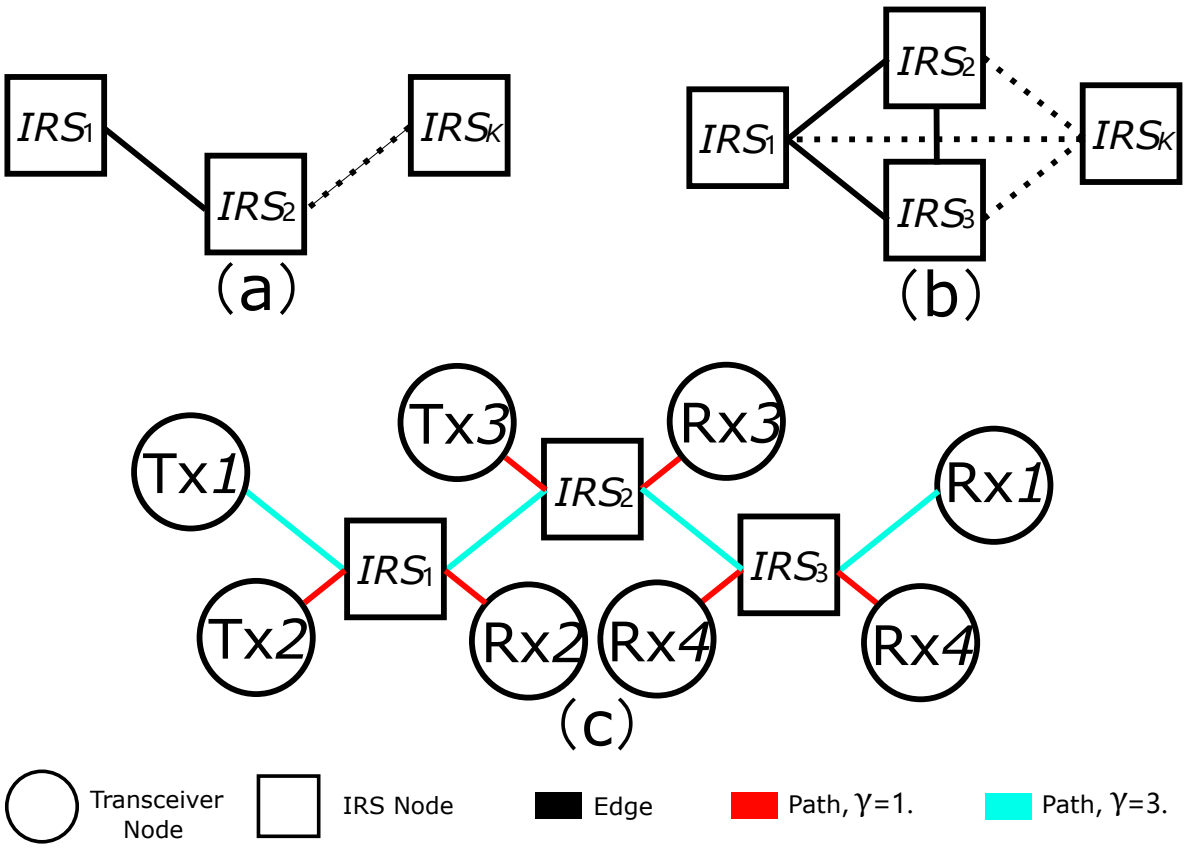


Figure 4.1: MUMOR Transmission based on IRS network with K single IRS. (a) An LG topology. (b) A CG topology. (c) An example of a network shaping an LG to serve MU where K=3, $M = 2$, $\Gamma = 3$, N=4. The solid line represents the edge that connects two adjacent nodes. The dashed line represents a series of other adjacent connections that are omitted.

graphs, which are linear graph (LG)[3] and complete graph (CG)[4], as shown in Fig. 4.1(a) and Fig. 4.1(b) respectively. In addition, a special topology without edges is considered, which is called the null graph (NG)[5] and means each IRS only forms FOR paths locally without LoS between any two IRS nodes in the network.

4.2.1 The IRS Network In Linear Graph

With the signal of Tx_i passing along an LG with a length of K order, where $\Gamma = K$, we can write the received signal of Rx_i after K orders reflection as

$$y_i = \mathbf{a}(\phi_{out,i,X_{K,1K}})[\prod_{v=1}^{K-1} \mathbf{W}_{X_{K,1(v+1)}} \mathbf{E}_{X_{K,1v}X_{K,1(v+1)}}]\mathbf{W}_{X_{K,11}} \mathbf{a}(\phi_{in,i,X_{K,11}})s_i + n_i, \tag{4.9}$$

where $\mathbf{a}(\phi_{out,i,X_{K,1K}})$ and $\mathbf{a}(\phi_{in,i,X_{K,11}})$ are the corresponding steering vector in matrix $\mathbf{A}_{out,X_{K,1K}}$ and $\mathbf{A}_{in,X_{K,11}}$ for i-th pair transceiver. Since each transceiver pair now communicates orthogonally in the network following optimal conditions, the index matrix $\mathbf{X}_K$ now is simplified to contain only one row, holding one specific sequence of one C-LoS path. Note that the dual reflection exists within the LG network as well. As $\Gamma = K$, there is only one path with the maximal effective length that can reach Rx_i. Moreover, equation (4.9) can be written in a similar form with equation (2.7) such that

$$y_i = [\prod_{v=1}^{K} \mathbf{w}_{X_{1v}}^H \mathbf{a}_{C,i,X_{1v}}]s_i + n_i, \tag{4.10}$$

where $\mathbf{a}_{C,i,k}$ means the equivalent channel of IRS_k for i-th pair transceiver and $\mathbf{w}_k = diag(\mathbf{W}_k)$ is the corresponding weights vector on IRS_k for $k = 1, ..., K$. As the optimal power gain for a single pair transceiver is M^2 from a single IRS, with K order reflection where each IRS applying weights to realize maximal power gain in equation (4.10), the cascaded power gain would be M^{2K}. In this case, EE is maximized for a single pair in an LG network. Thus, based on equation (4.10), the sum rate upper bound for one transceiver pair is

$$C_{SU,LG,(K)} = log(1 + \frac{P_T M^{2K}}{N_0}) , \tag{4.11}$$

[3]**Linear Graph/Path Graph**: a linear graph is a graph whose vertices/nodes can be listed in the order $v_1, v_2, ..., v_n$ such that the edges exist between v_i and v_{i+1} where $i = 1, 2, ..., N$. Paths are often important in their role as subgraphs of other graphs, in which case they are called paths in that graph.

[4]**Complete Graph**: A complete graph is one in which every two vertices/nodes are adjacent: all edges that could exist are present.

[5]**Null Graph**: A null graph, or an empty graph, is a graph in which there are no edges between its vertices.

where the subscripts SU, LG, and (K) mean a single pair, linear graph, and a K-order reflection. Since an edge between two nodes is a rank-one channel from Lemma 1, we cannot realize multi-stream information transmission based on one edge, and thus the cascaded channel of an LG is rank one. Nevertheless, MU transmission in an LG network is still available as each IRS node can have $2M$ degrees. With the topology of the network, as shown in Fig. 4.1(c), the sum rate upper bound can be reached by combining three 1-length paths and a 3-length path where each IRS node has 4 degrees. By including the sum rates from all 1-length paths and one K-length path, we have the MU sum rate upper bound as

$$C_{MU,LG,(1,K)} = log(1+\frac{P_T M^{2K}}{N_0}) + K(M-1)log(1+\frac{P_T M^2}{N_0}), \tag{4.12}$$

where the power of all Txs is equal to P_T and the subscript $(1, K)$ means only the paths whose lengths are equal to 1 and K are involved. In this case, spatial multiplexing has been maximized while these paths would not introduce extra interference from reflections or dual reflections since all paths still keep spatially orthogonal.

4.2.2 The IRS Network In Complete Graph

For CG networks, though multiple paths can be leveraged by one pair transceiver, this is equivalent to transferring spatial multiplexing into power gain, which introduces a trade-off. For the maximal spatial multiplexing gain of the network, each transceiver should send one stream via one path. Thus, the network sum rate limit is determined by the number of Eulerian paths[6] without revisiting nodes in the CG network. Eulerian paths with revisiting nodes are excluded because these transmissions are unnecessary for the network. To clarify the number of paths with different lengths in the network, we denote N_γ as the number of transceiver pairs whose Eulerian paths have γ-length. Thus, following equation (4.11), we have

$$C_{MU,CG,(1,\dots,\Gamma)} = \sum_{\gamma=1}^{\Gamma} N_\gamma log(1+\frac{P_T M^{2\gamma}}{N_0}), \tag{4.13}$$

which is the sum rate upper bound of MUMOR transmission assisted by the CG network. Note that the number of total transceiver pair N that can achieve interference-free transmission is a variable, where

$$N = \sum_{\gamma=1}^{\Gamma} N_\gamma. \tag{4.14}$$

[6]**Eulerian path**: or Eulerian trail is a trail in a finite graph that visits every edge exactly once (allowing for revisiting vertices/nodes).

Since there are multiple ways to decompose a CG into the different number of Eulerian paths with different lengths, the value of N_γ, $\gamma = 1, 2, ..., \Gamma$ are to be determined by a specific graph decomposition. To rewrite N in a general expression, we decompose the CG into paths where all of their lengths are equal to γ. To ensure the upper bound is reached at maximal SE, all these Eulerian paths should pass through all edges. Note that a class of graph decomposition problem is introduced here, which is determining if the CG network can be completely decomposed into paths of γ length equally, which has been proven to be NP-complete [147]. Therefore, it is hard to determine N_γ and write N_γ in a general expression.

In order to obtain a general expression of the sum rate upper bound of CG network, we denote $\Lambda_i, i \in [1, N]$, as the path length for i-th pair transceiver, and we consider $\Lambda_i = \tau, i = 1, ..., N$, and $\Gamma = \tau > 1$, where τ is a specific value of path length. In addition, we denote

$$N_\tau = \frac{\binom{K}{2}}{\tau - 1} = \frac{K(K-1)}{2(\tau-1)}, \tag{4.15}$$

where $\binom{K}{2}$ is the total edges' number of a K-nodes CG. To completely decompose the CG, we should satisfy

$$N_\tau \in \mathbb{Z}\,, \tag{4.16}$$

as it is a necessary and sufficient condition for the existence of an edge-disjoint decomposition of a K-nodes CG into simple isomorphic paths consisting of $(\tau - 1)$ edges each [148]. With $N_\tau \in \mathbb{Z}$, the edge number of a CG can be equally divided up into paths with τ-length. Thus, N_τ is the multiplexing gain while the cascading power gain of a corresponding pair is $M^{2\tau}$. Following equation (4.13), the sum rate now becomes

$$C_{MU,CG,(\tau)} = N_\tau log(1 + \frac{P_T M^{2\tau}}{N_0}), \tag{4.17}$$

which is the sum rate upper bound for the MUMOR transmission for N_τ pairs transceivers with length of τ. By combining the sum rate upper bound of 1-length paths, we have

$$C_{MU,CG,(1,\tau)} = N_\tau log(1 + \frac{P_T M^{2\tau}}{N_0}) + (KM - N_\tau \tau) log(1 + \frac{P_T M^2}{N_0}), \tag{4.18}$$

and now the upper bound is reached for $N = KM + N_\tau(1 - \tau)$ pairs of transceiver. The form of the second term can be derived similarly as to derive equation (4.12).

For the sum rate upper bound in a general case, i.e., path lengths are different for different pairs, the sum rate upper bound can still be computed as long as the graph decomposition is determined. Then, the value of N_γ is fixed, and the sum rate upper bound can be computed

using equation (4.13).

4.2.3 The IRS Network In Null Graph

When $\tau = 1$, since the graph of the network has no edges, we can call it a null graph (NG). In this case, each IRS serves a local network in different cells, and no edges connect any two IRS nodes. The sum rate upper bound can be straightforwardly obtained from Theorem 2 as $C_{MU,NG} = KC_{Max}$, which is directly scaled by K-folds. Since each IRS node is isolated locally, inter-user interference is not induced.

Also, as one proof has been shown in [4] that leveraging Jensen's inequality, we know at low SNR, the sum rate reaches an upper bound if equal decomposition is realized for the K nodes CG with largest τ. In addition, at high SNR, the upper bound is reached with $\tau = 1$.

4.3 Discussion on Channel Estimation for IRS Networks

For wireless transmission in MUMOR networks, it is required to have the CSI. Practically, acquiring CSI requires sophisticated channel estimation by sending training pilot through direct BS-user link and BS-IRS-user link [149]. During channel estimation, the on-off statement and phase shifts of elements are jointly controlled in a predefined way by BS along with the training pilot sent by transmitters [67, 150–152]. Once the channel is estimated, it would be sent to IRS via the internet link between BS-IRS for beamforming [153]. Since the channel estimation will not be the focus of the thesis, the global CSI availability at each IRS is commonly assumed as did in lots of works [43, 44, 64, 85, 150–153].

A basic channel estimation strategy for obtaining the relative parameters of IRS networks is provided. In addition, we discuss the scenario with non-IRS scatters/reflectors presented in the transmission environment. Specifically, we mainly leverage the geometry of scatters' position, beam sweeping and feedback which will be discussed as follows. First, the channel estimation for the IRS networks can be decomposed into several tasks where different methods can complete each. Specifically, they are

- Task 1. Locate the UE and scatters.
- Task 2. Measuring the maximum reflection order, Γ.
- Task 3. Measuring the scatters' weights and orientation.
- Task 4. The MIMO channel establishment and validation.

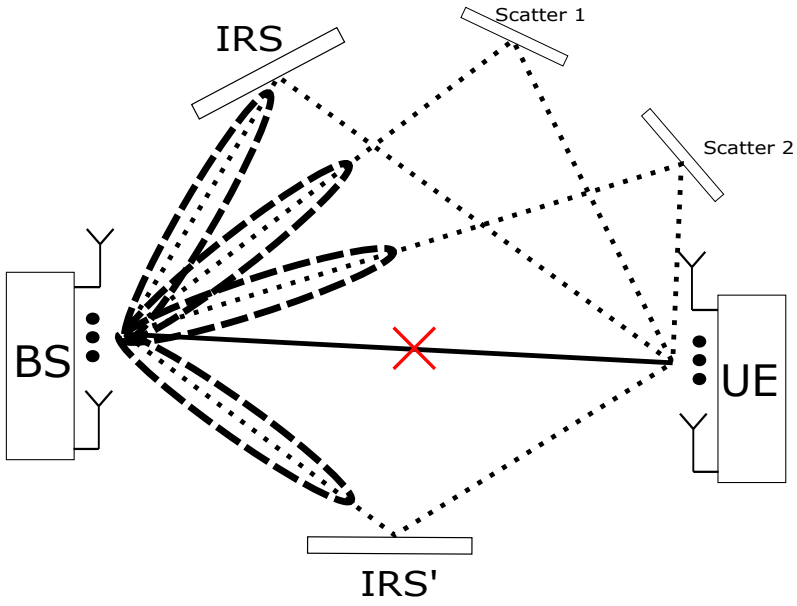


Figure 4.2: The illustration of considered channel estimation scheme.

We conceive a typical scenario to show the basic principle, as shown in Fig. 4.2. We illustrate the channel estimation with the MIMO scheme as the BS and UE can have multiple antennas for high-frequency transmission. There are two IRSs deployed in the environment, and the BS perfectly knows the location of IRSs (mainly about the distance and surface orientation). The locations of scatters, weights of scatters and the UE are unknown for BS, and they are to be estimated. The LoS channel between BS and UE is assumed to be blocked, and only FOR paths are illustrated for simplicity. It is also believed that FOR paths have relatively stronger signal intensity than MOR paths following Friis transmission formula [1] and can be estimated with lower error in the presence of noise.

4.3.1 Locate the UE and Scatters

In Task 1, at least two IRSs are necessary for trilateration [154] to obtain the distance, due to the LoS channel[7] is blocked between BS and UE. Specifically, the following steps can be implemented.

- Step 1, activate beam sweeping at BS such that the beam with maximal power gain at BS can sweep through the entire space.
- Step 2, let UE feedback the beam indices with receive SNR peaks to the BS.

[7]If LoS channel is available, the camera-based method will also work [155] and does not require any help of IRSs.

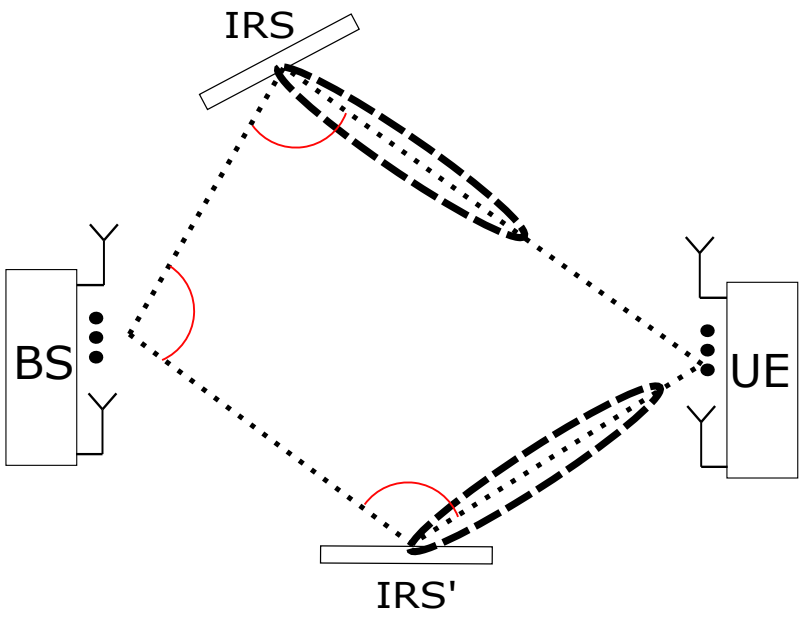


Figure 4.3: Estimating the position of UE by trilateration.

In the steps above, four beam indices are expected to be obtained in this case since there are four objects in total in the environment which cause four SNR peaks by four FOR paths. The dashed line marks out the beams in Fig. 4.2. By corresponding the beam indices to the AOD of BS, the directions of these objects can be obtained. Different algorithms can be employed in this case. It is expected that BS can employ MRC for beam sweeping. IRS can leverage a broad-beam coverage algorithm, which does not require any foreknowledge of channel, such that IRSs can have similar behaviour as the scatters. The UE can use a broadcasting algorithm for uploading feedback.

The expected error induced in this case is due to MOR paths undermining the truth such that the estimation here might not be very accurate. This issue can be mitigated by averaging. In addition, corrections by further feedback in the following steps can also help to reduce the estimation error.

- Step 3, let BS focus the maximal power to a single IRS while activating the beam sweeping on the focused single IRS.
- Step 4, let UE feedback the beam index with the peak value of SNR to the BS such that the AOD on the IRS towards the UE is obtained.
- Step 5, repeat Steps 1 and 2 on another IRS.
- Step 6, by triangle geometry shaped between IRS/IRS', BS, and UE, the distances can be obtained analytically.

As we can see, the steps above first obtain the estimated AOD of two IRSs towards the UE, which forms an intersection point of UE by these two beams, as shown in Fig. 4.3. The quadrilateral is a determined shape made up of edges between UE, IRS, IRS' and BS. The distance between UE and IRS/IRS' can be obtained since the angles marked red are known, and the distances between IRS/IRS' and BS are also known. Reciprocally, as UE is located, the uplink of UE can estimate the parameters above in the same way. The intersection points of scatters can be determined, and thus, the location of scatters can be initially estimated by FOR signal.

4.3.2 Measure the Maximum Reflection Order

Further, in Task 2, we would like to measure the maximum reflection order, Γ. Specifically, as shown in Fig. 4.4, we first use a single IRS for measuring Γ, and another IRS is omitted here. The following steps may do the work.

- Step 1, let BS/UE nullify the transmit/receive power towards/from other objects, except the single IRS for measuring Γ.
- Step 2, let the IRS remain the signal power of FOR path of BS-IRS-UE and the dual reflection paths between IRS and Scatter 1.
- Step 3, estimate Γ and complex gain of Scatter 1 with parameter fitting from the received signal of UE.

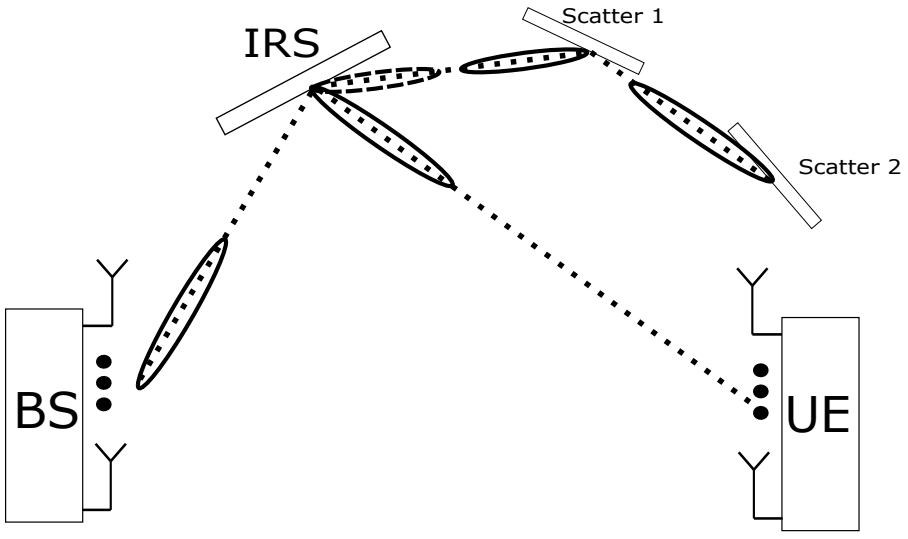


Figure 4.4: Maximum reflection order measuring scheme.

Now, the single IRS is typical scenario as we consider in [143], where the IRS reflects power towards Scatter 1 and UE while receiving power from BS and Scatter 1. In this case, the IRS

implements the MU-LCMV beamforming. BS and UE can employ ZF to nullify the signal transmit/receive from other scatter and other IRSs to filter the desired signal components from the only focused IRS. Ideally, with the cascading model, the UE will have the following receive signal

$$\hat{\mathbf{y}}_{r,UE} = \mathbf{a}_{UE}\mathbf{a}_{BS}^{H}(\underbrace{\mathbf{w}^{H}\mathbf{a}_{C}(\Omega_{out,UE},\Omega_{in,BS})}_{\text{FOR Path Gain}}+\dots$$
$$\underbrace{\Sigma_{v=1}^{V}[\mathbf{w}^{H}\mathbf{a}_{C}(\Omega_{out,UE},\Omega_{in,Sc1})g_{Sc1}]^{v}+\Sigma_{u=1}^{U}[\mathbf{w}^{H}\mathbf{a}_{C}(\Omega_{out,UE},\Omega_{in,Sc1})g_{Sc12}]^{u}}_{\text{MOR Path Gain}})\mathbf{s}_{BS}+\mathbf{n}, \quad (4.19)$$

where $\mathbf{a}_{UE}$ and $\mathbf{a}_{BS}$ are the steering vectors for UE and BS respectively. The $\mathbf{w}$ is the weight vector of IRS. The $\mathbf{s}_{BS}$ is the pilot symbol of BS and $\mathbf{n}$ is the receive noise vector at UE. The $g_{Sc1} = \mathbf{w}_{Sc1}^{H}\mathbf{a}_{C}(\Omega_{out,Sc1},\Omega_{in,Sc1})$ is the complex gain of one time reflection from Scatter 1 where $\mathbf{w}_{Sc1}^{H}$ is the weights of Scatter 1 and $\mathbf{a}_{C}(\Omega_{out,Sc1},\Omega_{in,Sc1})$ is the combining steering vector of Scatter 1 towards the IRS. And the combining vector $\mathbf{a}_{C}(\Omega_{out,Sc1},\Omega_{in,Sc1})$ has same incident and exit directions with respect to the orientation of Scatter 1. Though we can not directly estimate the weights and the combining vector of Scatter 1, the g_{Sc1} can be estimated as a whole. Similarly, we can regard the C-LoS channel of Scatter 1-Scatter 2-Scatter 1 as whole so we denote the cascading complex gain of Scatter 1 and Scatter 2 as g_{Sc12}. Note g_{Sc1} and g_{Sc12} are two complex constants. Also, we should notice the U and V is the cascading times of g_{Sc1} and g_{Sc12}. Notice that the existence of V, or $V = 1$, requires $\Gamma = 3$ while the existence of U requires $\Gamma = 5$. Due to the relationship of a number of graphs' edges, it is clear that $\Gamma = 2V + 1$ and $V = 2U$.

Therefore, we can notice the unknown parameters above are V, g_{Sc1} and g_{Sc12}. Therefore, in Step 1.3 of Task 2, these three parameters can be estimated by solving the following problem, which can be written as

$$\text{(P4)}: \quad \min_{V,g_{Sc1},g_{Sc12}} ||\hat{\mathbf{y}}_{r,UE} - \mathbf{y}_{r,UE}||_2\,, \quad (4.20)$$

where $\mathbf{y}_{r,UE}$ is the actual received signal at the receiver. Multiple samplings are required to estimate these three parameters. Of course, we can estimate parameters in (P4) from the relationship on other IRS and non-IRS scatters as well, using the steps above and taking the maximum value of Γ values as the Γ of the system. Though a similar mindset in the above steps can be used to measure the Γ between two non-IRS scatters, it is hard to have a precise measure because we cannot control them individually. As for the measurement between two arbitrary IRSs, they can be measured with higher accuracy and simpler way. By far, Task 2

is achieved.

4.3.3 Measure Scatters' Weights and Orientation

For Task 3, to have full knowledge of the IRS network channel with non-IRS scatters, the weights and the steering vectors of non-IRS should be respectively estimated as well. Since the cascading complex gain with finite reflection order can also be obtained in Task 2. For Task 3, the following steps can be leveraged.

- Step 1, select one scatter to be estimated and IRS as the main pair.
- Step 2, construct one C-LoS path link via the main pair.
- Step 3, record the specific complex gain of this one C-LoS path.
- Step 4, change another IRS and pair it with the targeted scatter, then repeat Step 1 to 4 above until the sufficiency for estimation.
- Step 5, estimate the weights and orientation of the selected scatter by measured cascaded complex gains of different paths.

For exemplification, we choose IRS and Scatter 1 as the main pair in Step 1.

In Step 2, we can keep the power response of the SOR path of BS-IRS-Scatter1-UE for measurement, as shown on the left-hand side of Fig. 4.5. We can achieve it by nullifying other channel components. The channel components without spatial aliasing can be cancelled based on the knowledge we have obtained so far. In particular, the channel components can be symmetrically cancelled by employing the symmetry of graph topology in the considered scenario, as there are two IRSs and two unknown scatters. As shown in Fig. 4.6, the FOR signal and MOR signal on IRS' and Scatter 2 are able to be cancelled as long as the complex gain from IRS' has the same amplitude but inverse phase. Similarly, other SOR paths that do not require recording can be cancelled with the same condition as long as these paths have symmetry on the graph topology. In this case, except for the selected IRS in the main pair, another one is used to equalize the channel components of Scatter 2 inversely. For channel components with spatial aliasing, where signals are transmitted/received in the same direction, the filtering in the spatial domain does not work. However, signals with spatial aliasing can be nullified by time-domain filters because they must have different path lengths. Thus, the received signal to be recorded in Step 3 is expected to contain the single C-LoS path, BS-IRS-Scatter 1-UE, without any other superposition of signal components. Similarly, another single C-LoS path, BS-Scatter 1-IRS-UE, can be measured in a similar way, as shown on the right-hand side of Fig.4.5.

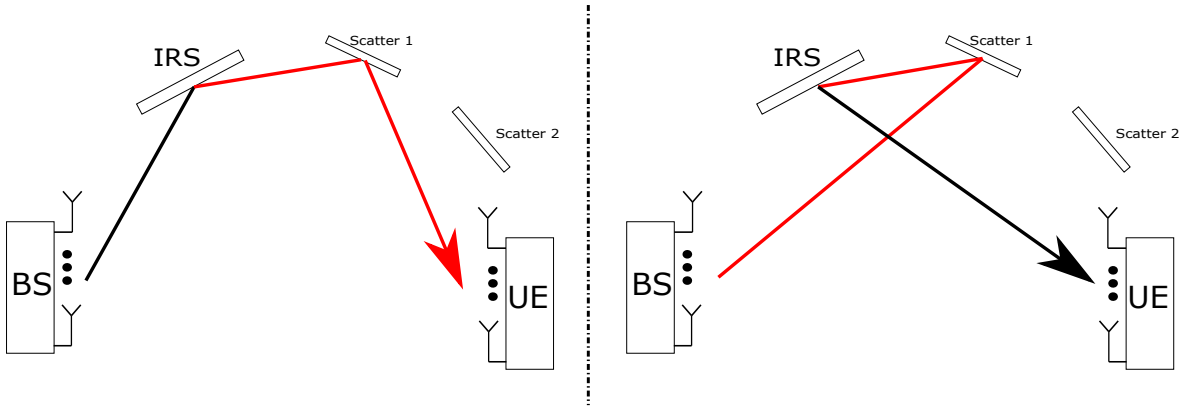


Figure 4.5: Specific C-LoS paths' measurement.

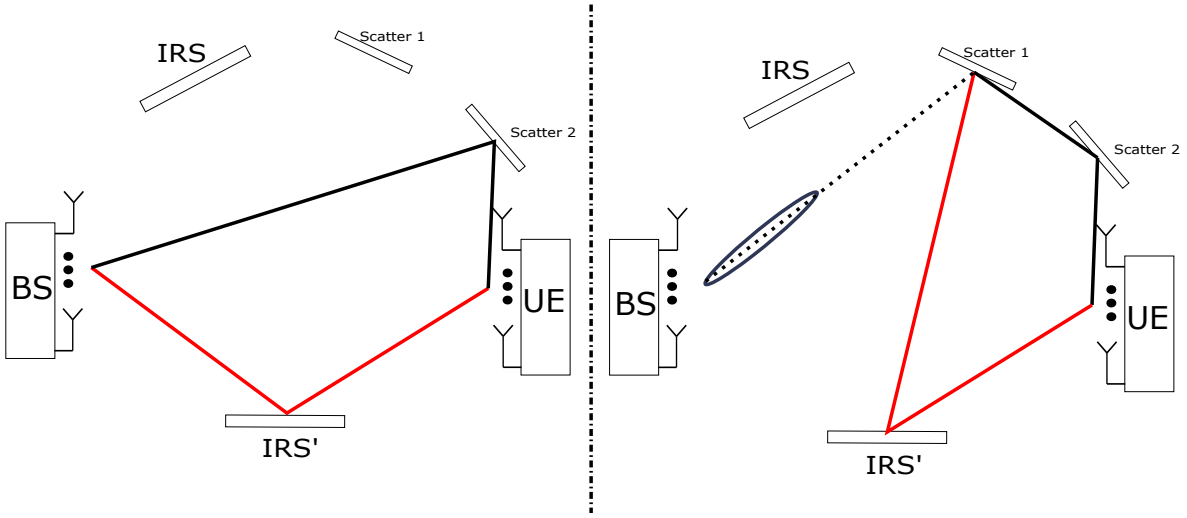


Figure 4.6: Channel components cancellation. Left: Cancellation of two FOR paths. Right: Cancellation of two SOR paths

Following Step 4, another main pair can be selected, and these recordings are used for estimation, which is similarly achieved by solving (P4). By far, Task 3 is completed.

4.3.4 IRS Networks Model Validation and Summary

Task 4 can be finished by leveraging Theorem 1. The recordings obtained in the previous tasks above can be used for the overall model validation.

4.4 Simulation Result

4.4.1 The Single IRS Optimal Capability

In Fig. 4.7, Fig. 4.8 and Fig. 4.9, we consider the MRC solution of beamforming to illustrate the optimal transceiver position of IRS depicted in the Section IV, where MRC is optimal for the 1st fixed pair, which is located at $\mathbf{a}_C(\phi_{in,1}, \phi_{out,1}) = (30°, 135°)$ considering ULA shape's IRS. SNR is assumed to be 10 dB. According to the theorems proposed in this Chapter, we can analytically calculate the optimal available positions for the 2nd pair, where it can harvest maximal power gain from a single IRS with nullified interference from the 1st fixed pair. Analytically, these positions are $(68.53°, 101.95°)$, $(97.70°, 72.97°)$, and $(129.34°, 37.54°)$, respectively when the relative length $L = 2$. Fig. 4.7, Fig. 4.8 and Fig. 4.9 show that theorems in Section IV accurately depict the optimal positions for other pairs. We can also observe that increasing the elements under the fixed-length L will not change the optimal positions. All the optimal positions remain in the same place but only with higher power gain. Fig. 4.9 shows that doubling L also doubles the number of optimal positions, and 8 pairs can be optimally supported in this case.

4.4.2 The Single IRS Interference Suppressing

To validate the interference-free transmission scheme is effective with only a single IRS, we simulate 10000 times the realization of three transmission schemes. The first one is the C-LoS channel with random weights on the single IRS. The second scheme still transmits through the C-LoS channel with random weights, but a 4 by 4 joint decoding matrix at Rxs' side using the ZF algorithm is leveraged as a benchmark (though it may not be practically implemented). The third one transmits through the C-LoS channel with weights obtained by multi-user linearly constrained minimum variance (MU-LCMV) algorithm [143], which can simultaneously support multiple streams by a single IRS. For a specific realization, 4 pairs of transceivers are distributed uniformly around an IRS and transmit normalized power. Since

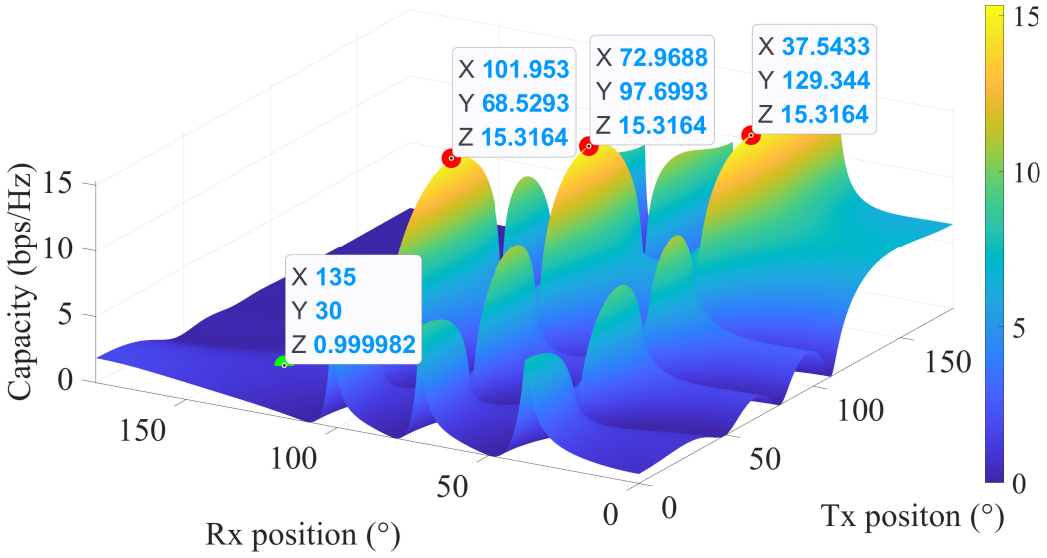


Figure 4.7: $M = 64$, $L = 2$.

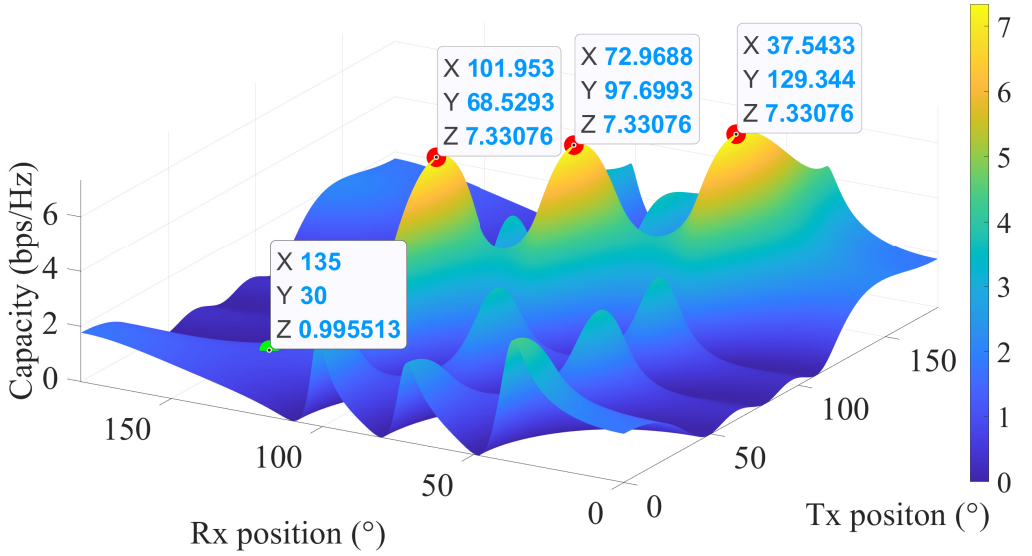


Figure 4.8: $M = 4$, $L = 2$

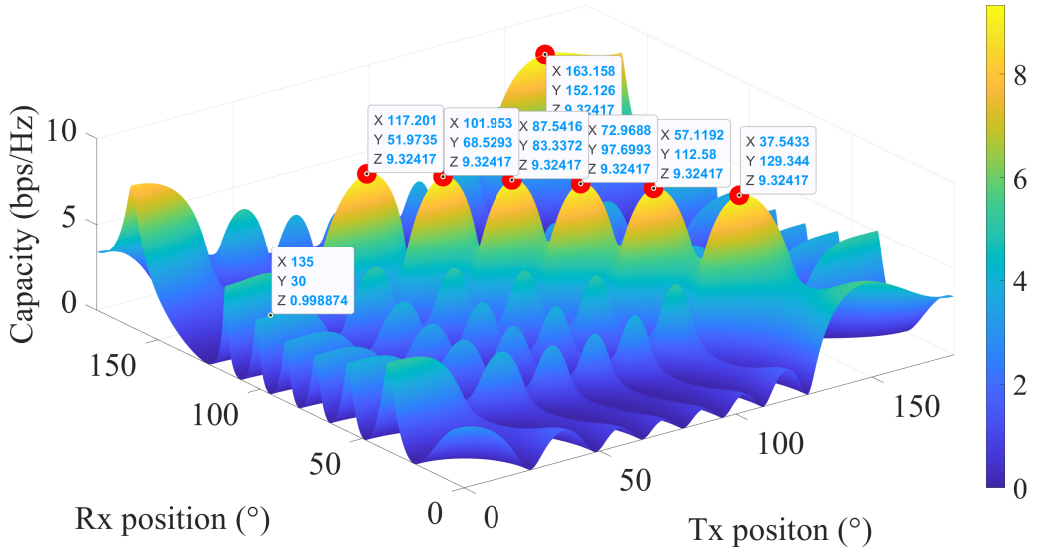


Figure 4.9: $M = 8$, $L = 4$.

ZF at Rxs causes the noise amplification of Rxs but MU-LCMV from IRS does not, the noise is neglected at Rxs for a fair comparison.

The sum rates of these three schemes, changing with the number of elements of a single IRS, are shown in Fig. 4.10. The lower bound of the sum rate for 4 pairs of transceivers is plotted for reference. It can be observed that with a relatively small amount of reflector elements, e.g., $4 < M < 16$, the IRS with the MU-LCMV algorithm is less likely to outperform the traditional MIMO ZF-decoding scheme. At this point, the capability of the IRS is less likely to manage the interference with a limited amount of elements. However, the IRS can suppress the interference effectively at $M = 16$, where the sum rate exhibits a jump. This is critical since the relation of $M = N^2$ in Lemma 3 is exactly satisfied. After that, the sum rate of MU-LCMV on the IRS also reaches a plateau and can have an equivalent performance with the ZF decoding scheme. Nevertheless, since the size of the decoding matrix is fixed, with sufficiently large M on the IRS, the MU-LCMV scheme can finally outperform the benchmark in terms of the power gain from the controlled channel.

4.4.3 The IRS Network Capability

For illustrating the sum rate upper bound of networks, the path loss is assumed to be 0 dB, while the scenario with path loss of -10 dB per edge is also involved. As shown in Fig. 4.11, the sum rate upper bound of LG ($\Gamma = K$) and NG ($\Gamma = 1$) are compared under different SNR, given $M = 6$ and $K = 4, 8$. When path loss is neglected, in a low SNR region, the sum rate upper bound of the LG outperforms that of the NG due to the power gain from each C-LoS

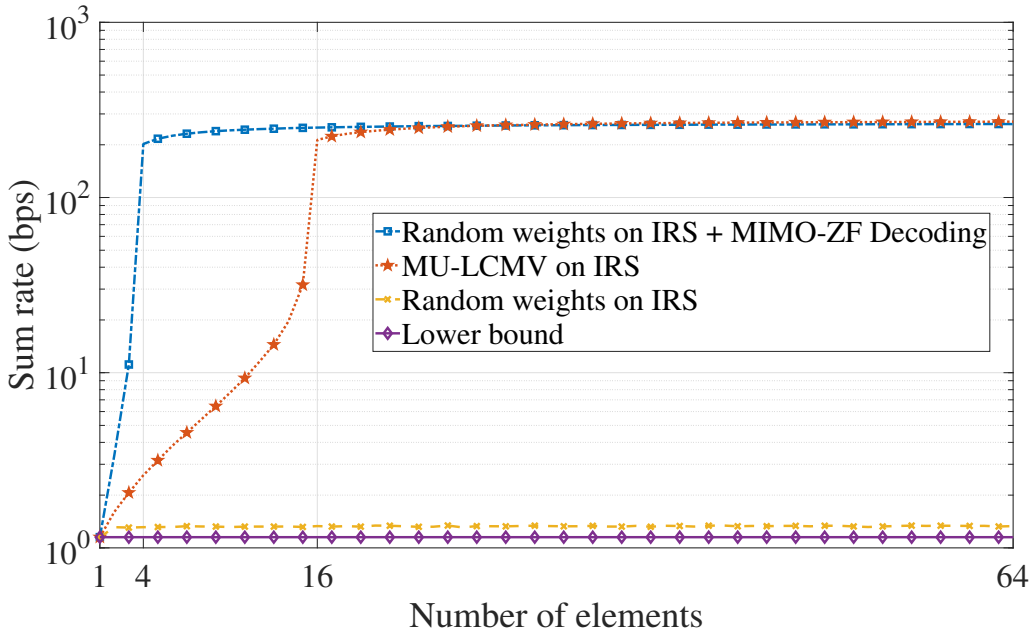


Figure 4.10: sum rates of three different transmission schemes changing with IRS elements number, given $N = 4$.

can be positively cascaded. In contrast, in a high SNR region, the sum rate upper bound of the NG performs better than that of the LG due to larger spatial multiplexing gain being leveraged. However, given an apparent path loss, the NG network achieves a better sum rate since each cascading only causes a larger loss on the cascaded power gain. Thus, the transmission leveraging the most FOR paths in the networks is preferred in this case.

Fig. 4.13 displayed the sum rate upper bound of CG networks with different path lengths and IRS nodes, where $M = 6$. For $K = 4$, we can have $\Gamma = 2, 4$, while for $K = 6$, we have $\Gamma = 2, 4, 6$ such that the graph decomposition into Eulerian paths with equal length is complete. Note that the sum rate upper bound of CG is dominated by the spatial multiplexing gain. Since the CG network can shape more FOR paths with fewer transceiver nodes leveraging edges of the CG, decomposing the CG with the largest path length should result in the least number of MOR paths. Hence, the sum rate upper bound is also related to the value of the maximum order of reflections Γ. In addition, both Fig. 4.11 and Fig. 4.13 verify the sum rate of networks increases substantially with K folds scaling, as we analyzed in Section 4.2. For showing the influence of the number of elements on each IRS, we also plot the corresponding sum rate limit for $M = 128$.

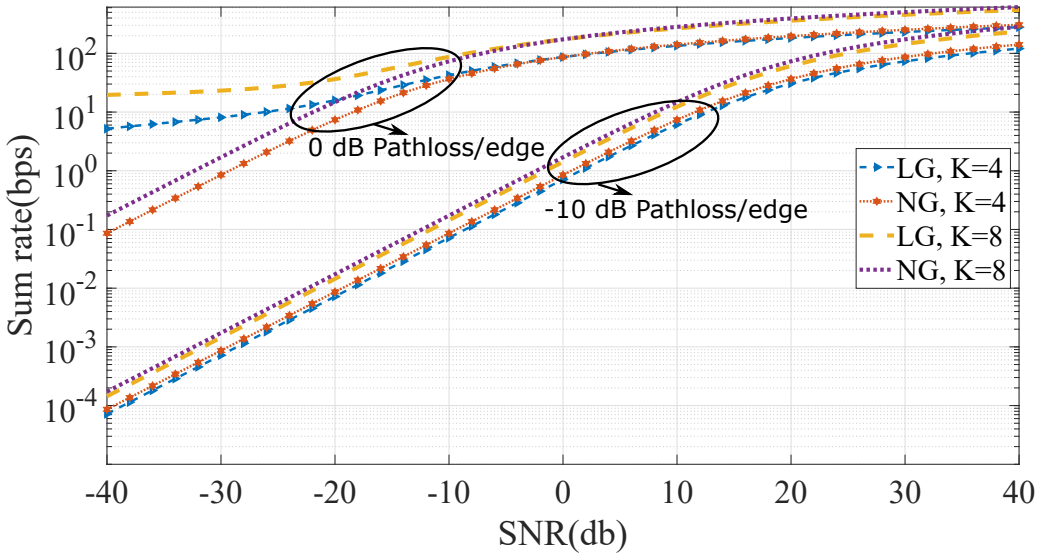


Figure 4.11: The sum rate upper bound of LG and NG network with optimal condition, $M = 6$.

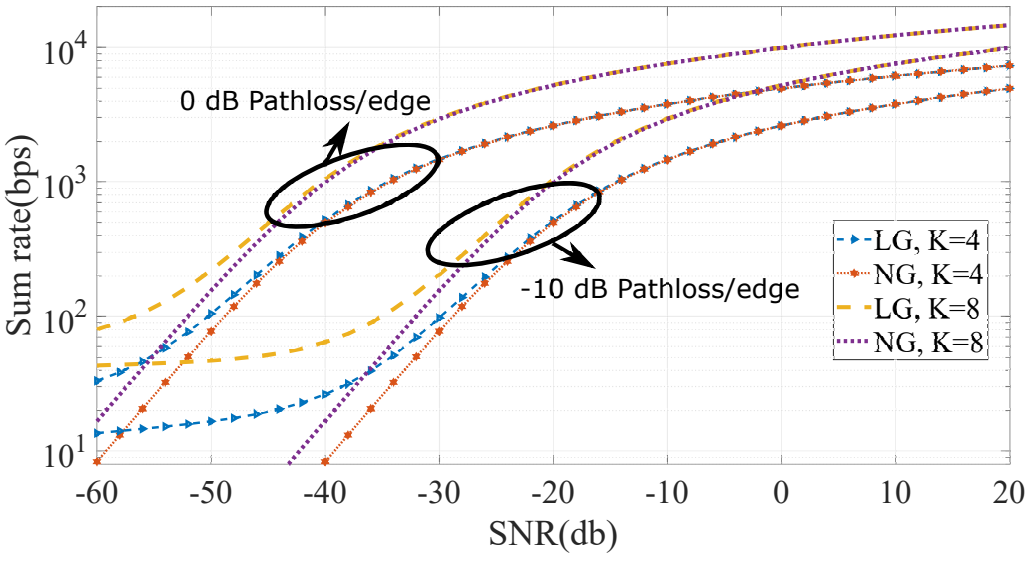


Figure 4.12: The sum rate upper bound of LG and NG network with the optimal condition, $M = 128$.

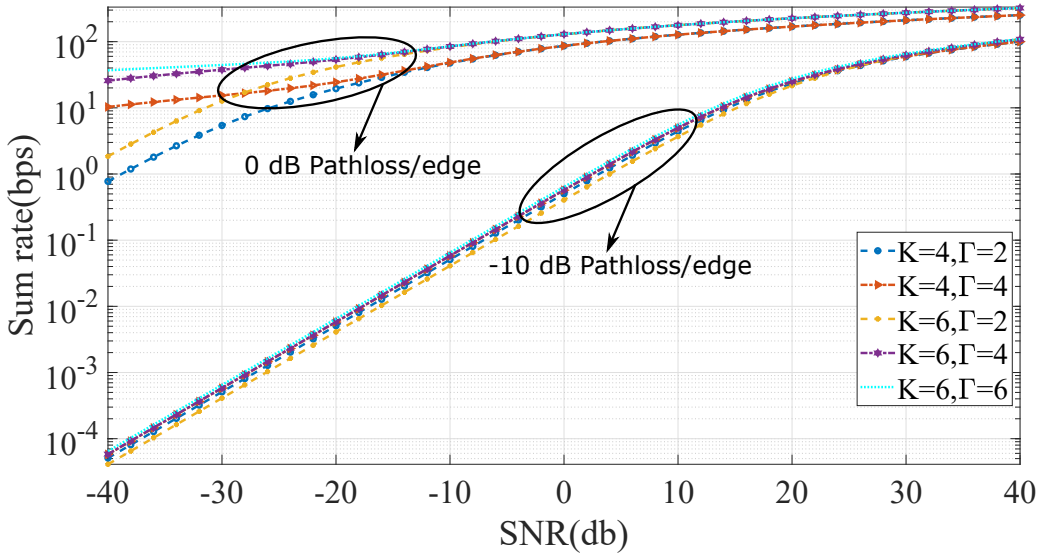


Figure 4.13: The sum rate upper bound of MUMOR CG network with the optimal condition, $M = 6$.

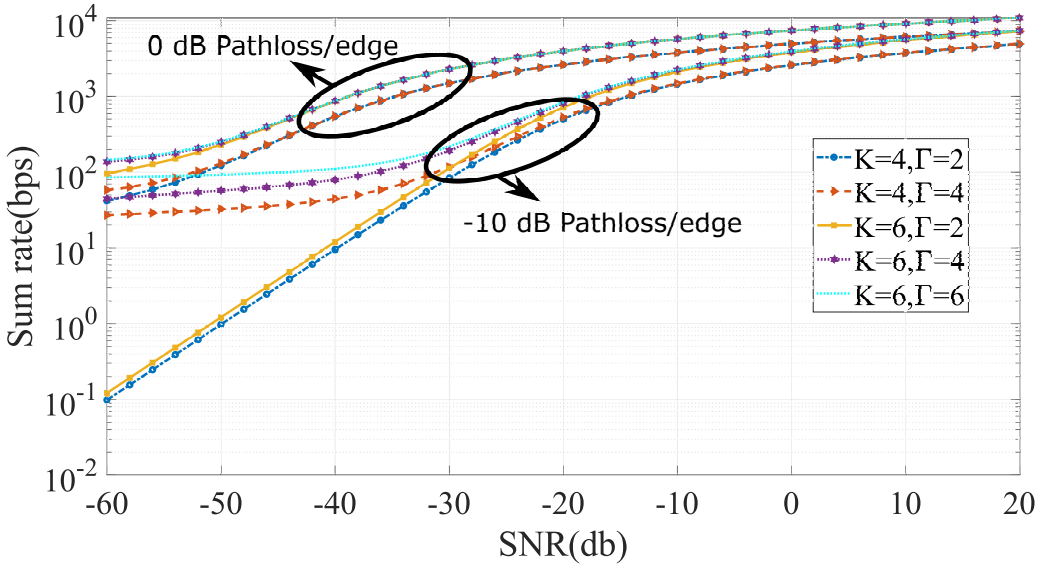


Figure 4.14: The sum rate upper bound of MUMOR CG network with the optimal condition, $M = 128$.

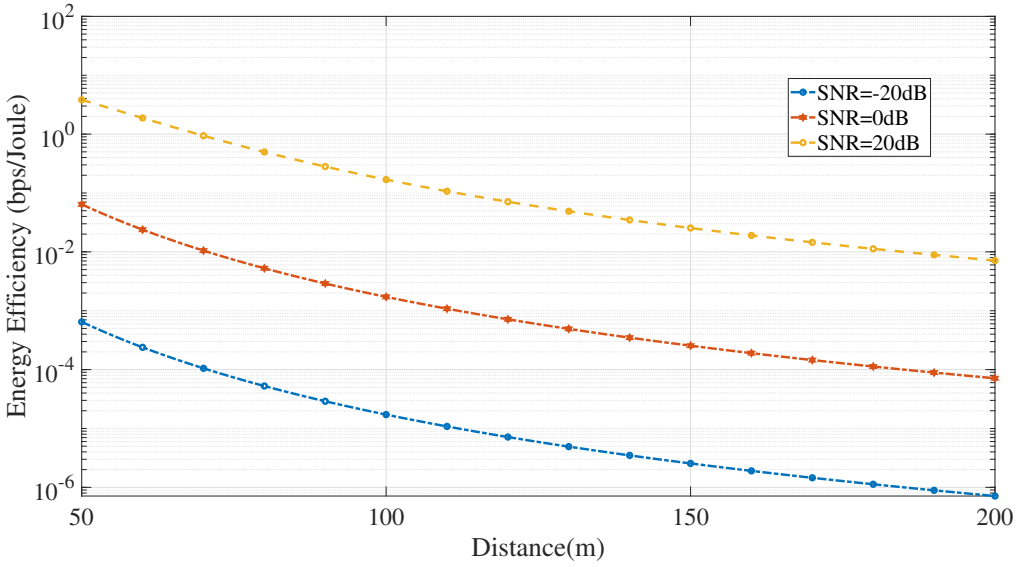


Figure 4.15: The energy efficiency with FSPL, $N = 4, K = 2, M = 128, \Gamma = 2$.

4.4.4 The Energy Efficiency of IRS Network

We illustrate the EE by leveraging MOR under the optimal condition. In particular, 2 IRSs make up the IRS network to support 4 pair transceivers. Each IRS has 128 elements, i.e., $K = 2, M = 128$. Transceivers only have LoS conditions to two IRSs, and they transmit with unified power. In addition, we consider the maximum reflection order to be $\Gamma = 2$. We let the distances of edges in the network by an equal value. We notice that the EE decrease as the FSPL gradually increases with the edge distance, as shown in Fig. 4.15. The receiver with a better design over noise can also have a better performance on the EE.

We also illustrate the single-user multi-path (SUMP) transmission, as shown in Fig. 4.16, where a single pair transceiver can combine multiple C-LoS paths for transmission. This scheme can work more directly to observe the benefits of combining SOR reflections. Note the sum rate limits of SUMP and MUSP can be derived equivalently on typical deployment conditions. An multi-user single-path (MUSP) with $\Gamma = 2$ and $N = 4$ can be derived equivalently with a SUMP, as shown in Fig. 4.17, though the difference in the gain from combining the MOR embodies in the spatial multiplexing gain and the power gain respectively. Nevertheless, the EE and SE can still be simultaneously optimal and our derived theorems are effective. In order to show the EE performance with FSPL and benefits to combining the MOR reflection in this case, we consider one pair transmission with multi C-LoS paths, given 2 IRSs, each IRS has 128 elements, i.e., $N = 1, K = 2, M = 128, \Gamma = 1, 2, 3$. Similarly, we consider all edges have an equal length with a variation. As shown in Fig. 4.18, it can be noticed that the increase in MOR is more significant when IRSs are deployed at a closer

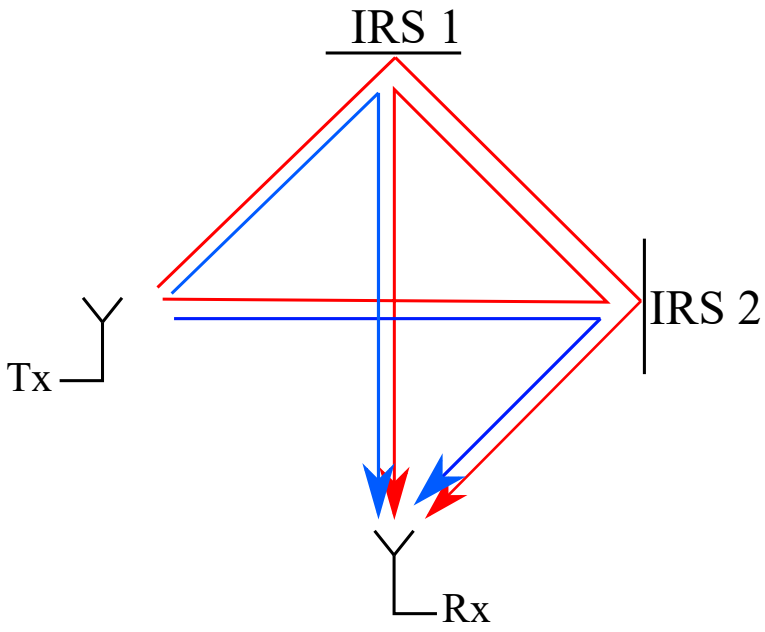


Figure 4.16: The simulation set up for single-user-multi-path transmission.

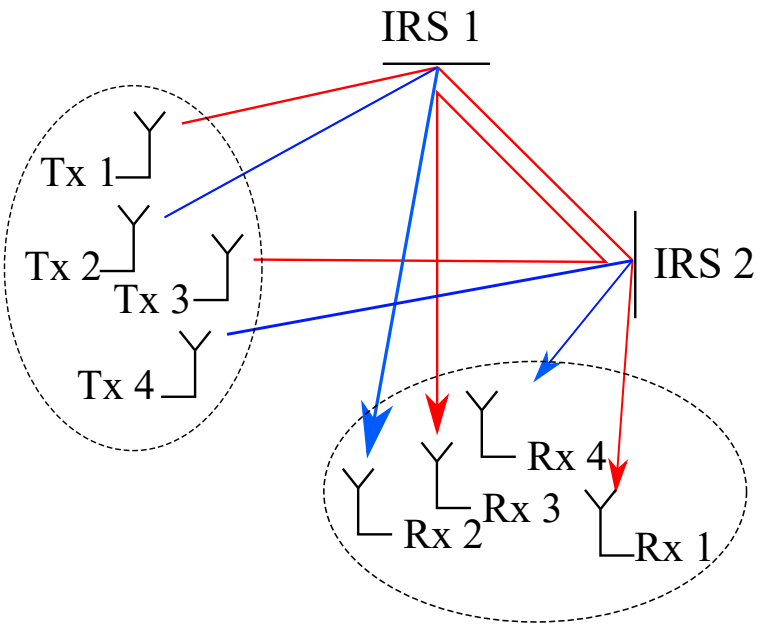


Figure 4.17: The MUSP graph topology, $N = 4, K = 2, \Gamma = 2$.

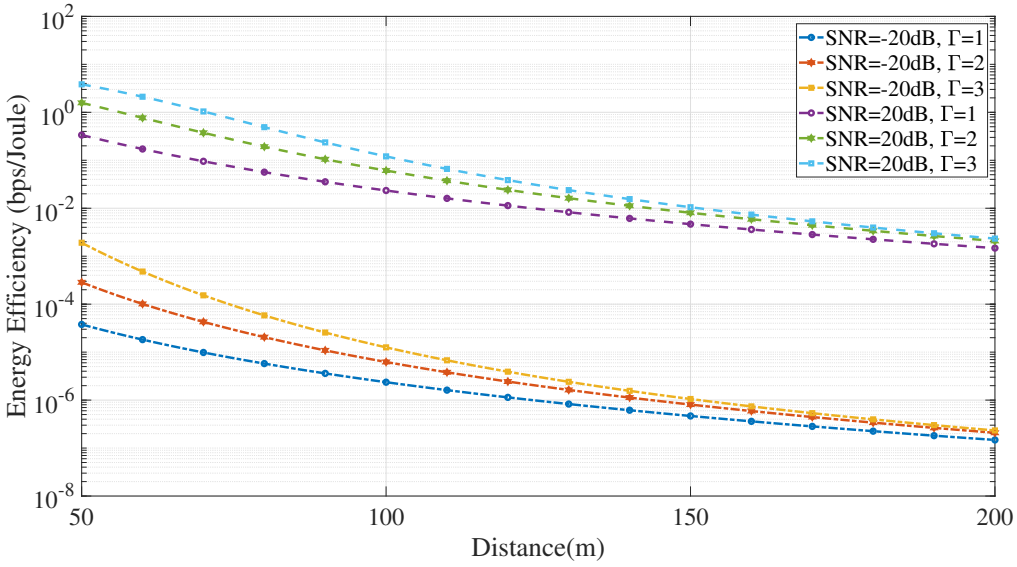


Figure 4.18: The EE of SUMP with different edge distance.

distance. In addition, the enhancement of EE is more limited when SNR is sufficiently large. To conclude, the benefit of leveraging MOR paths is promising, where it is better to combine the power of these paths in a relatively short distance with FSPL for SUMP.

4.5 Summary

In this chapter, we study the MUMOR transmission assisted by the IRS network. Firstly, we analytically establish a complete model of the IRS network by permutationally combining two fundamental models. Secondly, the optimal condition to reach the sum rate upper bound is derived, where the function of optimal positions for the transceivers is written in a closed form. In addition, we found that to sufficiently realize interference-free transmission, $M \geq N^2$ should be satisfied. Lastly, the sum rate upper bound, which the IRS network can provide, is analyzed, where different topologies can enhance the sum rate concerning different numbers of users and SNR. The IRS network topology with the longer cascading path from MOR is shown to perform better on sum rate limit with SNR condition while that of with shorter cascading but with higher spatial multiplexing gain can perform better. The simulation results verify our proposed theorems and indicate a promising K folds scaling from the IRS network.

Chapter 5

IRS Beamforming with Practical Effects

In the previous two Chapters, the IRS and networks are designed and analyzed with conventional and classic assumptions. However, practical performance can be affected by many objective issues. This chapter presents further analysis of IRS with practical effects. In particular, we involve the mutual coupling effect and near-field effect. The mutual coupling effect brings bias to the deployed weights. The near-field effect brings a mismatch to the far-field transmission model.

5.1 Mutual Coupling Effect

5.1.1 Weights bias

Once a mismatch becomes too large by more practical effects, the algorithms proposed in previous Chapters may not guarantee effective solutions for MU transmission. Further analysis of this case is required for specific scenario. Related articles for considering mutual coupling include but are not limited to [125, 156–159]. We notice that the mutual coupling is induced by mutual impedance between each element/antenna circuit. However, an analytical expression for the mutual impedance matrix and the antenna impedance is difficult to obtain [159]. As there are electronic impairments and specific amplitude-phase relationships in practical element circuits, we need to consider the impairment plus mutual coupling into the nonlinear relationship of electronic components from capacitance, impedance, and reflecting coefficient. However, as the mutual coupling does not have an analytical expression, it is rare to incorporate the mutual coupling with the amplitude-phase relationship and impairment together.

Though a physically exact model is yet to be developed for the coupling effect on the IRS, a statistical correlation model is indicated in [125] and we are able to analyze the mutual

coupling effect here. As the authors in [156] indicates that the mutual impedance of element only depends on the geometry, following the modelling method in [125, 157], the mutual coupling effect on the weights vector of IRS is expressed as

$$\mathbf{w}_{\text{coup}} = \mathbf{M}\mathbf{w} \tag{5.1}$$

where $\mathbf{w} = [w_1, \ldots, w_M]^H$ is the raw weights vector (weights before considering mutual coupling effect), $\mathbf{M}$ is the mutual coupling matrix and $\mathbf{w}_{\text{coup}}$ is the weights vector with mutual coupling effect. The term in the i-th row and j-th column of $\mathbf{M}$ is expressed as

$$\mathbf{M}_{ij} = \frac{100^{-d_{i,j}/\lambda}}{\sum_{l=1}^{M} 100^{-d_{i,l}/\lambda}} \ . \tag{5.2}$$

Considering other possible impairments, we have

$$\mathbf{w}_{\text{emp}} = \mathbf{M}\mathbf{w} + \mathbf{n}_g \ , \tag{5.3}$$

where $\mathbf{w}_{\text{emp}}$ is the actual solution of the IRS to employ, $\mathbf{n}_g$ is the noise term due to any other general error, e.g. it can be modelled to equally represent the electronic impairment.

5.1.2 Weights Equalization

To tackle the general variation, which originates from mutual coupling effects and other practical effects of the IRS weights, we can leverage the equalization method. Here, equalization based on MMSE can be applied by choosing $\mathbf{w}$, while obeying the particular amplitude-phase relationship, such that

$$\text{(P5)}: \quad \min_{\mathbf{w}} ||\mathbf{w}_{\text{emp}} - \mathbf{w}_{\text{des}}||_2 \ . \tag{5.4}$$

The aim of solving the problem (P4) is to reduce the gap between desired weights, $\mathbf{w}_{\text{des}}$, and actual weights $\mathbf{w}_{\text{emp}}$ which will be truly employed on the IRS.

5.2 Near-field Effect

Regarding the near-field effect[1], there are four types of transmission scenarios, with respect to the physical transmission distance, for one-pair transceivers. The most common one would be the far-in far-out (FIFO) scheme, where Tx and Rx have equal incident/exit directions/distances to all elements of IRS, as shown in Fig. 5.1(a). In previous chapters, we complete MU beamforming, system limitation and networks capacity analysis of FIFO.

[1]The near-field effect mentioned here should be distinguished from the transient nearfield concept.

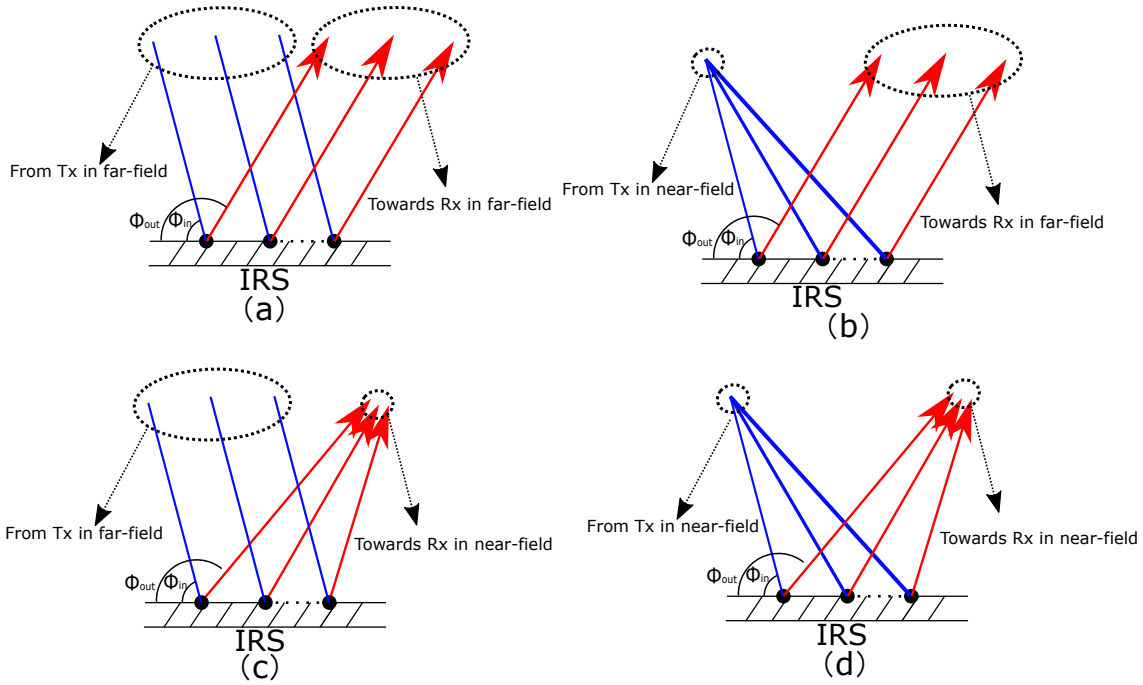


Figure 5.1: Single IRS supporting one pair transceiver. (a) Far-In far-out scheme. (b) Near-in far-out scheme. (c) Far-in near-out scheme. (d). Near-in near-out scheme, where incident/exit propagation paths are marked in blue/red color.

Since a Tx/Rx can also be located in the near-field, Tx/Rx should have different incident/exit directions and distances to elements. Thus, there are other three scenarios, which are near-in far-out (NIFO),far-in near-out (FINO) and near-in near-out (NINO) schemes corresponding to Fig. 5.1(b), (c) and (d) respectively.

5.2.1 Near-field and Far-field Boundary

One common criterion has been mentioned in previous work, which is the Fraunhofer distance, to distinguish the range of near-field and far-field in practical applications. In particular, Fraunhofer distance is obtained in equation (5.5) is based on $\frac{\pi}{8}$ phase error criteria [22], which can be written as

$$D_{Frau} = \frac{2D_{Array}^2}{\lambda}, \tag{5.5}$$

where D_{Array} is the largest dimension from the aperture of the array. As long as the distance between Tx/Rx and the nearest element on the array is greater than D_{Frau}, the plane wave assumption can hold. As indicated by authors in [22], the phase error of $\frac{\pi}{8}$ is not detrimental when the aperture of an antenna is greater than the wavelength of a carrier, thus the far-field assumption is justified. However, once the distance from Tx/Rx to the farthest element is

smaller than D_{Frau}. The near-field effect should be considered, since over close distance will incur significant phase error if one still regards the incident/exit directions as equal value for modelling the wireless channel.

5.2.2 Near-field Channel

To differentiate the near-field IRS model from the far-field one, the far-field steering vector, **a** in equations (2.2) and (2.6), needs to be replaced with the steering vector in the near-field. For simplicity, we let Ω contain one-dimensional information of azimuth angle for an ULA in far-field and contain both azimuth angle and distance for that of in near-field. We denote vector **b** as the near-field steering vector, which can be written as

$$\mathbf{b} = [\frac{1}{D_1}e^{-jkD_1}, \ldots, \frac{1}{D_m}e^{-jkD_m}, \ldots, \frac{1}{D_M}e^{-jkD_M}]^T \,, \tag{5.6}$$

where D_m, $m = 1, 2, ..., M$, is the distance between the point source to the m-th element on the IRS. And

$$D_m = \sqrt{D_1^2 + 2D_1 l \cos\phi_m + l^2}, \tag{5.7}$$

where $l = (m-1)d$, ϕ_m is the incident/exit angle of AOA/AOD from the point source to the m-th element on the IRS. Therefore, we can observe that the distance and angle of directions are coupled in the near-field steering vector[2]. As long as $D_m < D_{fraun}$, by replacing steering matrices $\mathbf{A}_{in}$ and $\mathbf{A}_{out}$, composed by far-field steering vectors with that of the near-field, we can similarly obtain

$$\mathbf{y}_B = \mathbf{B}_{out}^T \mathbf{W} \mathbf{B}_{in} \mathbf{s} + \mathbf{n} \,, \tag{5.8}$$

and

$$y_{B,i} = \mathbf{w}^H \mathbf{B}_{C,i} \mathbf{s} + n_i \,, i = 1, 2, ..., N \;\; , \tag{5.9}$$

where **w** is a column vector whose elements are the main diagonal elements of **W**. Meanwhile, n_i is the noise at Rx_i. The $\mathbf{B}_{C,i}$ is the combined steering matrix for i-th Rx which can be expressed as

$$\mathbf{B}_{C,i} = [\mathbf{b}_C(\phi_{out,i}, \phi_{in,1}), \ldots, \mathbf{b}_C(\phi_{out,i}, \phi_{in,N})] \in \mathbb{C}^{M \times N} \,. \tag{5.10}$$

where

$$\mathbf{b}_C(\phi_{out,v}, \phi_{in,u}) = \mathbf{b}(\phi_{out,v}) \odot \mathbf{b}(\phi_{in,u}) \, u, v = 1, \ldots, N \,. \tag{5.11}$$

[2]In order to obtain the analytical expression of equation (5.6), at least information from one reference point should be obtained. Regarding the 1st element as the reference point, with D_1 and ϕ_1 would be sufficient to express the near-field steering vector. The default reference point is considered to be the 1st element.

5.3 Near-field Beamforming

5.3.1 Steering Error Vector

In order to study the phase error, we derive the phase error between the far-field and near-field steering vector. Following definition in equation (5.7), let $D_{in,p}$ be the distance of Tx to the p-th reference element on the IRS and $D_{in,q}$ is the distance to the q-th element, where $p \neq q$. By deducing the actual phase difference between p-th and q-th element to the first-order phase term of far-field, we have the incident phase error between two elements, which can be written as

$$\epsilon_{in,pq}(\phi_{in,p}, D_{in,p}) = (|D_{in,p} - D_{in,q}| - |p-q|d\cos\phi_{in,p})k, \tag{5.12}$$

thus, we can transfer the near-field steering vector with the far-field vector that

$$\mathbf{a}(\phi_{in}) \odot \mathbf{e}(\phi_{in}, D_{in}) = \mathbf{b}(\phi_{in}, D_{in}), \tag{5.13}$$

where the phase error vector, $\mathbf{e}(\phi_{in}, D_{in})$, is

$$\mathbf{e}(\phi_{in}, D_{in}) = [e^{j\epsilon_{in,11}}, \ldots, e^{j\epsilon_{in,1m}}, \ldots, e^{j\epsilon_{in,1M}}]^T . \tag{5.14}$$

In this case, the phase error difference between the actual phase and the far-field phase can be precisely measured.

5.3.2 Weights Design

As the performance of MU-LCMV has been designed and evaluated to provide a minimal MU interference, we propose to reuse the design for the NINO beamforming. In particular, for the N pair transceiver, we similarly formulate the optimization in near-field

$$\text{(P6)}: \min_{\mathbf{w}} J(\mathbf{w}) = E\{\Sigma_{i=1}^{N} |\hat{y}_{r,i}|^2\} \triangleq \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} + \mathbf{I} N \sigma_n^2 \tag{5.15}$$

$$s.t. \;\; \hat{\mathbf{C}}^H \mathbf{w} = \hat{\mathbf{f}} , \tag{5.16}$$

where σ_n^2 is the noise power, $\hat{\mathbf{R}}$ in equation (5.15) is the covariance matrix and can be written as

$$\hat{\mathbf{R}} = \mathbf{B}_C \mathbf{B}_C^H , \tag{5.17}$$

where $\mathbf{B}_C \in \mathbb{C}^{M\times N^2}$ contains all cascaded quasi far-field vectors of N pair transceiver. And $\hat{\mathbf{C}}$ in (5.16) is the constraint matrix and can be expressed as

$$\hat{\mathbf{C}} = [\hat{\mathbf{C}}_1, \hat{\mathbf{C}}_2] \,, \tag{5.18}$$

$$\hat{\mathbf{C}}_1 = [\mathbf{b}_{C,11}, \ldots, \mathbf{b}_{C,nn}, \ldots \mathbf{b}_{C,NN}] \,, \tag{5.19}$$

$$\hat{\mathbf{C}}_2 = [\ldots, \mathbf{b}_{C,pq}, \ldots] \,, p \neq q, p = 1, \ldots, N \,, q = 1, \ldots, N \quad . \tag{5.20}$$

where $\hat{\mathbf{C}}_1$ is the cascaded quasi matrix of MU transmission and $\hat{\mathbf{C}}_2$ is the cascaded quasi matrix of MU interference. And still, the close-from solution remains as

$$\mathbf{w}_{MU,LCMV} = \hat{\mathbf{R}}^{-1}\hat{\mathbf{C}}(\hat{\mathbf{C}}^H\hat{\mathbf{R}}^{-1}\hat{\mathbf{C}})^{-1}\hat{\mathbf{f}} \,, \tag{5.21}$$

where the computational complexity is mainly subjected to matrices inversion and can be obtained as $O(N^6 + N^4M + M^2N^2 + M^3 + MN^2 + M)$. The constrained response in $\hat{\mathbf{f}}$ may result in the violation of fundamental power constraint for IRS, i.e., $||\mathbf{w}^H\mathbf{w}||_2 \leq M$. This result entails the required signal-to-interference ratio is infeasible for MU given a *passive* IRS, thus it is helpful to judge if an IRS with fixed size and power can satisfy the SINR of MU. Nevertheless, the optimality holds for a normalized solution to satisfy the power constraint with linear scaling.

5.3.3 Redundant Beam Analysis

In a single IRS based MU transmission, the redundant beam is addressed as an intrinsic issue, which causes unavoidable power loss. Specifically, the redundant beam originated from the Bragg condition of IRS in far-field [143]. If the power is maximized for i-th pair, then the Bragg condition for i-th pair is satisfied such that

$$kd(\cos\phi_{out,i} + \cos\phi_{in,i}) = 2\pi n_0, n_0 \in \mathbb{Z}, \tag{5.22}$$

Therefore, as long as equation (5.22) is satisfied, constructive interference occurs in the space domain of another pair. For another j-th Tx at $\phi_{out,j}$, there is one direction $\phi_{out,ij}$ also satisfies this condition, which means the redundant beam at $\phi_{out,ij}$ is caused by the same spatial correlation of the i-th pair's channel.

With near-field effect, the Bragg conditions are expanded into M equations with coupled distance and phase term. For near-field Bragg condition on m-th element, we have

$$D_{in,i,m} + D_{out,i,m} + \frac{2\pi n_m}{k} = D_{in,j,m} + D_{out,ij,m}, n_m \in \mathbb{Z}, \tag{5.23}$$

where $D_{in,i,m}$, $D_{out,i,m}$ and $D_{in,j,m}$ are constants as they are fixed in their locations. The $D_{out,ij,m}$ are to be determined for j-th Tx, which contains the possible direction and distance for the redundant beam to appear since

$$D_{out,ij,m} = \sqrt{D_{out,ij,1}^2 + 2D_{out,ij,1} l \cos\phi_{out,ij,m} + l^2}, \tag{5.24}$$

As the distance should be determined with a reference point, we let $D_{out,ij,1}$ be the reference distance, where the redundant beam can occur. Without loss of generality, we can let $D_{out,ij,1} = D_{out,j,1}$, which means the reference distance for the redundant beam is chosen to be equal to the distance of j-th Rx's to its reference point. With fixed reference distance, the existence of redundant beams is

$$\cos\phi_{out,ij,1} = \cos\phi_{out,ij,2} = \cdots = \cos\phi_{in,out,M}. \tag{5.25}$$

Nevertheless, the equality between M directions above may not strictly hold and the closed-form expression of redundant beams can not be derived as in far-field. With the arbitrary value of the reference distance $D_{out,ij,1}$, the constructive superposition of the reflected spherical wave can reach a peak amplitude response slightly lower than the main beam. As long as the strict equality in equation (5.25) does not hold, the locations of the redundant beam are also subjected to the designed weights vector $\mathbf{w}$, as the phase from the weights can enlarge/reduce the phase difference from equation (5.25) indefinitely after weighting. In this case, we call the resulting redundant beam the quasi-redundant beam, as redundant beams that occur in the near-field is more tricky to be determined than that of the far-field.

Theorem 3. *The quasi-redundant beam of IRS in the near-field occurs as long as*

$$\left|\frac{\theta_p + \theta_q}{k} - D_{out,ij,p} - D_{out,ij,q}\right|_{max} \leq \frac{\pi}{8k}, \tag{5.26}$$

where $p \neq q, p, q = 1, 2, \ldots, M$, and θ_p and θ_q are the designed phase terms from weights such that the weighted phase error between arbitrary two elements, p and q, are constrained below $\frac{\pi}{8}$.

To determine the precise location of the quasi-redundant beam, grid-searching on both distance and direction domain is necessary to solve total M near-field Bragg conditions in equation (5.23). To judge the existence of m-th equation has a solution $\cos\phi_{in,ij,m}$, we can transform equation (5.23) into a series of discriminant. In particular, denote the $D_{ref,ij,m} =$

$D_{in,i,m} + D_{out,i,m} - D_{in,j,m}$, and taking the square from equation (5.23), we have

$$(D_{ref,ij,m} + \frac{2\pi n_m}{k})^2 = D_{out,ij,1}^2 + 2D_{out,ij,1} l \cos\phi_{in,i,m} + l^2. \tag{5.27}$$

Note that, $\cos\phi_{in,i,m} \in [-1, 1]$. By substituting the $\cos\phi_{in,i,m}$ with the range to form the inequality, we have a discriminant on n_m to obtain the feasible range of n_m, we have

$$\begin{cases} n_{m,lower} \geq (\Delta_{-1} - D_{ref,ij,m})\frac{k}{2\pi} \\ n_{m,upper} \leq (\Delta_{+1} - D_{ref,ij,m})\frac{k}{2\pi}, \end{cases} \tag{5.28}$$

where $\Delta_{-1} = |D_{out,j,m} - l|$ and $\Delta_{+1} = D_{out,j,m} + l$. Therefore, the feasible range of n_m is addressed such that we have at least one solution for $\phi_{in,i,m}$. Notice that a relation

$$n_m \geq -\frac{k}{2\pi} D_{ref,ij,m} \tag{5.29}$$

should be remained to keep all distances used in equation (5.23) non-negative. Thus, we can remove the other two imaginary expressions of discriminant due to the square and square root operations. For the feasible range of n_m, it can be understood that only a certain amount of points on the referenced wavefront with the radius $D_{out,j,m}$ to the 1st reference element can be intersected by wavefronts reflected from other $M-1$ elements. Each other $M-1$ element may have multiple indices of wavefronts that can have intersections with referenced wavefront, thus n_m can have multiple values. This introduces a more necessary condition for the existence of redundant beams: given a certain value n_m^* of n_m in its feasible range, the corresponding solutions by solving equation (5.23), $\phi_{in,ij,m}, m = 1, ..., M$ are exactly the same. Thus, with the necessary condition from the discriminants, we can now find M wavefronts constructively interference at the same point on the referenced exit distance $D_{out,j,1}$. To complete the sufficiency, we shall notify all wavefronts that can add constructively at the point of the redundant beam, including but not limited to the case that n_m is an integer. In particular, as long as M wave fonts have the same phase while they intersect on the reference wavefront, the redundant beam can be produced. Thus, with the discriminant relationship of redundant beams, the search process of redundant beams can be implemented more efficiently, as discriminants enable more efficient algorithms than exhaustive algorithms to apply for determining the positions of redundant beams, such as the binary search algorithm.

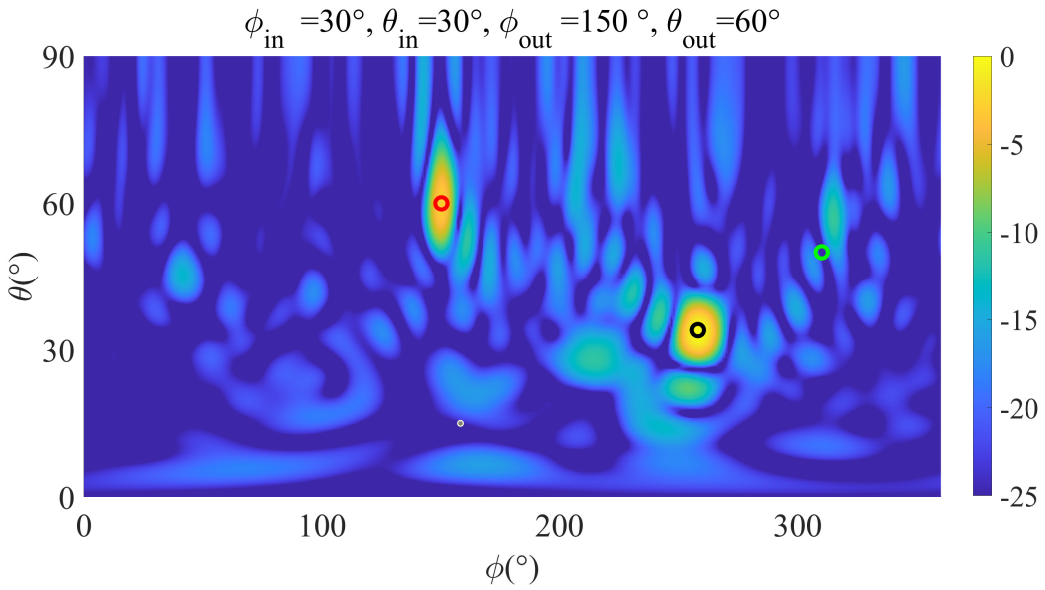


Figure 5.2: 1st pair beampattern with mutual coupling effect, given ACPC of SQP.

5.4 Simulation Results

5.4.1 ACPC with Coupling effect

Considering involving the mutual coupling matrix and using the ACPC solution of SQP as the raw weights vector and following the same simulation setting of Chapter 3.4, the change in amplitude response of beams can be observed from Fig. 5.2 and Fig. 5.4. It is clear that though we aim to let two receivers receive the same amplitude response of signal, however, the unbalanced power allocation is caused where one transceiver pair can have better performance than another one. Thus, the coupling effect in the multi-user transmission affects users' weighting priority. However, as the position of marks we computed are still precisely labelled on the peaks of each beam, the discussion and result on the direction of the redundant beam in Chapter 3.3 still hold correct. It can be found that redundant beams are actually the compromised result of implementing equal power allocation on a single piece IRS for multi-user transmission. The SQP specifically minimizes the interference for multi-user while MRC only maximizes received power for a single user. For multi-user transmission, we also constrain the energy for each pair of users and the cost of doing so is to induce redundant beams such that all users can receive a satisfying signal power. This phenomenon can also be seen as a result that users share the aid of a single IRS fairly. As grating lobes in the antenna array are produced due to the large element spacing, the redundant beams and the grating lobes are incurred in different ways.

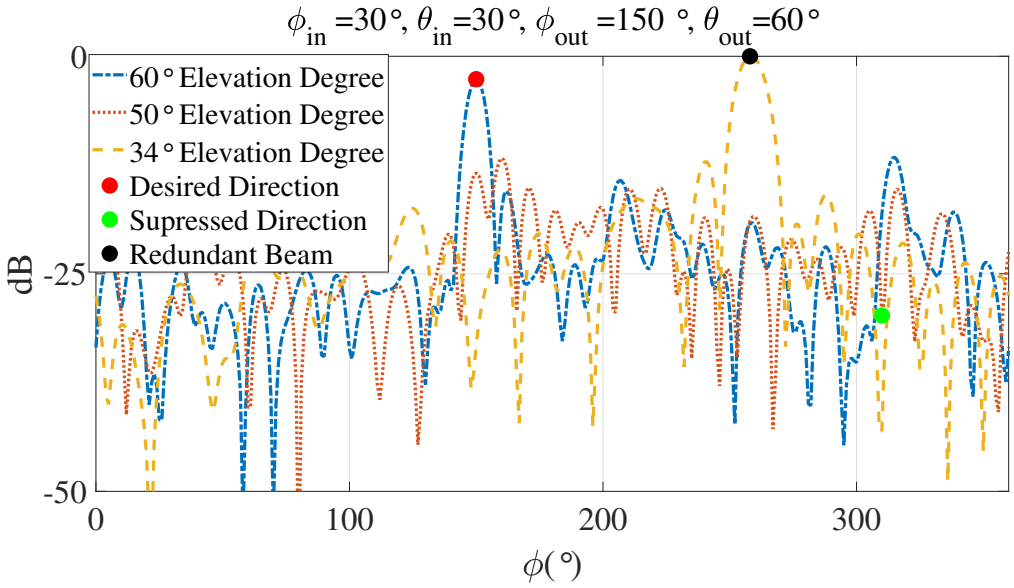


Figure 5.3: 1st pair's cross-section of beampattern.

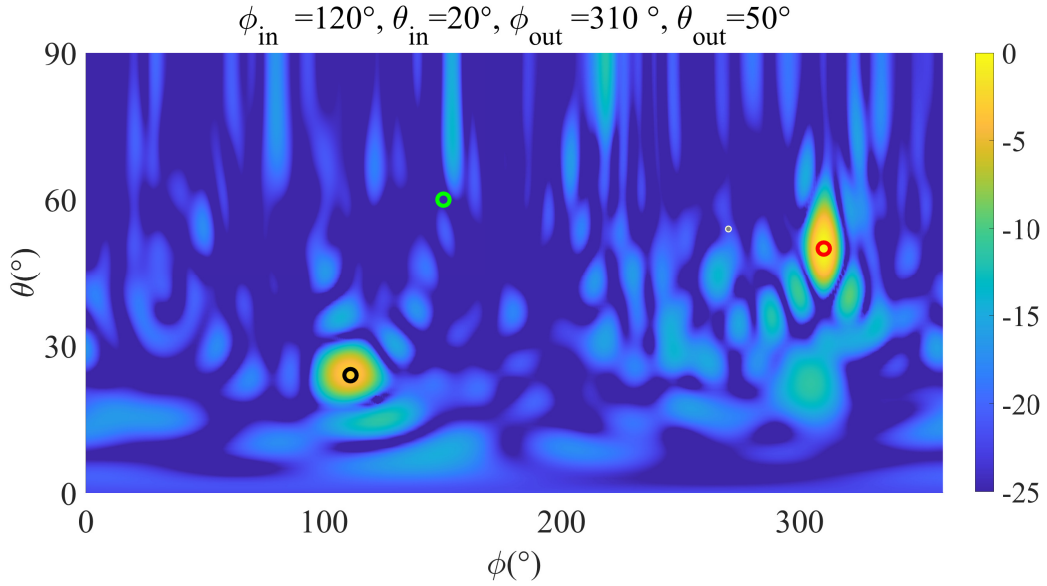


Figure 5.4: 2nd pair beampattern with mutual coupling effect, given ACPC of SQP.

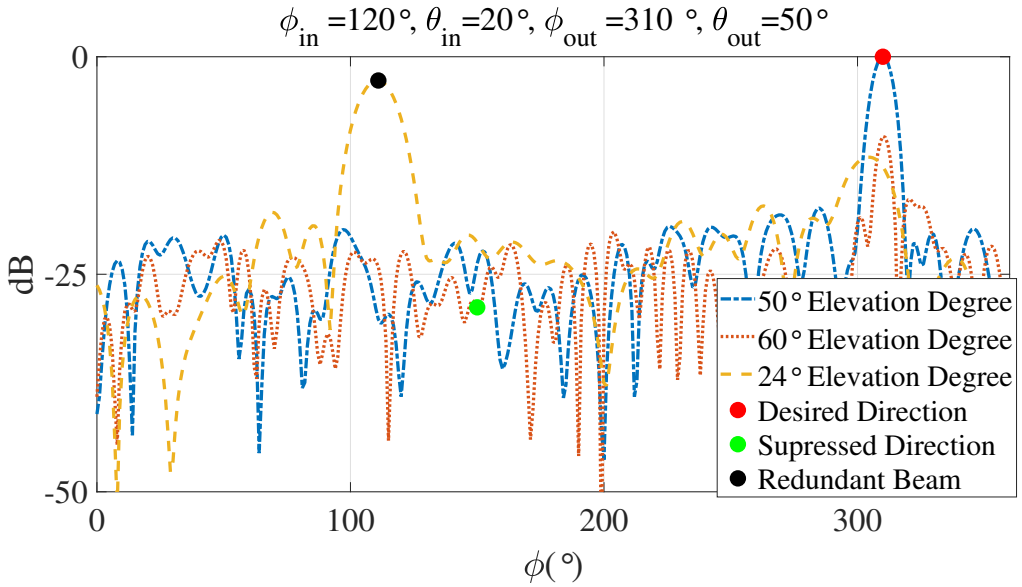


Figure 5.5: 2nd pair's cross-section of beampattern.

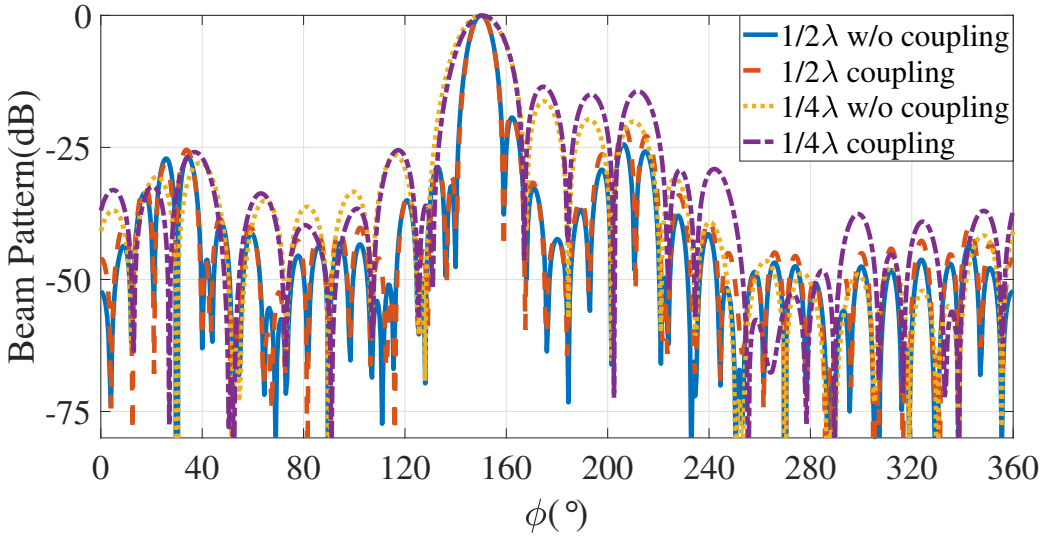


Figure 5.6: Beam pattern with or without mutual coupling effect at $\theta = 60°$

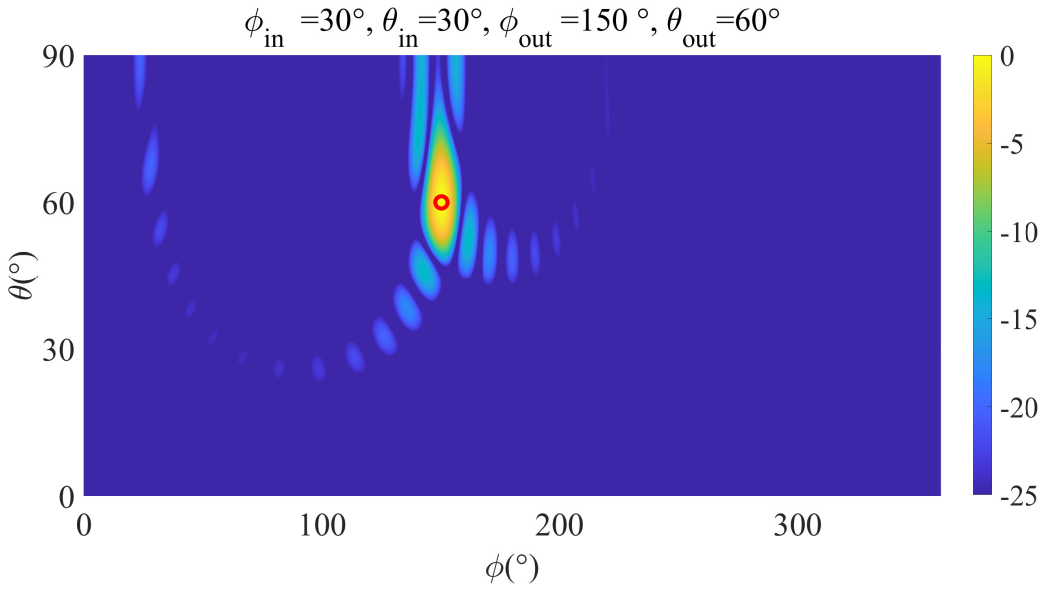


Figure 5.7: Without mutual coupling effect, given $d = \frac{1}{2}\lambda$

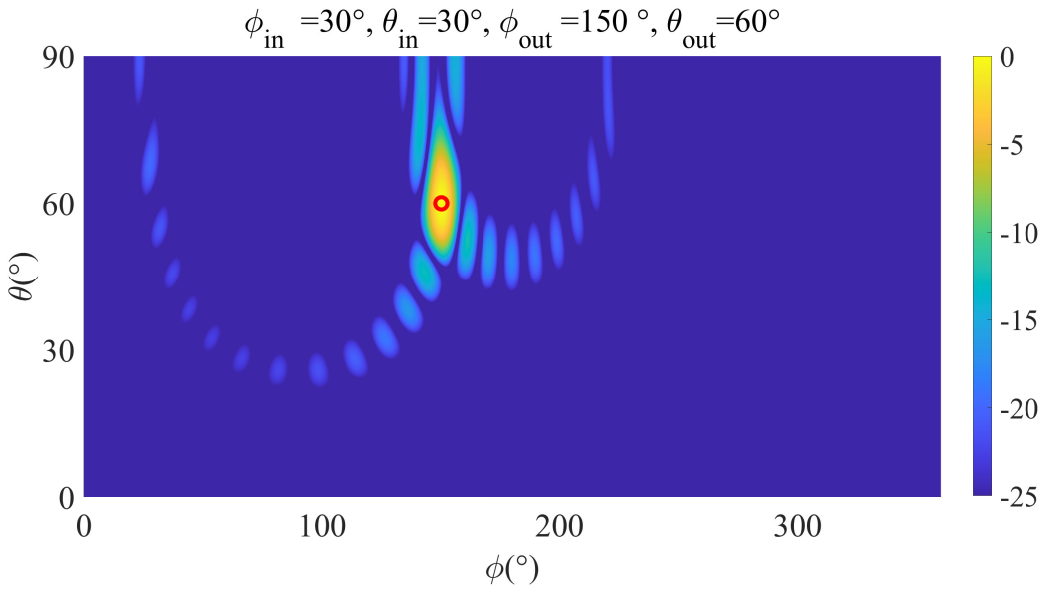


Figure 5.8: With mutual coupling effect, given $d = \frac{1}{2}\lambda$

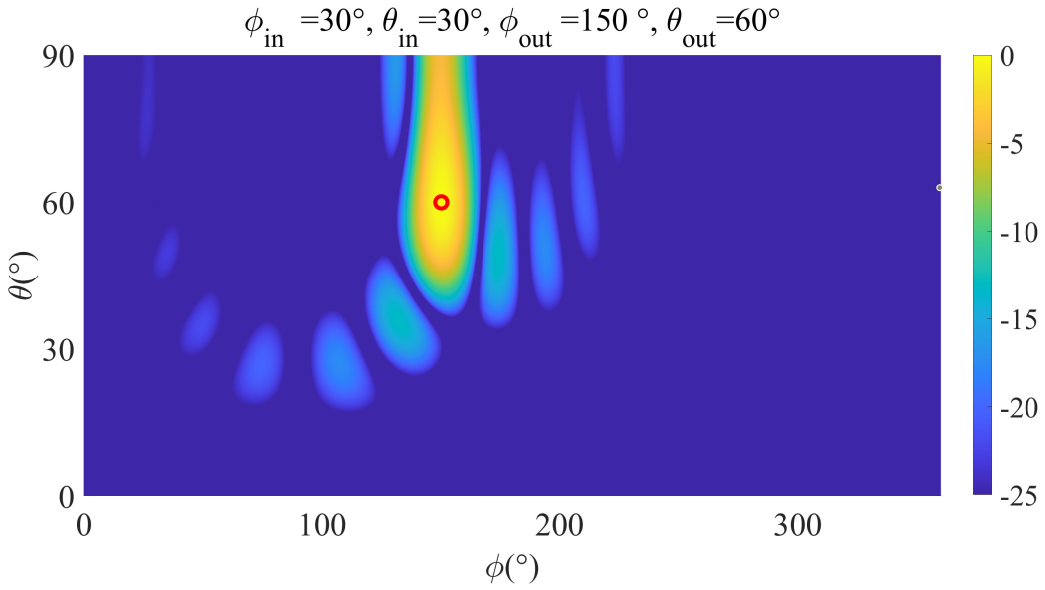


Figure 5.9: Without mutual coupling effect, given $d = \frac{1}{4}\lambda$.

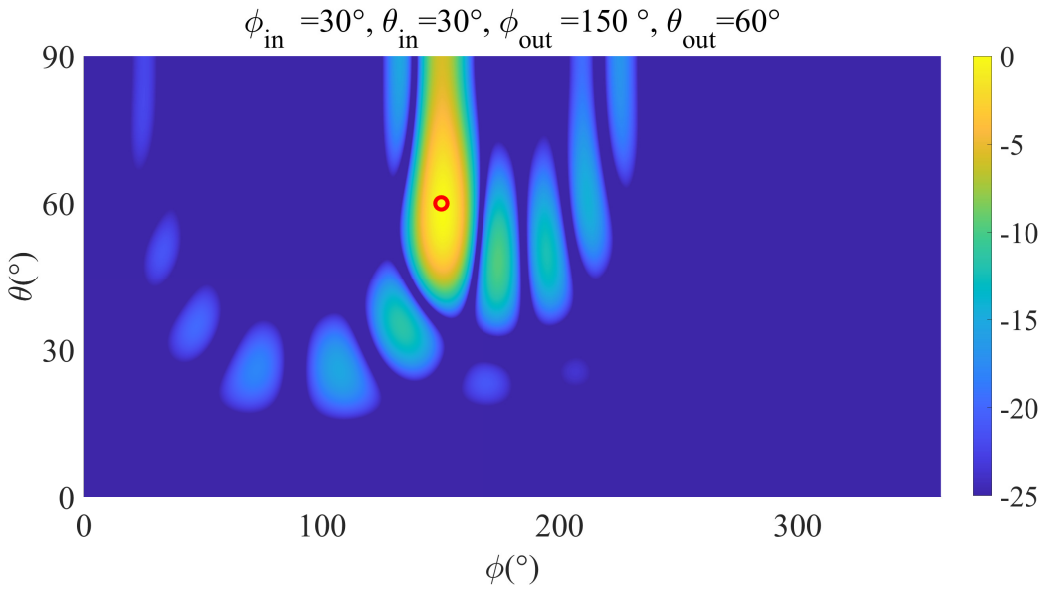


Figure 5.10: With mutual coupling effect, given $d = \frac{1}{4}\lambda$.

5.4.2 Side lobes with Coupling effect

To study the effect of mutual coupling on the side lobes of IRS system specifically, we leverage MRC to obtain ACPC solution to maximize the SNR for only 1st pair user at $\Omega_{in,1} = (30°, 30°), \Omega_{out,1} = (150°, 60°)$. The comparisons are given by considering with (or without) the mutual coupling effects on the weights and different spacing between elements. As shown in Fig. 5.6, the mutual coupling causes higher side lobes than that without a mutual coupling effect. With smaller element spacing, the higher side lobe can be observed on the beam pattern since sparsity in the mutual coupling matrix gets smaller when the element spacing gets smaller. Therefore, it is necessary to equalize the mutual coupling effect if the IRS leverage the small element spacing. But for spacing with half-wavelength, the coupling effect does not affect the directivity of beams.

Fig. 5.7, Fig. 5.8, Fig. 5.9 and Fig. 5.10 also present the coupling effect on the side lobes in 3D pattern. More apparent side lobes can be observed if the mutual coupling is considered. In particular, Fig. 5.8 and Fig. 5.10 shows the 3D beam pattern of URA IRS with mutual coupling effect while Fig. 5.7 and Fig. 5.9 does not involve the mutual coupling. The amplitude of side lobes are increased by coupling effect in Fig. 5.8 and Fig. 5.10, thus less number of side lobes are observed in Fig. 5.7 and Fig. 5.9.

5.4.3 Near-field MU Beamforming with AXPC solution

To study the MU transmission with near-field effect, an ULA IRS of 512 elements with the half-wavelength element spacing is considered to support three transceiver pairs simultaneously. Since the Fraunhofer distance and Fresnel distance are both dependent on frequency and aperture size, 30 GHz is considered as the frequency of mmWave[3]. In this case, the Fraunhofer distance is equal to around 1.3 km, and the Fresnel distance is equal to around 25 m. Three pairs of transceivers are considered and their locations can be represented by Ω notation as $\Omega_{in,1} = (70°, 14.7\text{m}), \Omega_{out,1} = (60°, 14.7\text{m}), \Omega_{in,2} = (80°, 14.7m), \Omega_{out,2} = (110°, 44.7\text{m})$ and $\Omega_{in,3} = (90°, 14.7\text{m}), \Omega_{out,3} = (130°, 29.7\text{m})$, where only azimuth directions and reference distances are included in Ω. These receivers are set to have equal power response by constraints, but the nullification for MU interference is omitted for reducing the complexity of computation and simplicity to obtain the AXPC solution. To show the beam behaviour clearly, the distance range for plotting is considered from $0m$ to $55m$. In addition, the Fresnel distance is plotted as the black dash line in the middle of Fig. 5.11, Fig. 5.12 and Fig. 5.13.

[3]Nevertheless, the result can be general for other configurations as long as the ratio of frequency and aperture size keep a constant.

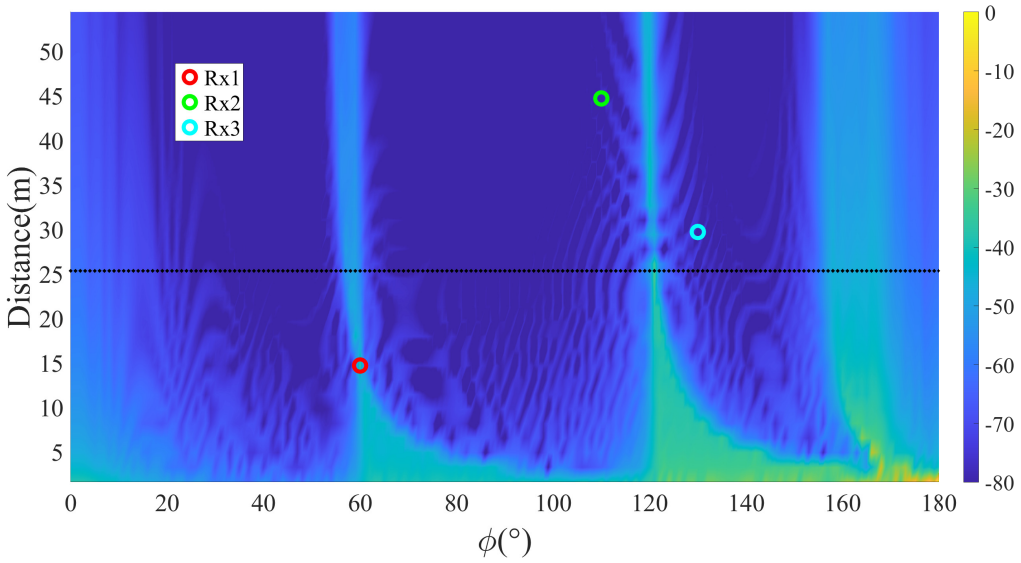


Figure 5.11: Beampattern of Tx1.

In Fig. 5.11, Fig. 5.12 and Fig. 5.13, it can be observed that each Rx has received the main beam from the corresponding Tx. With distance variations on the wireless channel in the near-field, the focus point of beams at different distances introduces the beam scattering at a different level. For instance, as shown in Fig. 5.11, a clear focal point is addressed, marked by the red circle, on the main beam to the location of Rx1. However, the main beam is scattering in other positions except for the focus point. The main beam becomes scattering after penetrating the position of Rx1 while the main beam is also scattering before reaching the distance where Rx1 is located. As shown in Fig. 5.12 and Fig. 5.13, for the main beams where the focus points are farther than the Fresnel distance, the beam scattering occurs in the distance shorter than the distance of the focal point.

Since three pairs of transceivers are leveraged, two redundant beams can be observed in Fig. 5.11, Fig. 5.12 and Fig. 5.13. For the redundant beams, their locations are mainly subjected to the focus points of MU. The locations of redundant beams also shifted with different positions of transceivers. As shown in Fig. 5.11, the redundant beams still exist but they avoid the positions of Rx2 and Rx3, which are marked by the green and cyan circles. Due to the focus points of Rx2 and Rx3 for Tx1 being a shifted position, the beam scattering and focusing of redundant beams also shift in this case. Due to these shifted focus points can not being derived in close form, it is not convenient to determine the specific positions as did in far-field conditions with analytical results. Numerical methods like grid-searching can apply to address the positions of redundant beams.

Fig. 5.14 shows the cross-section of beampattern at different reference distances. Specifically, for each Rx, the beampattern with respect to the reference distance corresponding to

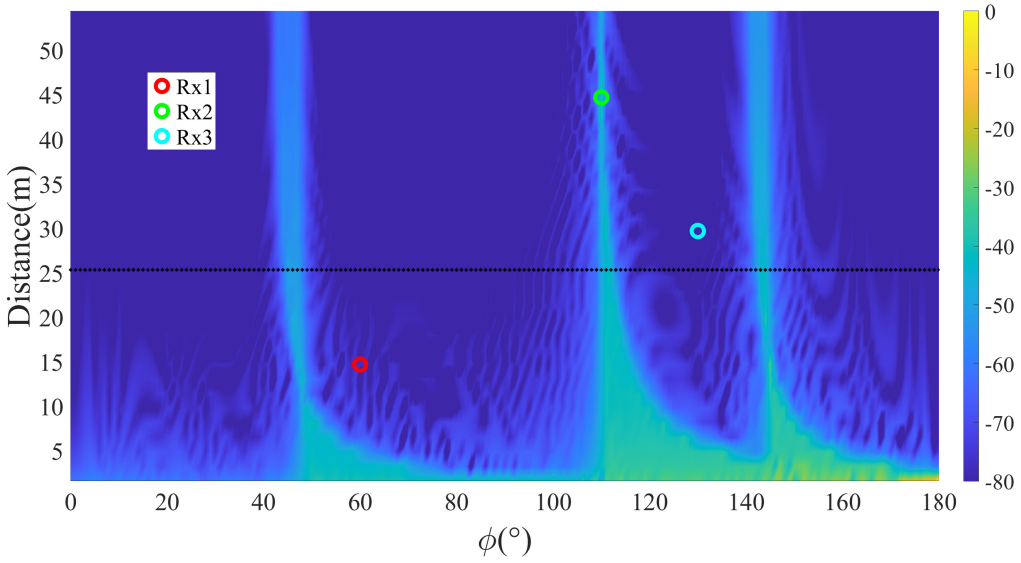


Figure 5.12: Beampattern of Tx2.

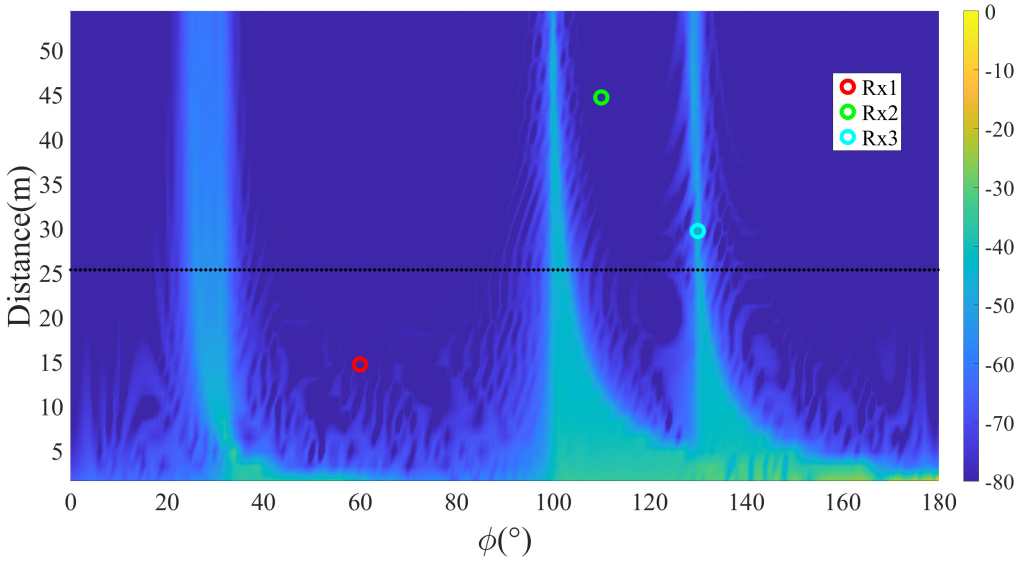


Figure 5.13: Beampattern of Tx3.

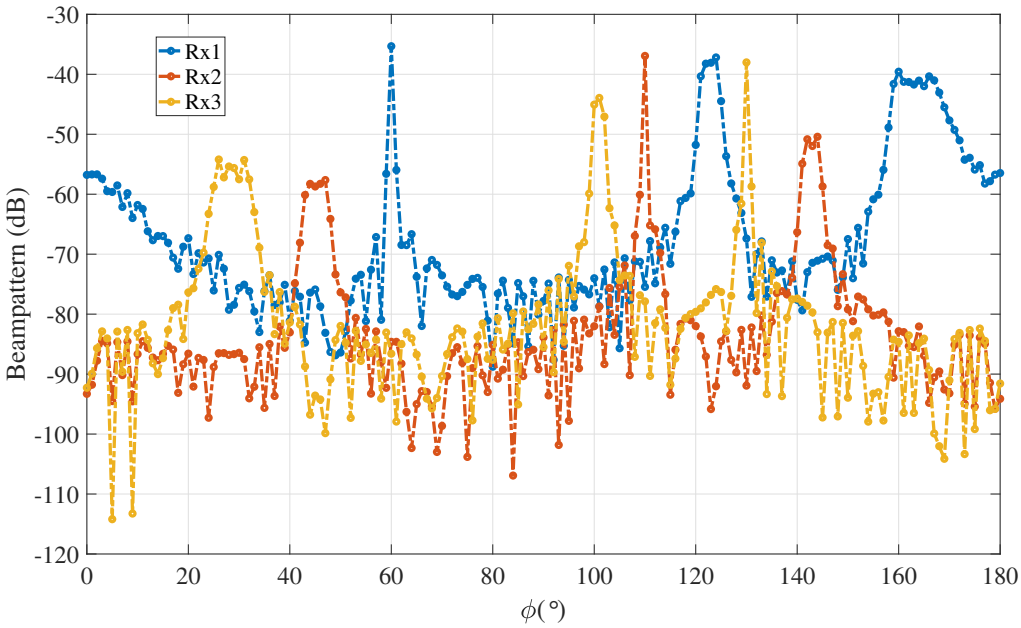


Figure 5.14: Cross section at reference distances of different receivers.

such an Rx is plotted. The blue, red, and yellow dot-dash lines represent the beampattern of Rx1, Rx2 and Rx3 respectively. It can be observed that with FSPL, the received power response is around -36 dB and there is approximately the power response of beam scattering in other distances. Though nullification of interference is not leveraged in AXPC solution of MU-LCMV, due to the formulation of power minimization, the MU interference remained at a maximum -67 dB. Due to the focus point can shift with different observation points, i.e., location of Tx1, Tx2 and Tx3, the redundant beam can have different amplitude and bandwidth at the different reference distances.

5.4.4 Error of Redundant Beams in Near-field

Fig. 5.15 shows the three different methods for calculating the directions of redundant beams at different distances. The locations of transceivers are the same as in the previous near-field simulation. Here, the relative distance range is sampled equally from half of the Fraunhofer distance with 1000 dots. The analysis is made for the redundant beam caused by the 1^{st} pair to the 2^{nd} pair transceiver, where the redundant beam is roughly at $47°$ and can be seen in Fig. 5.12.

The first method directly leverages the formula derived in Chapter 3.3 for far-field conditions. The second method is based on numerical computation with Fresnel approximation. Specifically, the direction is computed by solving the M equations from equation (5.23) with Fresnel approximation and taking the average value from the M solved directions. The third

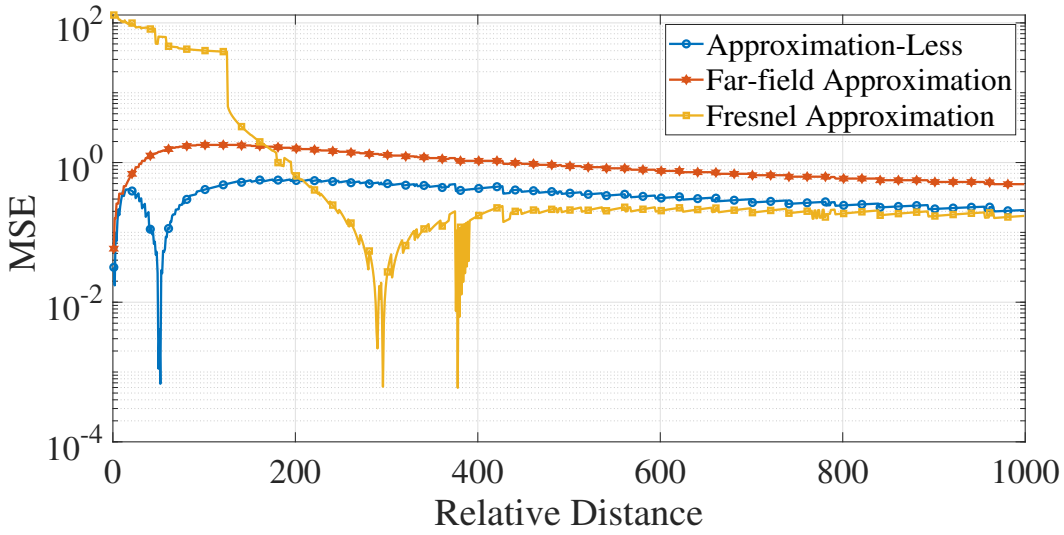


Figure 5.15: A single IRS model for MU transmission. Different colors mark the signal transmission path from different Txs.

method keeps the same computation process as the second one but without any approximation. The true value of the redundant beam is obtained by the grid searching method, where the error (in a unit of degree) is obtained lower than 10^{-4}. As can be seen from the result, the far-field approximation method becomes precise gradually as the increment relative distance. This is due to the far-field approximation becoming precise with increasing distance. The Fresnel approximation fails to obtain the direction of the redundant beam but becomes precise farther than the 200 units relative distance. This is due to the Fresnel approximation also failing to hold and the non-linear characteristics of the Fresnel approximation also amplify the error. In this case, the approximation-free analysis method can obtain a more precise result.

5.5 Summary

In this chapter, we evaluate the mutual coupling effect and near-field effect on the transmission performance and redundant beams of IRS. The mutual coupling effect is addressed to be an intrinsic issue that can happen to different IRS hardware structures, but it may not be critical to affecting the MU transmission and the locations of redundant beams with proper element spacing. Nevertheless, the mutual coupling effect should become significant and need to be equalized once the element spacing is smaller than the half wavelength. The near-field effect becomes significant with a close transmission distance and over the large size of the IRS. In this case, the distance-induced phase variation will be apparent and the transmission design on the weights should be specifically designed, given the different locations

of transceivers. Though the locations of redundant beams in the near-field do not have a closed-form expression, they can be determined through grid-search methods.

Chapter 6

Conclusion and Future Work

The IRS can be a potential and promising enabler in 5G and beyond generations of wireless communication to support programmable channels. The concept of the programmable channel can generate an all-around disruptive change to wireless communications ranging from transceiver design, system and architecture design, network deployment, optimization and operation. This thesis comprehensively studied the MU beamforming and its limitation, deployment of IRS networks with upper bound of EE and SE and made a specific analysis of mutual coupling effect and near-field effect. This chapter concludes this thesis's main work and contributions, together with a brief discussion on future works about the IRS.

6.1 Conclusion

In Chapter 1, together with the introduction of some critical wireless techniques that have been used, an overview of IRS, from the basic concept to its detailed hardware realization and application scenarios, are presented. Specifically, the necessary background about physical wireless channel, MIMO and antenna array theorems are introduced. The current literature indicated the promising study area in MU transmission with single IRS and IRS networks.

A comprehensive model of IRS is presented in Chapter 2, in terms of the basic transmission model of single IRS to multiple IRS composed networks. For the networks model, we further address the multi-order-reflection affected by the dual reflection between two IRSs as well. The proposed IRS networks model is able to completely express transmission environments with an arbitrary number of reflectors in arbitrary sizes and weights and an arbitrary order of signal reflections.

In Chapter 3, an idea of achieving MU transmission using a programmable channel based on single IRS is proposed. The single IRS can use a single set of weights to support interference-free multiple data stream transmission. A novel MU-LCMV beamformer

is proposed to solve the optimization problem to obtain the optimal AUPC weight vector in close form with minimal interference. The ACPC and ACPQ are proposed for IRS implementation with less control ability on phase and amplitude coding by constraining the amplitude and phase of the weight vector. Limitations on HPBW of IRS with MRC beamformer are obtained. It is addressed that the beamwidth employing MRC can match with a traditional array system. A critical observation of this chapter is that redundant beams exist in the MU transmission on single IRS, which brings significant power loss but also can be used for multi-cast transmission. Under different IRS configurations with different element spacings, the redundant beams are mathematically proved and analyzed. The directions of redundant beams are derived in close form as well, which can provide important guidance when designing with multi-cell aided communication to avoid strong interference on the cell edge. Results verified the finding and algorithms and suggested MU to be more spatially distributed for satisfactory system performance.

Towards the communication aided by the IRS networks, the MUMOR transmission based on the IRS network is studied in Chapter 4. Based on the complete model of the IRS network, where networks permutationally combine single IRS and IRS to IRS transmission model, the optimal condition to reach the sum rate upper bound is derived, where the function of optimal positions for the transceivers is written in a closed-form. In addition, An observation is obtained that to sufficiently realize interference-free transmission, $M \geq N^2$ should be satisfied. It is also addressed that different topologies can enhance the sum rate concerning different numbers of transceivers and SNR. Moreover, leveraging MOR of IRS networks with a relatively close range can be more efficient in improving the EE. The simulation results verify our proposed theorems, indicate a promising K folds scaling from the IRS network and illustrate the promising improvement on the EE and SE by leveraging the MOR effect of IRS networks.

Involving the mutual coupling effect and near-field effect, the MU beamforming and the corresponding behaviour of redundant beam are studied in Chapter 5. As a result. mutual coupling effect of IRS indicates increasing on side lobes level with smaller element spacing, where sidelobes can be more significant when element spacing becomes smaller. Except for more power dissipation in a trivial direction, the mutual coupling may not be critical to affecting the MU transmission and the redundant beams in directivity. For the near-field effect, it becomes significant with close transmission distance and over the large size of IRS. The MU beamforming for transceivers located at different distances is achieved with MU-LCMV. It is also addressed that locations of redundant beams, in this case, become tricky to determine explicitly and should be determined by grid-search methods. The discriminant for the existence of a redundant beam, in this case, is derived and can be used to reduce the

complexity of grid search. Simulations are provided to show the performance of MU design, and numerical result is given to validate the analysis of the redundant beam.

6.2 Future Work

6.2.1 Extension of this work

Although the work in this thesis completes the design of MU beamforming based on IRS and networks, there are a couple of following works that remain to be discussed and solved in the future.

- Following the MU transmission, it can be common that EE and SE may not achieve their corresponding maximum with MU directions does not satisfy the optimal condition in Chapter 4. In this case, EE and SE have a trade-off relationship and require further investigation. More advanced optimization problems can be formulated to maximize EE and SE respectively.

- More refined discussion on the positions of transceivers should be made for MU transmission in the networks. In particular, the analysis of the sum rate limit in Chapter 4 is derived from single-user single-path (SUSP) scheme and MUSP scheme. Without some special conditions, the sum rate limit may not hold for SUMP transmission scheme. In addition, the analysis for the general scheme multi-user multi-path (MUMP) is another interesting topic for further investigation in our future works to have a more in-depth analysis. Also, the efficient routing of MU within a specific network graph topology may also be needed.

- The practical conditions should be further extended since only mutual coupling and near-field effects are analyzed in this thesis. In addition, following the near-field transmission, the weights should also be properly considered to leverage the redundant beams. The future design may also require a different design to avoid strong MU interference since the locations of redundant beams in near-field are subjected to different weights designs. It is also noticed by the author that exploring the co-design, including the IRS's architecture, circuit, and algorithms is also very interesting but challenging.

- In most of the theoretical works [63,160], including the work in this thesis, the effective area of a single IRS is normalized with a far-field assumption; therefore, the received power can be equivalent to the peak value of power gain. Considering effective area [130], the power gain should be normalized with respect to the total power impinging

on the surface. And the reflected power, which the next object will receive, is the integral of the reflected beam pattern in the angular domain, where the lower bound and upper bound of the integral is related to the effective area of the next object. Within the near-field, the effective area of IRS elements for transceivers is different given different transceivers' positions, e.g., directions and distances concerning the IRS. Thus, the optimal power gain of MU will be subject to the distance in this case. However, the effective area and the specific path loss should involve a specific discussion on the near-field effect [77, 130], which is temporally out of the scope of this thesis but will be one of the focuses in the future.

6.2.2 Promising Future Direction

It is no doubt that the application of IRS has a very promising prospect in future wireless communication networks. Based on the knowledge of the author, following directions should be of great importance to exploit more features and value about IRS based programmable wireless channel.

- For leveraging the benefits brought by the IRS networks, efficient estimation is necessary for transceivers. In particular, the accuracy for achieving it should be maximized, and the time slots required for estimation should be minimized. Thus, enhancing the efficiency can be seen as optimizing the tasks' efficiency concluded in Chapter 4. Advanced channel estimation methods may be also needed since the methods proposed in this work require a relatively larger number of IRS than the number of non-IRS scatters. Or it is very challenging to estimate every single parameter of these non-IRS scatters. The complexity for channel estimation increases as the number of non-IRS scatters increases. On the other hand, the overhead for channel estimation can be huge or even forbidden once we have too many reflections with many non-IRS scatters. We think one way to mitigate the channel estimation overhead is using hierarchical estimation with respect to reflection orders. For example, there is no need to wait to get all channel parameters in these tasks before data transmission. The initial communication can be activated after obtaining the FOR channel, though this may not fully leverage the spatial multiplexing gain and channel capacity. With more cycles of uplink and downlink data transmissions, the channel estimation can be gradually completed. Efficiency can be improved using the transmitted data instead of just estimation pilots, which can be regarded as one part of the integrated sensing and communication (ISAC) [161].

- Limitation of IRS networks in time domain analysis should be implemented. For instance, the MOR effect should be cut off since it is trivial to consider an infinite re-

sponse in the time domain. In addition, specific deployments of IRS introduce selective fading channel, since multi-path components with different order of reflections superpose and cause different effects in the signal processing in the time domain. Thus, the system performance should be investigated from this view.

- Due to the abundant spectrum resource in the high-frequency band, broadband design should also be valuable for investigation. However, the beam squint effect, which is induced by the correlation of elements with a specific frequency, can be a potential issue to fundamentally affect the transmission with the aid of IRS, thus it requires an in-depth analysis [36]. Some interesting results may be obtained by combining the physical characteristics of beam squint and redundant beams for MU transmission.

Appendix A

Derivation of Theorems

A.1 Proof for Redundant Beams in ULA

A.1.1 Proof of Corollary 2

Under condition of $d = \lambda/2$ with ULA IRS, equation (3.35) becomes

$$B(\phi_{out,v}, \phi_{in,u}) = \left| \sum_{m=0}^{M-1} w_m e^{-j\pi m(cos\phi_{out,v} + cos\phi_{in,u})} \right| . \tag{A.1}$$

For $N = 2$, from equation (3.15) and (3.16), we have the constraints:

$$\begin{cases} \delta_{lower} \leq \mathbf{w}^T \mathbf{a}_C(\phi_{out,1}, \phi_{in,1}) \\ \delta_{upper} \geq \mathbf{w}^T \mathbf{a}_C(\phi_{out,2}, \phi_{in,1}) \\ \delta_{lower} \leq \mathbf{w}^T \mathbf{a}_C(\phi_{out,2}, \phi_{in,2}) \\ \delta_{upper} \geq \mathbf{w}^T \mathbf{a}_C(\phi_{out,1}, \phi_{in,2}) \end{cases} . \tag{A.2}$$

Let

$$F(\phi_{out,v}, \phi_{in,u}) = e^{-j\pi f(\phi_{out,v}, \phi_{in,u})} , \tag{A.3}$$

and

$$f(\phi_{out,v}, \phi_{in,u}) = cos\phi_{out,v} + cos\phi_{in,u} . \tag{A.4}$$

It is clear that the term $F(\phi_{out,v}, \phi_{in,u})$ directly correlate to $\mathbf{a}_C(\phi_{out,v}, \phi_{in,u})$ from equation (2.11) since this term is the base of the new merged steering vector, i.e.,

$$\mathbf{a}_C(\phi_{out,v}, \phi_{in,u}) = [F(\phi_{out,v}, \phi_{in,u})^0, F(\phi_{out,v}, \phi_{in,u})^1, \ldots, F(\phi_{out,v}, \phi_{in,u})^{(M-1)}]. \quad \text{(A.5)}$$

Thus, if the following relationship

$$F(\phi_{out,1}, \phi_{in,1}) = F(\phi_{12}, \phi_{in,2}) \quad \text{(A.6)}$$

holds, we have

$$\mathbf{a}_C(\phi_{out,1}, \phi_{in,1}) = \mathbf{a}_C(\phi_{12}, \phi_{in,2}) \quad \text{(A.7)}$$

and

$$\delta_{lower} \leq \mathbf{w}^T \mathbf{a}_C(\phi_{12}, \phi_{in,2}) \, . \quad \text{(A.8)}$$

Equation (A.8) implies that the redundant beam is at ϕ_{12} for 2^{nd} pair transceiver. It can be explained as the numerical equivalence of two merged steering vectors of the desired direction and the direction of the redundant beam. As long as one constraint of the steering vector is satisfied in one direction, this constraint would broadcast to another direction. In a general sense, any other steering vector which numerically equals the constrained steering vectors of minimum secured power response would indicate redundant beams. Note that the phenomenon above also applies to the constrained steering vector under suppressed response.

Next, let's discuss the condition to make equation (A.6) hold, or in other words, the existence of the redundant beam. Since the function F in equation (A.3) is actually the phasor with respect to $\pi f(\phi_{out,v}, \phi_{in,u})$. Therefore, we have

$$e^{-j\pi f(\phi_{out,v}, \phi_{in,u})} = e^{-j\pi(f(\phi_{out,v}, \phi_{in,u})+2n)}, n \in Z \quad \text{(A.9)}$$

due to its periodicity. And n is the integer indicating the n-th order of equivalence.

If equation (A.6) holds, then it is necessary and sufficient that

$$f(\phi_{out,1}, \phi_{in,1}) + 2n = f(\phi_{12}, \phi_{in,2}), n \in Z \quad \text{(A.10)}$$

holds. If $n = 0$ and assume the redundant beam exist, then from equation (A.10), we set a middle variable γ_{12},

$$\gamma_{12} = \cos\phi_{in,1} + \cos\phi_{out,1} - \cos\phi_{in,2} = \cos\phi_{12} \, . \quad \text{(A.11)}$$

So ϕ_{12} is the solution of this equation and it equals to $\arccos(\gamma_{12})$, meaning the 0^{th} order of

redundant beam at ϕ_{12} is caused by the 1st transceiver pair and can be "seen" by Rx2. Note that n can equal other integers, and this may cause multiple orders of redundant beams as long as solution ϕ_{12} comes from n-th order equivalence exists. Namely, by substituting equation (A.11) into (A.10), we have

$$\phi_{12} = \arccos(\gamma_{12} + 2n) = \arccos(\cos\phi_{in,1} + \cos\phi_{out,1} - \cos\phi_{in,2} + 2n). \tag{A.12}$$

the discussion of classification based on the value range of γ_{12} which can make ϕ_{12} exist from function arccos is as follows.

Case **1**, when $\gamma_{12} \in [-3,-1)$. If $n \geq 2$, $\gamma_{12} + 2n > 1$ holds and ϕ_{12} does not exist. If $n \leq 0$, $\gamma_{12} + 2n < -1$ holds and ϕ_{12} does not exist. If $n = 1$, $-1 \leq \gamma_{12} + 2n < 1$ holds, $\phi_{12} = \arccos(\gamma_{12} + 2)$.

Case **2**, when $\gamma_{12} \in (1,3]$, which is similar to last case, if $n = -1$, $\phi_{12} = \arccos(\gamma_{12} - 2)$ or otherwise ϕ_{12} does not exist.

Case **3**, when $\gamma_{12} \in (-1,1)$. If $n = 1$, $\phi_{12} = \arccos(\gamma_{12})$ or otherwise ϕ_{12} does not exist.

Case **4**, when $\gamma_{12} = -1$. If $n = 1$, $\phi_{12} = 0°$. If $n = -1$, $\phi_{12} = 180°$. Otherwise, ϕ_{12} does not exist. So, two redundant beams occur at the margin angles, which are almost parallel to the surface.

Case **5**, which is similar to case **4**. When $\gamma_{12} = 1$, two redundant beams occur at $\phi_{12} = 180°$ and $0°$.

Case **4** and Case **5** are on the boundary, so we merge it in Case **1** and Case **2** respectively. In the cases analysis of the following discussion, the middle variable equals to 1 or -1 will also be merged into cases of $(1,3]$ and $[-3,-1)$ respectively. Last, it can be summarized as **Corollary 2**.

The results above provide insight into the redundant beam. The phenomenon is actually caused by considering the reflection process where one more path is involved compared to traditional transceivers. As we can find, the redundant beams are similar to the concept of grating lobes in the traditional antenna array theory. However, they are different and it might be tricky to distinguish those two concepts. For instance, if we consider the setting of such corollary, in scenarios of the antenna array, the condition of grating lobes are similar to equation (A.10), but in the form of

$$f(\phi_{out,1}) + 2n = f(\phi_{out,11}), n \in Z, \tag{A.13}$$

, where $\phi_{out,11}$ is the grating lobes' position. E.g., 0th order equivalence means the 0th side lobe is the main beam and the 1st order equivalence means the position of the first side lobe. In the IRSs' scenarios, as we have mentioned, the 0th order equivalence of equation (A.10)

means the main beam of Tx1 is actually equivalent to producing the 0^{th} order redundant beam on the beampattern of Tx2. Interchangeably, Tx2 can also produce the 0^{th} order redundant beam on the beampattern of Tx1. However, there are also grating lobes, as long as the grating lobes position $\phi_{out,11}$ satisfy

$$f(\phi_{out,1}, \phi_{in,1}) + 2n = f(\phi_{out,11}, \phi_{in,1}), n \in Z. \tag{A.14}$$

Therefore, we stress that redundant beams are caused by pairs of each other while grating lobes are self-introduced.

A.1.2 Proof of Corollary 3

When $d \in (0, \lambda/4)$, equation (A.10) can be rewritten as

$$f(\phi_{out12}, \phi_{in,1}) + \frac{2\pi n}{kd} = f(\phi_{out12}, \phi_{in,2}), n \in Z\ . \tag{A.15}$$

The form of equation (A.15) is more general than that of **Corollary 2**. However, base on the value range of γ_{12}, the redundant beam may not exist if no integer n can make $-1 \leq \gamma_{12} + \frac{2\pi n}{kd} \leq 1$, such that $\gamma_{12} + \frac{2\pi n}{kd} = cos\phi_{12}$ have the solution. By substituting equation (A.11) into (A.15), the discussion based on the value range of γ_{12} which can make ϕ_{12} exist is as follows.

Case **1**, when $\gamma_{12} \in (-1, 1)$. If $n = 0$, $\phi_{12} = \arccos(\gamma_{12})$. If $n \neq 0$, since $\gamma_{12} + \frac{2\pi n}{kd} \notin (-1, 1)$ so ϕ_{12} does not exist.

Case **2**, when $\gamma_{12} \in [-3, -1]$. If $n \geq 1$, $\gamma_{12} + \frac{2\pi n}{kd} > 1$ holds. ϕ_{12} does not exist. If $n \leq 0$, $\gamma_{12} + \frac{2\pi n}{kd} < -1$ holds. ϕ_{12} also does not exist.

Case **3**, when $\gamma_{12} \in [1, 3]$. It is similar as the last case, so ϕ_{12} does not exist.

Therefore, it can be summarized as **Corollary 3**.

A.1.3 Proof of Corollary 4

In the case of $d \in [\lambda/4, \lambda/2)$, base on equation (A.15), to make ϕ_{12} exist, we have the following discussion.

Case **1**, when $\gamma_{12} \in (-1, 1)$. If $n = 0$, then ϕ_{12} exists and $\phi_{12} = \arccos(\gamma_{12})$. If $n \neq 0$, since $\gamma_{12} + \frac{2\pi n}{kd} \notin (-1, 1)$, ϕ_{12} does not exist.

Case **2**, when $\gamma_{12} \in [-3, -1]$. If $\gamma_{12} + \frac{2\pi}{kd} > 1$, $\gamma_{12} + \frac{2\pi n}{kd} > 1$ holds when $n \geq 1$. And $\gamma_{12} + \frac{2\pi n}{kd} < -1$ holds when $n \leq 1$. So when $\frac{\lambda}{d} > 1 - \gamma_{12}$, ϕ_{12} does not exist.

When $n = 1$, if $\gamma_{12} + \frac{2\pi}{kd} \leq 1$, since $k = 2\pi/\lambda$, $\phi_{12} = \arccos(\gamma_{12} + \frac{\lambda}{d})$. When $n \neq 1$, ϕ_{12} does not exist since $\gamma_{12} + \frac{2\pi n}{kd} \notin (-1, 1)$.

Case **3**, when $\gamma_{12} \in [1,3]$, similar to last case, if $\frac{\lambda}{d} < \gamma_{12}+1$, $\phi_{12} = \arccos(\gamma_{12} - \frac{\lambda}{d})$ or ϕ_{12} does not exist which result in **Corollary 4**.

A.2 Proof for Redundant Beams in URA

A.2.1 Proof of Corollary 5

Consider $d = \lambda/2$ and URA IRS. Then equation (3.35) becomes:

$$\mathbf{B}(\Omega_{out,v}, \Omega_{in,u}) = \mathbf{B}(\phi_{out,v}, \theta_{out,v}, \phi_{in,u}, \theta_{in,u}) , \tag{A.16}$$

$$\mathbf{B}(\Omega_{out,v}, \Omega_{in,u}) = \left| \sum_{m_y=0}^{M_y-1} \sum_{m_x=0}^{M_x-1} w_{m_x m_y} \cdot e^{-jk(f_{cs}(\Omega_{out,v},\Omega_{in,u})d_x m_x + f_{ss}(\Omega_{out,v},\Omega_{in,u})d_y m_y)} \right| . \tag{A.17}$$

For $N = 2$, from equation (3.15) and (3.16), the constraints can be expressed as follows

$$\begin{cases} \delta_{lower} \leq \mathbf{w}^T \mathbf{a}_C(\Omega_{out,1},\Omega_{in,1}) = \mathbf{w}^T \mathbf{a}_C(\phi_{out,1},\theta_{out,1},\phi_{in,1},\theta_{in,1}) \\ \delta_{lower} \leq \mathbf{w}^T \mathbf{a}_C(\Omega_{out,2},\Omega_{in,2}) = \mathbf{w}^T \mathbf{a}_C(\phi_{out,2},\theta_{out,2},\phi_{in,2},\theta_{in,2}) \\ \delta_{upper} \geq \mathbf{w}^T \mathbf{a}_C(\Omega_{out,2},\Omega_{in,1}) = \mathbf{w}^T \mathbf{a}_C(\phi_{out,2},\theta_{out,2},\phi_{in,1},\theta_{in,1}) \\ \delta_{upper} \geq \mathbf{w}^T \mathbf{a}_C(\Omega_{out,1},\Omega_{in,2}) = \mathbf{w}^T \mathbf{a}_C(\phi_{out,1},\theta_{out,1},\phi_{in,2},\theta_{in,2}) \end{cases} \tag{A.18}$$

Let

$$\begin{aligned} F_1(\Omega_{out,v}, \Omega_{in,u}) &= F_1(\phi_{out,v}, \theta_{out,v}, \phi_{in,u}, \theta_{in,u}) \\ &= e^{-j\pi f_{cs}(\phi_{out,v},\theta_{out,v},\phi_{in,u},\theta_{in,u})} , \end{aligned} \tag{A.19}$$

$$\begin{aligned} F_2(\Omega_{out,v}, \Omega_{in,u}) &= F_2(\phi_{out,v}, \theta_{out,v}, \phi_{in,u}, \theta_{in,u}) \\ &= e^{-j\pi f_{ss}(\phi_{out,v},\theta_{out,v},\phi_{in,u},\theta_{in,u})} . \end{aligned} \tag{A.20}$$

It is clear that the term $F_1(\Omega_{out,v}, \Omega_{in,u})$ and $F_2(\Omega_{out,v}, \Omega_{in,u})$ correlate to $\mathbf{a}_C(\Omega_{out,v}, \Omega_{in,u})$ from equation (2.12) directly since the multiplication of this two terms is the base of new merged steering vector. Thus, if the following relationship

$$F_1(\phi_{out,1}, \theta_{out,1}, \phi_{in,1}, \theta_{in,1}) = F_1(\phi_{12}, \theta_{12}, \phi_{in,2}, \theta_{in,2}) \tag{A.21}$$

and

$$F_2(\phi_{out,1}, \theta_{out,1}, \phi_{in,1}, \theta_{in,1}) = F_2(\phi_{12}, \theta_{12}, \phi_{in,2}, \theta_{in,2}) \tag{A.22}$$

hold, we have following equation

$$\mathbf{a}_C(\phi_{out,1}, \theta_{out,1}, \phi_{in,1}, \theta_{in,1}) = \mathbf{a}_C(\phi_{12}, \theta_{12}, \phi_{in,2}, \theta_{in,2}) \ . \tag{A.23}$$

We can get

$$\delta_{lower} \leq \mathbf{w}^T \mathbf{a}_C(\phi_{12}, \theta_{12}, \phi_{in,2}, \theta_{in,2}) \ , \tag{A.24}$$

which implies that the redundant beam of URA at ϕ_{12}, θ_{12}. And it is clear that if and only if equation (A.21) and (A.22) are both satisfied can both azimuth and elevation angle of redundant beam exist. It can be explained as the numerical equivalence of two merged steering vectors of the desired direction and the direction of redundant beam. As long as one constraint of steering vector is satisfied in one direction, this constraint would also affect another direction. In a general sense, any other steering vector which numerically equal to the constrained steering vectors of minimum secured power response would indicate redundant beams.

Next let's discuss the condition to make equation (A.21) and (A.22) hold respectively. To be noted that there is the periodicity of function F_1 and F_2 for its phasor form. If equation (A.21) and (A.22) hold, then it is necessary and sufficient that

$$f_{cs}(\phi_{out,1}, \theta_{out,1}, \phi_{in,1}, \theta_{in,1}) + 2n_1 = f_{cs}(\phi_{12}, \theta_{12}, \phi_{in,2}, \theta_{in,2}), n_1 \in Z \tag{A.25}$$

and

$$f_{ss}(\phi_{out,1}, \theta_{out,1}, \phi_{in,1}, \theta_{in,1}) + 2n_2 = f_{ss}(\phi_{12}, \theta_{12}, \phi_{in,2}, \theta_{in,2}), n_2 \in Z \tag{A.26}$$

hold respectively. Base on the definition of α_{12} and β_{12}, making equation (A.25) and (A.26) hold is equivalent to the following two equations

$$\alpha_{12} + 2n_1 = \cos\phi_{12} \sin\theta_{12}, n_1 \in Z \ , \tag{A.27}$$

$$\beta_{12} + 2n_2 = \sin\phi_{12} \sin\theta_{12}, n_2 \in Z \ , \tag{A.28}$$

have the correspondent solutions of ϕ_{12} and θ_{12}. Only under conditions of both left part of equation (A.27) and (A.28) have the value range of $[-1, 1]$ can they be solved out.

Specifically, when $\alpha_{pq}, \beta_{pq} \in [-3, -1]$, only if $n_1 = 1, n_2 = 1, -1 \leq \alpha_{12} + 2n_1 \leq 1$ and $-1 \leq \beta_{12} + 2n_2 \leq 1$ hold, $\phi_{12} = \arctan(\frac{\beta_{12}+2}{\alpha_{12}+2}) + n_3\pi, n_3 \in Z, \phi_{12} \in [0, 2\pi]$. Otherwise, ϕ_{12} does not exist. Then by further substituting ϕ_{12} into equation (A.28), $\theta_{12} = \arcsin(\frac{\beta_{12}+2}{\sin\phi_{12}}), \theta \in [0, \frac{\pi}{2}]$.

In this case, there are usually two solutions of ϕ_{12}. There might not be a correspondent solution of θ_{12} which means the redundant beam does not exist. Likewise, the discussion on the condition of ϕ_{12} and θ_{12} can be concluded in TABLE 3.1 and TABLE 3.2. Therefore, it can be summarized as **Corollary 5**.

From the results above, it provides insight into the redundant beam. As we can find, the redundant beams are similar to the concept of grating lobes in the traditional antenna array theory. However, they are different and it might be tricky to distinguish those two concepts. I.e, the grating lobes are caused by the large spacing between elements of antenna arrays such that the beamforming result can repeat in the angular domain of AOA/AOD more than once. For IRS, grating lobes are caused by the large spacing between IRS elements as well. However, due to IRS being equivalent to a cascading of antenna arrays, one more degree of freedom of AOA/AOD is introduced. E.g. for IRS in ULA, the "Bragg condition" can be written as

$$kd(\cos\theta_{out} + \cos\theta_{in}) = 2\pi n_0, n_0 \in \mathbb{Z}, \tag{A.29}$$

where θ_{in} and θ_{out} are incident direction and exit direction with respect to IRS for one pair transceiver. k and d are the wave number and spacing of elements, respectively. The Bragg condition for the traditional linear array is just

$$kd\cos\theta = 2\pi n_0, n_0 \in \mathbb{Z}, \tag{A.30}$$

where θ can be transmitted/received direction of the ULA [162]. And it is not hard for us to find the direction of transceivers and the element spacing jointly determining the existence of the "Bragg condition" for IRS in equation (A.29).

A.2.2 Proof of Corollary 6

Under condition of $d \in (0, \lambda/4)$, the equation (A.27) and (A.28) would be rewritten as

$$\alpha_{12} + \frac{2\pi n_1}{kd} = \cos\phi_{12}\sin\theta_{12}, n_1 \in Z\,, \tag{A.31}$$

$$\beta_{12} + \frac{2\pi n_2}{kd} = \sin\phi_{12}\sin\theta_{12}, n_2 \in Z\,, \tag{A.32}$$

which is the general form. The discussion on getting both ϕ_{12} and θ_{12} to make equation (A.31) and (A.32) hold is as follows.

When $\alpha_{12} \in [-1, 1]$ and $\beta_{12} \in [-1, 1]$. If $n_1 = 0$, $n_2 = 0$, $\phi_{12} = \arctan(\frac{\beta_{12}}{\alpha_{12}})$ and $\theta_{12} = \arcsin(\frac{\beta_{12}}{\sin\phi_{12}})$. If $n_1 \neq 0$ or $n_2 \neq 0$, since $\alpha_{12} + \frac{2\pi n_1}{kd} \notin (-1, 1)$ or $\beta_{12} + \frac{2\pi n_2}{kd} \notin (-1, 1)$, both ϕ_{12} and θ_{12} would not exist. Therefore, it can be summarized as **Corollary 6**.

A.2.3 Proof of Corollary 7

Consider URA IRS under condition of $d \in [\lambda/4, \lambda/2)$, the discussion of ϕ_{12} and θ_{12} base on value range of α_{12} and β_{12} to make equation (A.31) and (A.32) hold is as follows.

Case **1**, when $\alpha_{12} \in (-1,1)$ and $\beta_{12} \in (-1,1)$. If $n_1 = 0$, $n_2 = 0$, $\phi_{12} = \arctan(\frac{\beta_{12}}{\alpha_{12}})$ and $\theta_{12} = \arcsin(\frac{\beta_{12}}{\sin\phi_{12}})$. If $n_1 \neq 0$ or $n_2 \neq 0$, since $\alpha_{12} + \frac{2\pi n_1}{kd} \notin (-1,1)$ or $\beta_{12} + \frac{2\pi n_2}{kd} \notin (-1,1)$, both ϕ_{12} and θ_{12} would not exist.

Case **2**, when $\alpha_{12} \in [-3,-1]$ and $\beta_{12} \in [-3,-1]$. If $\alpha_{12} + \frac{2\pi}{kd} > 1$, $\alpha_{12} + \frac{2\pi n_1}{kd} > 1$ holds when $n_1 > 1$. And $\alpha_{12} + \frac{2\pi n}{kd} < -1$ holds when $n_1 < 1$. So when $\frac{\lambda}{d} > 1 - \alpha_{12}$, ϕ_{12} does not exist. And in another hand, when $\frac{\lambda}{d} > 1 - \beta_{12}$, θ_{12} does not exist as well.

Only if $\alpha_{12} + \frac{2\pi}{kd} \leq 1$ and $\beta_{12} + \frac{2\pi}{kd} \leq 1$, when $n_1 = 1$ and $n_2 = 1$, $\phi_{12} = \arctan(\frac{\beta_{12}+\frac{\lambda}{d}}{\alpha_{12}+\frac{\lambda}{d}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}+\frac{\lambda}{d}}{\sin\phi_{12}})$. Otherwise when $n_1 \neq 1$ or $n_2 \neq 1$, ϕ_{12} or θ_{12} do not exist since $\alpha_{12} + \frac{2\pi n_1}{kd} \notin (-1,1)$ and $\beta_{12} + \frac{2\pi n_2}{kd} \notin (-1,1)$.

Case **3**, when $\alpha_{12} \in [-3,-1]$ and $\beta_{12} \in (-1,1)$.

$\phi_{12} = \arctan(\frac{\beta_{12}}{\alpha_{12}+\frac{\lambda}{d}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}}{\sin\phi_{12}})$, under the condition that $\alpha_{12} + \frac{2\pi}{kd} \leq 1$.

Case **4**, when $\alpha_{12} \in [-3,-1]$ and $\beta_{12} \in [1,3]$.

$\phi_{12} = \arctan(\frac{\beta_{12}-\frac{\lambda}{d}}{\alpha_{12}+\frac{\lambda}{d}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}-\frac{\lambda}{d}}{\sin\phi_{12}})$, under the condition that $\alpha_{12} + \frac{2\pi}{kd} \leq 1$ and $\beta_{12} - \frac{2\pi}{kd} \geq -1$.

Case **5**, when $\alpha_{12} \in (-1,1)$ and $\beta_{12} \in [-3,-1]$.

$\phi_{12} = \arctan(\frac{\beta_{12}+\frac{\lambda}{d}}{\alpha_{12}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}+\frac{\lambda}{d}}{\sin\phi_{12}})$, under the condition that $\beta_{12} + \frac{2\pi}{kd} \leq 1$.

Case **6**, when $\alpha_{12} \in (-1,1)$ and $\beta_{12} \in [1,3]$.

$\phi_{12} = \arctan(\frac{\beta_{12}-\frac{\lambda}{d}}{\alpha_{12}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}-\frac{\lambda}{d}}{\sin\phi_{12}})$, under the condition that $\beta_{12} - \frac{2\pi}{kd} \geq -1$.

Case **7**, when $\alpha_{12} \in [1,3]$ and $\beta_{12} \in [-3,-1]$.

$\phi_{12} = \arctan(\frac{\beta_{12}+\frac{\lambda}{d}}{\alpha_{12}-\frac{\lambda}{d}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}+\frac{\lambda}{d}}{\sin\phi_{12}})$, under the condition that $\alpha_{12} - \frac{2\pi}{kd} \geq -1$ and $\beta_{12} + \frac{2\pi}{kd} \leq 1$.

Case **8**, when $\alpha_{12} \in [1,3]$ and $\beta_{12} \in (-1,1)$.

$\phi_{12} = \arctan(\frac{\beta_{12}}{\alpha_{12}-\frac{\lambda}{d}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}}{\sin\phi_{12}})$, under the condition that $\alpha_{12} - \frac{2\pi}{kd} \geq -1$.

Case **9**, when $\alpha_{12} \in [1,3]$ and $\beta_{12} \in [1,3]$.

$\phi_{12} = \arctan(\frac{\beta_{12}-\frac{\lambda}{d}}{\alpha_{12}-\frac{\lambda}{d}}) + n_3\pi$ and $\theta_{12} = \arcsin(\frac{\beta_{12}-\frac{\lambda}{d}}{\sin\phi_{12}})$, under the condition that $\alpha_{12} - \frac{2\pi}{kd} \geq -1$ and $\beta_{12} - \frac{2\pi}{kd} \geq -1$. Therefore, it can be summarized as **Corollary 7**.

A.3 Proof of Optimal Condition on Transceivers' Direction

Lemma 2 can be proved by analysing the IRS channel as a whole. I.e, we start by analysing the channel of a single transceiver pair assisted by a single IRS. By referring to equation (4.2), we denote $\mathbf{H}_{I,1} = \mathbf{H} = \mathbf{A}_{out}^T \mathbf{W} \mathbf{A}_{in}$ for simplicity. Thus, the channel between Tx_i and Rx_i

in **H** can be written as

$$h_{ii} = \mathbf{w}^H \mathbf{a}_C(\phi_{out,i}, \phi_{in,i}). \tag{A.33}$$

We can observe that the diagonal terms in the matrix of equation (4.2) are signal gains for each Rx and these terms are required to be maximized. Other off-diagonal terms are the interference gain which should be minimized. Therefore, by calculating an optimal weights vector **w** such that the diagonal terms are maximized while nullifying off-diagonal terms, the optimal IRS-based channel can be obtained and the optimal sum rate can be achieved. Note that if the single IRS is considered as the ULA or URA specification, which has the characteristic of equal spacing between each element, the optimal weights can be analytically obtained simply by the MRC algorithm. Specifically, for ULA scenario and we let the i-th pair user locate at $\phi_{in,i} = \alpha_i^\circ, \phi_{out,i} = \beta_i^\circ$, equation (A.33) can be rewritten as

$$h_{ii} = \mathbf{w}^H \mathbf{a}_C(\phi_{out,i}, \phi_{in,i}) = \sum_{m=0}^{M-1} w_m e^{-jkd(\cos\alpha_i + \cos\beta_i)m}, \tag{A.34}$$

where $k = \frac{2\pi}{\lambda}$ is the wave number, d is the distance between each element, and λ is the carrier wavelength. The path loss here is assumed to be a constant value. Thus, with unit power constraint on each IRS element, the weight on an IRS can then be expressed as

$$w_m = e^{j\Xi_m}, \Xi_m \in (0, 2\pi], m = 1, \ldots, M \ . \tag{A.35}$$

As we can find, a necessary condition for $|h_{ii}| = M$ is that the weights need to guarantee each term in the summation in phase by writing the channel gain as

$$h_{ii} = \mathbf{w}^H \mathbf{a}_C(\phi_{out,i}, \phi_{in,i}) = \sum_{m=0}^{M-1} e^{-jkd(\cos\alpha_i + \cos\beta_i + \zeta_m)m}, \tag{A.36}$$

where ζ_m is an arbitrary term that comes from $\angle[w_m]$, the phase design on each element of IRS, we can observe the maximal value of $|h_{ii}| = M$ is guaranteed as long as

$$kd(cos\alpha_i + cos\beta_i + \zeta_m) = 2\pi n_1, n_1 \in Z. \tag{A.37}$$

Denote $\Delta r = \frac{d}{\lambda}$, which is the normalized spacing between each element. We can compute the weight value on m-th element such as

$$\zeta_m = -\cos\alpha_i - \cos\beta_i + \frac{K}{\Delta r}, \tag{A.38}$$

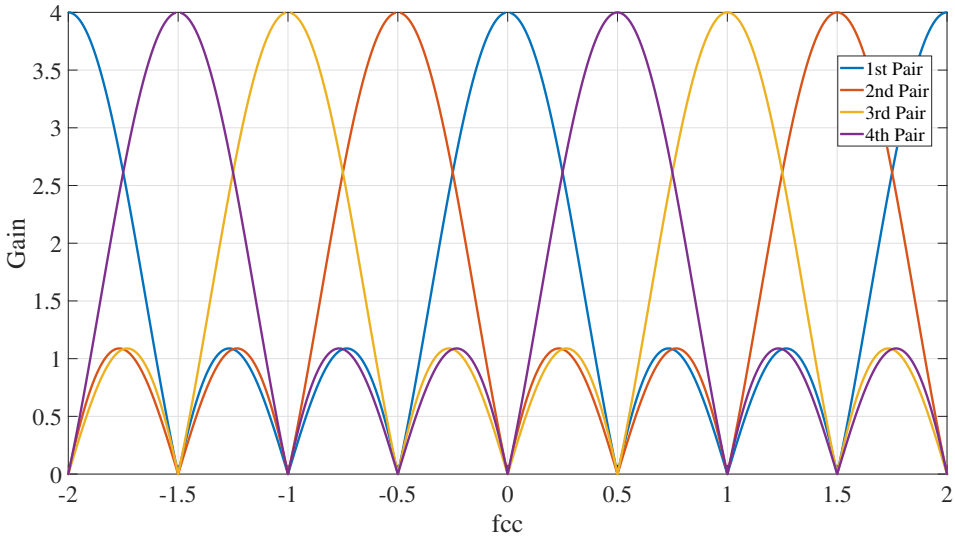


Figure A.1: Optimal Spatial Multiplexing of h_{1j},h_{2j},h_{3j} and h_{4j}, M=4,d=$\frac{\lambda}{2}$, L=2.

to equalize the phase shifts. This is essentially the same as using the MRC algorithm to calculate the weights vector. Actual phase of weights can be obtained by $\Xi_m = -\zeta_m k dm$. Then, after applying the result of MRC, since the weights have been determined, we can analyze other terms in the i-th column of the matrix in equation (4.2) and write them as

$$h_{ji} = \mathbf{w}^H \mathbf{a}_C(\phi_{out,j}, \phi_{in,i}) = \sum_{m=0}^{M-1} e^{-j2\pi\Delta r(\cos\beta_j - \cos\beta_i + \frac{K}{\Delta r})m}, \tag{A.39}$$

where $\Delta r = \frac{d}{\lambda}$. Denote $f_{cc} = (\cos\beta_j - \cos\beta_i + \frac{K}{\Delta r})$ and $L = M\Delta r$ which are the variable in angular domain and normalized length of IRS. Therefore, h_{ji} can be generalized as the beam-pattern and thus becomes a function of f_{cc}

$$h_{ji}(f_{cc}) = \mathbf{w}^H \mathbf{a}_C(\phi_{out,j}, \phi_{in,i}) = \sum_{m=0}^{M-1} e^{-j2\pi\Delta r(\cos\beta_j - \cos\beta_i + \frac{K}{\Delta r})m} = e^{-j\Delta r f_{cc}(M-1)} \frac{\sin(\pi f_{cc} L)}{\sin(\pi f_{cc} \frac{L}{M})}. \tag{A.40}$$

We can simply verify that h_{ji} is a periodic function of f_{cc} and the period is $\frac{1}{\Delta r}$. If the period of $h_{ji}(f_{cc})$ is within the visible angular range, which is $f_{cc} \in [-2, 2]$ in this case, there can be $M-1$ other pairs of transceivers communicating at the same time. These pairs can use the same frequency of carrier since they are orthogonal in the angular domain, which is shown in Fig. A.1. The nullifying point of h_{ji} is also in the period of $\frac{1}{\Delta r}$, separated by $\frac{1}{L}$. Therefore, we can determine other Rx's position β_j such that there is no interference from the i-th Tx

where the position can be calculated by

$$\beta_j = \cos^{-1}\left(\frac{j}{L} - \zeta_m - \cos\alpha_i \pm \frac{1}{\Delta r}\right). \tag{A.41}$$

This also means if other Rxs are standing in the same position as the nullifying position of Tx_i, there will be no interference from Tx_i, so other terms in the i-th column of channel matrix can be nullified. In addition, since the weights have been calculated as ζ_m is set by the first pair, given the position of Rx_j, we can calculate the optimal position of Tx_j correspondingly leveraging the equation (A.37) which is

$$\alpha_j = \cos^{-1}\left(\frac{j}{L} - \zeta_m - \cos\beta_i \pm \frac{1}{\Delta r}\right). \tag{A.42}$$

A.4 Proof for Interference-free condition

To make the proof easy to follow, we assume $M = 4$ and $N = 2$, where M is the number of elements on IRS and N is the number of transceiver pairs. However, it is worth noting that this conclusion can be extended to arbitrary numbers of N and M. Following the definition in Chapter 2, we have

$$\mathbf{A}_{in} = \mathbf{A} = [\mathbf{a}(\phi_{in,1}), \mathbf{a}(\phi_{in,2})] = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \\ a_{14} & a_{24} \end{bmatrix}, \tag{A.43}$$

which is the steering matrix of incident direction toward IRS, and $\mathbf{a}(\phi_{in,i}), i = 1,2$ is the steering vector of the incident directly on the IRS. Note that this is also the channel from the Txs to the IRS. Similarly, we define the steering matrix of exit directions, which also is the channel from the IRS to Rxs, as

$$\mathbf{A}_{out} = [\mathbf{a}(\phi_{out,1}), \mathbf{a}(\phi_{out,2})] = \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix}^T, \tag{A.44}$$

where we change the notations of $\mathbf{A}_{out}$ by $\mathbf{B}$ for easy understanding. The weight matrix of IRS is defined as $\mathbf{W}$, which is

$$\mathbf{W} = \begin{bmatrix} w_1^* & 0 & 0 & 0 \\ 0 & w_2^* & 0 & 0 \\ 0 & 0 & w_3^* & 0 \\ 0 & 0 & 0 & w_4^* \end{bmatrix}. \tag{A.45}$$

By ignoring the noise term, we can write the received signal vector as

$$\hat{\mathbf{y}}_r = \mathbf{B}^T\mathbf{W}\mathbf{A}\mathbf{s} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix} \begin{bmatrix} w_1^* & 0 & 0 & 0 \\ 0 & w_2^* & 0 & 0 \\ 0 & 0 & w_3^* & 0 \\ 0 & 0 & 0 & w_4^* \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \\ a_{14} & a_{24} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}. \tag{A.46}$$

where the vector $\mathbf{s} = [s_1\ s_2]^T$ is the vector of transmitted signal from Tx_1 and Tx_2.

Next, some terms can be rewritten into a more regular form in order to have a channel expression which is similar to a traditional MIMO model. Therefore, by factorizing the term above in equation (A.46), we can have

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} [w_1^* & w_2^* & w_3^* & w_4^*] \begin{bmatrix} \mathbf{A}_{C,1} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \\ [w_1^* & w_2^* & w_3^* & w_4^*] \begin{bmatrix} \mathbf{A}_{C,2} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \end{bmatrix}, \tag{A.47}$$

where

$$\mathbf{A}_{C,1} = \begin{bmatrix} \mathbf{a}_C(\phi_{out,1}, \phi_{in,1}) & \mathbf{a}_C(\phi_{out,1}, \phi_{in,2}) \end{bmatrix} = \begin{bmatrix} b_{11}a_{11} & b_{11}a_{21} \\ b_{12}a_{12} & b_{12}a_{22} \\ b_{13}a_{13} & b_{13}a_{23} \\ b_{14}a_{14} & b_{14}a_{24} \end{bmatrix}, \tag{A.48}$$

and

$$\mathbf{A}_{C,2} = \begin{bmatrix} \mathbf{a}_C(\phi_{out,2}, \phi_{in,1}) & \mathbf{a}_C(\phi_{out,2}, \phi_{in,2}) \end{bmatrix} = \begin{bmatrix} b_{21}a_{11} & b_{21}a_{21} \\ b_{22}a_{12} & b_{22}a_{22} \\ b_{23}a_{13} & b_{23}a_{23} \\ b_{24}a_{14} & b_{24}a_{24} \end{bmatrix}, \tag{A.49}$$

where $\mathbf{a}_C(\phi_{out,1}, \phi_{in,1}) = \mathbf{a}(\phi_{in,1}) \odot \mathbf{a}(\phi_{out,1})$, $\mathbf{a}_C(\phi_{out,1}, \phi_{in,1}) = \mathbf{a}(\phi_{in,2}) \odot \mathbf{a}(\phi_{out,1})$. There-

fore we can get the i-th user received signal as in [144]

$$\hat{y}_{r,i} = \mathbf{w}^H \mathbf{A}_{C,i}\mathbf{s} + n_i, i = 1, 2, ..., N \ . \tag{A.50}$$

where the n_i is the additive noise term at each Rx and $\mathbf{w}$ is obtained by taking all the diagonal terms in $\mathbf{W}$, which is

$$\mathbf{w} = [w_1 \ w_2 \ w_3 \ w_4]^T. \tag{A.51}$$

Note that that in equation A.47, for different Rxs, their received signal is obtained along different steering matrix $\mathbf{A}_{C,i}$ but processed by the same weight vector $\mathbf{w}$. Due to $\mathbf{A}_{C,1}$ and $\mathbf{A}_{C,2}$ sharing the same incident matrix, we can combine them further and move the difference between the two different steering matrices to the weight vector. Thus, through dividing $\mathbf{A}_{C,2}$ by $\mathbf{A}_{C,1}$ element-wisely, we can have a matrix $\mathbf{C}$ which can be regarded as a factor of Hadamard product such that

$$\mathbf{A}_{C,1} \odot \mathbf{C} = \mathbf{A}_{C,2}, \text{and} \, \mathbf{C} = \begin{bmatrix} \frac{b_{21}}{b_{11}} & \frac{b_{21}}{b_{11}} \\ \frac{b_{22}}{b_{12}} & \frac{b_{22}}{b_{12}} \\ \frac{b_{23}}{b_{13}} & \frac{b_{23}}{b_{13}} \\ \frac{b_{24}}{b_{14}} & \frac{b_{24}}{b_{14}} \end{bmatrix}. \tag{A.52}$$

Actually, the term $\frac{b_{21}}{b_{11}} = e^{-jkd(cos\phi_{out,2} - \cos\phi_{out,1})0}$ for ULA case is complex constant where k and d are wave number and distance between elements respectively. Then, we can have $\frac{b_{2m}}{b_{1m}} = e^{-jkd(cos\phi_{out,2} - \cos\phi_{out,1})(m-1)}, m = 1, 2, ..., M$. Although the terms' equivalence in the same column like $\frac{b_{21}}{b_{11}} = \frac{b_{22}}{b_{12}} = ... = \frac{b_{24}}{b_{14}}$ can be achieved with the increasing of the iterative power term $(m-1)$ which means the steering matrices are same, we can assume that d is small enough so that the overall complex term can not repeat in the period of itself and we can have $\frac{b_{21}}{b_{11}} \neq \frac{b_{22}}{b_{12}} \neq ... \neq \frac{b_{24}}{b_{14}}$ given the directions of angle $\phi_{out,1} \neq \phi_{out,2}$. Next, we note that the columns of the matrix $\mathbf{C}$ are the same, then we can rewrite equation (A.47) as

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{w}^H \mathbf{A}_{C,1}\mathbf{s} \\ \mathbf{w}^H \mathbf{A}_{C,1} \odot \mathbf{C}\mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{w}^H \mathbf{A}_{C,1}\mathbf{s} \\ \mathbf{w_C}^H \mathbf{A}_{C,1}\mathbf{s} \end{bmatrix}, \tag{A.53}$$

$$\mathbf{w_C} = [w_{c1} \ w_{c2} \ ... \ w_{c4}] = [w_1 \frac{b_{21}}{b_{11}} \ w_2 \frac{b_{22}}{b_{12}} \ ... \ w_2 \frac{b_{24}}{b_{14}}]^T. \tag{A.54}$$

$\mathbf{w_C}$ is the equivalent vector for the second Rx $\hat{y}_2$ and we can know that it has a mapping relationship to $\mathbf{w}$, which is the unique characteristic in the IRS's model. Therefore, by combining

the common term in equation (A.53), we have

$$\hat{\mathbf{y}}_r = \begin{bmatrix} \mathbf{w}^H \\ \mathbf{w_C}^H \end{bmatrix} \begin{bmatrix} \mathbf{A}_{C,1}\mathbf{s} \end{bmatrix} = \begin{bmatrix} w_1^* & w_2^* & w_3^* & w_4^* \\ w_{c1}^* & w_{c2}^* & w_{c3}^* & w_{c4}^* \end{bmatrix} \begin{bmatrix} b_{11}a_{11} & b_{11}a_{21} \\ b_{12}a_{12} & b_{12}a_{22} \\ b_{13}a_{13} & b_{13}a_{23} \\ b_{14}a_{14} & b_{14}a_{24} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}, \tag{A.55}$$

and by multiplying the weight matrix with the steering matrix, we have

$$\hat{\mathbf{y}}_r = \begin{bmatrix} \mathbf{w}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,1}) & \mathbf{w}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,2}) \\ \mathbf{w_C}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,1}) & \mathbf{w_C}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,2}) \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}. \tag{A.56}$$

To suppress the interference, we need to diagonalize the matrix in equation (A.56). Namely, the weights vector $\mathbf{w}$ should satisfy

$$\begin{cases} \mathbf{w}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,1}) = \delta_1 \\ \mathbf{w}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,2}) = 0 \\ \mathbf{w_C}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,1}) = 0 \\ \mathbf{w_C}^H\mathbf{a}_C(\phi_{out,1},\phi_{in,2}) = \delta_2 \end{cases}, \tag{A.57}$$

where δ_1 and δ_2 are non-zero values. Since $\mathbf{w_C}$ can be replaced by $\mathbf{w}$, we can present equation (A.57) by using matrix as

$$\begin{bmatrix} b_{11}a_{11} & b_{12}a_{12} & b_{13}a_{13} & b_{14}a_{14} \\ b_{11}a_{21} & b_{12}a_{22} & b_{13}a_{23} & b_{14}a_{24} \\ b_{21}a_{11} & b_{22}a_{12} & b_{23}a_{13} & b_{24}a_{14} \\ b_{21}a_{21} & b_{22}a_{22} & b_{23}a_{23} & b_{24}a_{24} \end{bmatrix} \begin{bmatrix} w_1^* \\ w_2^* \\ w_3^* \\ w_4^* \end{bmatrix} = \begin{bmatrix} \delta_1 \\ 0 \\ 0 \\ \delta_2 \end{bmatrix}. \tag{A.58}$$

As the left-hand matrix is full rank assured by the assumption above, 4 linear equations with 4 unknowns can be solved with a non-zero solution. Moreover, by increasing the element number such that $M >> N^2$, the solution space will be further enlarged. Thus, there must be multiple non-zero solutions to achieve the diagonalization of the matrix in the equation (A.56). In this case, the weights $\mathbf{w}$ and $\mathbf{w_C}$ can be nearly orthogonal to each other. As a result, the equivalence between traditional MIMO and IRS is established, and the interference can be suppressed among multiple transceiver pairs.

Bibliography

[1] H. T. Friis, A Note on A Simple Transmission Formula, Proceedings of the IRE 34 (5) (1946) 254–256.

[2] S. Geng, J. Kivinen, X. Zhao, P. Vainikainen, Millimeter-wave propagation channel characterization for short-range wireless communications, IEEE Transactions on Vehicular Technology 58 (1) (2009) 3–13. doi:10.1109/TVT.2008.924990.

[3] W. Liu, S. Weiss, Wideband Beamforming: Concepts and Techniques, John Wiley & Sons, 2010.

[4] D. Tse, P. Viswanath, Fundamentals of wireless communication, Cambridge university press, 2005.

[5] H. Yang, X. Cao, F. Yang, J. Gao, S. Xu, M. Li, X. Chen, Y. Zhao, Y. Zheng, S. Li, A Programmable Metasurface with Dynamic Polarization, Scattering and Focusing Control, Scientific reports 6 (2016) 35692.

[6] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, I. Akyildiz, A New Wireless Communication Paradigm through Software-Controlled Metasurfaces, IEEE Communications Magazine 56 (9) (2018) 162–169. doi:10.1109/MCOM.2018.1700659.

[7] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, G. Wunder, 5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice, IEEE Journal on Selected Areas in Communications 35 (6) (2017) 1201–1221. doi:10.1109/JSAC.2017.2692307.

[8] J. L. Pawsey, R. N. Bracewell, Radio Astronomy, Oxford (1955).

[9] F. Han, Y.-H. Yang, B. Wang, Y. Wu, K. J. R. Liu, Time-reversal division multiple access over multi-path channels, IEEE Transactions on Communications 60 (7) (2012) 1953–1965. doi:10.1109/TCOMM.2012.051012.110531.

[10] T. Tamir, On radio-wave propagation in forest environments, IEEE Transactions on Antennas and Propagation 15 (6) (1967) 806–817.

[11] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. Maccartney, T. S. Rappaport, A. Alsanie, Radio Propagation Path Loss Models for 5G Cellular Networks in the 28 GHZ and 38 GHZ Millimeter-wave Bands, IEEE Communications Magazine 52 (9) (2014) 78–86. doi:10.1109/MCOM.2014.6894456.

[12] M. Hata, Empirical formula for propagation loss in land mobile radio services, IEEE Transactions on Vehicular Technology 29 (3) (1980) 317–325. doi:10.1109/T-VT.1980.23859.

[13] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, F. Gutierrez, Millimeter Wave Mobile Communications for 5G cellular: It will work!, IEEE Access 1 (2013) 335–349. doi:10.1109/ACCESS.2013.2260813.

[14] V. Erceg, L. Greenstein, S. Tjandra, S. Parkoff, A. Gupta, B. Kulic, A. Julius, R. Bianchi, An empirically based path loss model for wireless channels in suburban environments, IEEE Journal on Selected Areas in Communications 17 (7) (1999) 1205–1211. doi:10.1109/49.778178.

[15] L. Zhang, W. Liu, R. J. Langley, A class of constrained adaptive beamforming algorithms based on uniform linear arrays, IEEE Transactions on Signal Processing 58 (7) (2010) 3916–3922. doi:10.1109/TSP.2010.2046078.

[16] Y. Lim, S. Parker, Fir filter design over a discrete powers-of-two coefficient space, IEEE Transactions on Acoustics, Speech, and Signal Processing 31 (3) (1983) 583–591. doi:10.1109/TASSP.1983.1164085.

[17] L. Zhang, W. Liu, Robust Beamforming for Coherent Signals Based On The Spatial-Smoothing Technique, Signal Processing 92 (11) (2012) 2747–2758.

[18] T. L. Marzetta, Massive MIMO: An Introduction, Bell Labs Technical Journal 20 (2015) 11–22.

[19] E. G. Larsson, O. Edfors, F. Tufvesson, T. L. Marzetta, Massive MIMO for Next Generation Wireless Systems, IEEE Communications Magazine 52 (2) (2014) 186–195. doi:10.1109/MCOM.2014.6736761.

[20] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, A. M. Sayeed, An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems, IEEE Journal of Selected Topics in Signal Processing 10 (3) (2016) 436–453.

[21] K. F. Braun, Electrical oscillations and wireless telegraphy, Nobel Lecture, December 11 (1909) (1909) 226–245.

[22] C. A. Balanis, Antenna Theory: Analysis and Design, John wiley & sons, 2015.

[23] A. Kaye, D. George, Transmission of multiplexed pam signals over multiple channel and diversity systems, IEEE Transactions on Communication Technology 18 (5) (1970) 520–526.

[24] W. van Etten, Maximum likelihood receiver for multiple channel transmission systems, IEEE Transactions on Communications 24 (2) (1976) 276–283.

[25] R. H. Roy III, B. Ottersten, Spatial division multiple access wireless communication systems, uS Patent 5,515,378 (May 7 1996).

[26] D. J. Rabideau, P. Parker, Ubiquitous mimo multifunction digital array radar, in: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, Vol. 1, IEEE, 2003, pp. 1057–1064.

[27] Y. Niu, Y. Li, D. Jin, L. Su, A. V. Vasilakos, A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges, Wireless networks 21 (8) (2015) 2657–2676.

[28] L. Zhang, W. Liu, J. Li, Low-complexity distributed beamforming for relay networks with real-valued implementation, IEEE Transactions on Signal Processing 61 (20) (2013) 5039–5048.

[29] L. Zhang, A. Ijaz, P. Xiao, M. M. Molu, R. Tafazolli, Filtered OFDM Systems, Algorithms, and Performance Analysis for 5G and Beyond, IEEE Transactions on Communications 66 (3) (2018) 1205–1218. doi:10.1109/TCOMM.2017.2771242.

[30] J. . van de Beek, O. Edfors, M. Sandell, S. K. Wilson, P. O. Borjesson, On Channel Estimation in OFDM systems, in: 1995 IEEE 45th Vehicular Technology Conference. Countdown to the Wireless Twenty-First Century, Vol. 2, 1995, pp. 815–819 vol.2. doi:10.1109/VETEC.1995.504981.

[31] P. Schniter, Low-Complexity Equalization of OFDM in Doubly Selective Channels, IEEE Transactions on Signal Processing 52 (4) (2004) 1002–1011. doi:10.1109/TSP.2004.823503.

[32] Qingwen Liu, Shengli Zhou, G. B. Giannakis, Cross-Layer Combining of Adaptive Modulation and Coding with Truncated ARQ Over Wireless Links, IEEE Transactions on Wireless Communications 3 (5) (2004) 1746–1755. doi:10.1109/TWC.2004.833474.

[33] S. Catreux, V. Erceg, D. Gesbert, R. W. Heath, Adaptive Modulation and MIMO Coding for Broadband Wireless Data Networks, IEEE Communications Magazine 40 (6) (2002) 108–115. doi:10.1109/MCOM.2002.1007416.

[34] A. Goldsmith, Wireless Communications, Cambridge university press, 2005.

[35] H. Artés, D. Seethaler, F. Hlawatsch, Efficient Detection Algorithms for MIMO Channels: A Geometrical Approach to Approximate ML Detection, IEEE transactions on signal processing 51 (11) (2003) 2808–2820.

[36] F. Gao, B. Wang, C. Xing, J. An, G. Y. Li, Wideband beamforming for hybrid massive mimo terahertz communications, IEEE Journal on Selected Areas in Communications 39 (6) (2021) 1725–1740. doi:10.1109/JSAC.2021.3071822.

[37] M. M. Molu, P. Xiao, M. Khalily, K. Cumanan, L. Zhang, R. Tafazolli, Low-complexity and robust hybrid beamforming design for multi-antenna communication systems, IEEE Transactions on Wireless Communications 17 (3) (2018) 1445–1459. doi:10.1109/TWC.2017.2778258.

[38] C. Kim, T. Kim, J.-Y. Seol, Multi-beam transmission diversity with hybrid beamforming for mimo-ofdm systems, in: 2013 IEEE Globecom Workshops (GC Wkshps), 2013, pp. 61–65. doi:10.1109/GLOCOMW.2013.6824962.

[39] S. Verdu, Spectral Efficiency in the Wideband Regime, IEEE Transactions on Information Theory 48 (6) (2002) 1319–1343. doi:10.1109/TIT.2002.1003824.

[40] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, R. W. Heath, Hybrid MIMO Architectures for Millimeter Wave Communications: Phase Shifters or Switches?, IEEE Access 4 (2016) 247–267. doi:10.1109/ACCESS.2015.2514261.

[41] S. V. Hum, J. Perruisseau-Carrier, Reconfigurable Reflectarrays and Array Lenses for Dynamic Antenna Beam Control: A Review, IEEE Transactions on Antennas and Propagation 62 (1) (2014) 183–198. doi:10.1109/TAP.2013.2287296.

[42] N. I. Zheludev, Y. S. Kivshar, From Metamaterials to Metadevices, Nature materials 11 (11) (2012) 917–924.

[43] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, C. Yuen, Reconfigurable Intelligent Surfaces for Energy Efficiency in Wireless Communication, IEEE Transactions on Wireless Communications 18 (8) (2019) 4157–4170.

[44] Q. Wu, R. Zhang, Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming, IEEE Transactions on Wireless Communications 18 (11) (2019) 5394–5409.

[45] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, et al., Smart Radio Environments Empowered by AI Reconfigurable Meta-surfaces: An Idea Whose Time Has Come, arXiv preprint arXiv:1903.08925 (2019).

[46] G. Memoli, M. Caleap, M. Asakawa, D. R. Sahoo, B. W. Drinkwater, S. Subramanian, Metamaterial bricks and quantization of meta-surfaces, Nature communications 8 (1) (2017) 1–8.

[47] G. Memoli, L. Chisari, J. P. Eccles, M. Caleap, B. W. Drinkwater, S. Subramanian, Vari-sound: A varifocal lens for sound, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 1–14. doi:10.1145/3290605.3300713.
URL `https://doi.org/10.1145/3290605.3300713`

[48] D. Lin, P. Fan, E. Hasman, M. L. Brongersma, Dielectric gradient metasurface optical elements, science 345 (6194) (2014) 298–302.

[49] L.-F. Chen, C. Ong, C. Neo, V. Varadan, V. K. Varadan, Microwave electronics: measurement and materials characterization, John Wiley & Sons, 2004.

[50] F. Costa, M. Borgese, Electromagnetic model of reflective intelligent surfaces, IEEE Open Journal of the Communications Society 2 (2021) 1577–1589. doi:10.1109/OJCOMS.2021.3092217.

[51] L. Zhang, X. Q. Chen, S. Liu, Q. Zhang, J. Zhao, J. Y. Dai, G. D. Bai, X. Wan, Q. Cheng, G. Castaldi, et al., Space-Time-Coding Digital Metasurfaces, Nature communications 9 (1) (2018) 4334.

[52] W. Saad, M. Bennis, M. Chen, A vision of 6g wireless systems: Applications, trends, technologies, and open research problems, IEEE network 34 (3) (2019) 134–142.

[53] T. Jiang, W. Yu, Interference nulling using reconfigurable intelligent surface, IEEE Journal on Selected Areas in Communications 40 (5) (2022) 1392–1406. doi:10.1109/JSAC.2022.3143220.

[54] A. Khalili, S. Zargari, Q. Wu, D. W. K. Ng, R. Zhang, Multi-objective resource allocation for irs-aided swipt, IEEE Wireless Communications Letters 10 (6) (2021) 1324–1328. doi:10.1109/LWC.2021.3065844.

[55] A. Wiesel, Y. C. Eldar, S. Shamai, Zero-Forcing Precoding and Generalized Inverses, IEEE Transactions on Signal Processing 56 (9) (2008) 4409–4418. doi:10.1109/TSP.2008.924638.

[56] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, R. Schober, Robust and Secure Wireless Communications via Intelligent Reflecting Surfaces, IEEE Journal on Selected Areas in Communications 38 (11) (2020) 2637–2652. arXiv:1912.01497, doi:10.1109/JSAC.2020.3007043.

[57] W. Tang, J. Y. Dai, M. Z. Chen, K.-K. Wong, X. Li, X. Zhao, S. Jin, Q. Cheng, T. J. Cui, Mimo transmission through reconfigurable intelligent surface: System design, analysis, and implementation, IEEE Journal on Selected Areas in Communications 38 (11) (2020) 2683–2699. doi:10.1109/JSAC.2020.3007055.

[58] H. Zhao, Y. Shuang, M. Wei, T. J. Cui, P. del Hougne, L. Li, Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals, Nature Communications 11 (1) (2020) 1–10. doi:10.1038/s41467-020-17808-y.
URL `http://dx.doi.org/10.1038/s41467-020-17808-y`

[59] M. H. Mazaheri, A. Chen, O. Abari, Millimeter Wave Backscatter: Toward Batteryless Wireless Networking at Gigabit Speeds, HotNets 2020 - Proceedings of the 19th ACM Workshop on Hot Topics in Networks (2020) 139–145doi:10.1145/3422604.3425948.

[60] M. H. Mazaheri, A. Chen, O. Abari, Millimeter Wave Backscatter: Toward Batteryless Wireless Networking at Gigabit Speeds, HotNets 2020 - Proceedings of the 19th ACM Workshop on Hot Topics in Networks (2020) 139–145doi:10.1145/3422604.3425948.

[61] Y. Rong, X. Tang, Y. Hua, A Unified Framework for Optimizing Linear Nonregenerative Multicarrier MIMO Relay Communication Systems, IEEE Transactions on Signal Processing 57 (12) (2009) 4837–4851. doi:10.1109/TSP.2009.2027779.

[62] Z. Ding, I. Krikidis, B. Rong, J. S. Thompson, C. Wang, S. Yang, On Combating The Half-Duplex Constraint in Modern Cooperative Networks: Protocols and Techniques, IEEE Wireless Communications 19 (6) (2012) 20–27.

[63] E. Björnson, Ö. Özdogan, E. G. Larsson, Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?, IEEE Wireless Communications Letters 9 (2) (2019) 244–248.

[64] W. Yan, X. Yuan, X. Kuai, Passive Beamforming and Information Transfer via Large Intelligent Surface, IEEE Wireless Communications Letters (2019).

[65] X. Yu, D. Xu, R. Schober, Miso wireless communication systems via intelligent reflecting surfaces : (invited paper), in: 2019 IEEE/CIC International Conference on Communications in China (ICCC), 2019, pp. 735–740. doi:10.1109/ICCChina.2019.8855810.

[66] N. S. Perović, M. D. Renzo, M. F. Flanagan, Channel Capacity Optimization Using Reconfigurable Intelligent Surfaces in Indoor mmWave Environments, in: ICC 2020 - 2020 IEEE International Conference on Communications (ICC), 2020, pp. 1–7. doi:10.1109/ICC40277.2020.9148781.

[67] D. Mishra, H. Johansson, Channel Estimation and Low-Complexity Beamforming Design for Passive Intelligent Surface Assisted MISO Wireless Energy Transfer, in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 4659–4663. doi:10.1109/ICASSP.2019.8683663.

[68] D. Dardari, N. Decarli, A. Guerra, F. Guidi, LOS/NLOS Near-Field Localization With a Large Reconfigurable Intelligent Surface, IEEE Transactions on Wireless Communications 21 (6) (2022) 4282–4294. doi:10.1109/TWC.2021.3128415.

[69] H. Taghvaee, S. Abadal, A. Pitilakis, O. Tsilipakos, A. C. Tasolamprou, C. Liaskos, M. Kafesaki, N. V. Kantartzis, A. Cabellos-Aparicio, E. Alarcón, Scalability analysis of programmable metasurfaces for beam steering, IEEE Access 8 (2020) 105320–105334. doi:10.1109/ACCESS.2020.3000424.

[70] H. Li, Q. Liu, Z. Wang, M. Li, Joint Antenna Selection and Analog Precoder Design with Low-Resolution Phase Shifters, IEEE Transactions on Vehicular Technology 68 (1) (2018) 967–971.

[71] Y. Shuang, H. Zhao, H. Li, M. Wei, L. Li, Dynamic Energy Allocation of Commodity Wi-Fi Signals Using Programmable Coding Metasurface, in: 2019 International

Applied Computational Electromagnetics Society Symposium - China (ACES), Vol. 1, 2019, pp. 1–2. doi:10.23919/ACES48530.2019.9060777.

[72] L.-H. Gao, Q. Cheng, J. Yang, S.-J. Ma, J. Zhao, S. Liu, H.-B. Chen, Q. He, W.-X. Jiang, H.-F. Ma, et al., Broadband Diffusion of Terahertz Waves by Multi-Bit Coding Metasurfaces, Light: Science & Applications 4 (9) (2015) e324.

[73] X. Wu, J. Yan, 3-D Mixed Far-Field and Near-Field Sources Localization With Cross Array, IEEE Transactions on Vehicular Technology 69 (6) (2020) 6833–6837. doi:10.1109/TVT.2020.2985903.

[74] A. Swindlehurst, T. Kailath, Passive Direction-of-Arrival and Range Estimation for Near-Field Sources, in: Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling, 1988, pp. 123–128. doi:10.1109/SPECT.1988.206176.

[75] R. A. Kennedy, T. Abhayapala, D. B. Ward, Broadband nearfield beamforming using a radial beampattern transformation, IEEE Transactions on Signal Processing 45 (1997) 287.

[76] V. Degli-Esposti, A Diffuse Scattering Model for Urban Propagation prediction, IEEE Transactions on Antennas and Propagation 49 (7) (2001) 1111–1113. doi:10.1109/8.933491.

[77] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, T. J. Cui, Wireless Communications With Reconfigurable Intelligent Surface: Path Loss Modeling and Experimental Measurement, IEEE Transactions on Wireless Communications 20 (1) (2021) 421–439. doi:10.1109/TWC.2020.3024887.

[78] E. Björnson, L. Sanguinetti, Power Scaling Laws and Near-Field Behaviors of Massive MIMO and Intelligent Reflecting Surfaces, IEEE Open Journal of the Communications Society 1 (2020) 1306–1324. doi:10.1109/OJCOMS.2020.3020925.

[79] E. Björnson, x. T. Demir, L. Sanguinetti, A Primer on Near-Field Beamforming for Arrays and Reconfigurable Intelligent Surfaces, in: 2021 55th Asilomar Conference on Signals, Systems, and Computers, 2021, pp. 105–112. doi:10.1109/IEEECONF53345.2021.9723331.

[80] S.-K. Chou, O. Yurduseven, H. Q. Ngo, M. Matthaiou, On the aperture efficiency of intelligent reflecting surfaces, IEEE Wireless Communications Letters 10 (3) (2021) 599–603. doi:10.1109/LWC.2020.3039483.

[81] K. Zetie, S. Adams, R. Tocknell, How does a mach-zehnder interferometer work?, Physics Education 35 (1) (2000) 46.

[82] Y. Liu, X. Mu, J. Xu, R. Schober, Y. Hao, H. V. Poor, L. Hanzo, STAR: Simultaneous transmission and reflection for 360° coverage by intelligent surfaces, IEEE Wireless Communications 28 (6) (2021) 102–109.

[83] X. Xie, C. He, H. Luan, Y. Dong, K. Yang, F. Gao, Z. J. Wang, A Joint Optimization Framework for IRS-assisted Energy Self-sustainable IoT Networks, IEEE Internet of Things Journal (2022) 1–1doi:10.1109/JIOT.2022.3142850.

[84] G. Zhou, C. Pan, H. Ren, K. Wang, A. Nallanathan, Intelligent Reflecting Surface Aided Multigroup Multicast MISO Communication Systems, IEEE Transactions on Signal Processing 68 (Xx) (2020) 3236–3251. arXiv:1909.04606, doi:10.1109/TSP.2020.2990098.

[85] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, M.-S. Alouini, Asymptotic Max-Min SINR Analysis of Reconfigurable Intelligent Surface Assisted MISO Systems, arXiv: Information Theory (2019).

[86] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, H. V. Poor, Hybrid Beamforming for Reconfigurable Intelligent Surface Based Multi-User Communications: Achievable Rates with Limited Discrete Phase Shifts, arXiv preprint arXiv:1910.14328 (2019).

[87] A. Papazafeiropoulos, C. Pan, P. Kourtessis, S. Chatzinotas, J. M. Senior, Intelligent reflecting surface-assisted mu-miso systems with imperfect hardware: Channel estimation and beamforming design, IEEE Transactions on Wireless Communications 21 (3) (2022) 2077–2092. doi:10.1109/TWC.2021.3109391.

[88] J. Chen, Y.-C. Liang, H. V. Cheng, W. Yu, Channel Estimation for Reconfigurable Intelligent Surface Aided Multi-User MIMO Systems, arXiv preprint arXiv:1912.03619 (2019).

[89] Z.-Q. He, X. Yuan, Cascaded Channel Estimation for Large Intelligent Metasurface Assisted Massive MIMO, IEEE Wireless Communications Letters 9 (2) (2020) 210–214. doi:10.1109/LWC.2019.2948632.

[90] C. You, B. Zheng, R. Zhang, Fast beam training for irs-assisted multiuser communications, IEEE Wireless Communications Letters 9 (11) (2020) 1845–1849. doi:10.1109/LWC.2020.3005980.

[91] X. Gao, Y. Liu, X. Liu, L. Song, Machine learning empowered resource allocation in irs aided miso-noma networks, IEEE Transactions on Wireless Communications 21 (5) (2022) 3478–3492. doi:10.1109/TWC.2021.3122409.

[92] X. Mu, Y. Liu, L. Guo, J. Lin, N. Al-Dhahir, Capacity and optimal resource allocation for irs-assisted multi-user communication systems, IEEE Transactions on Communications 69 (6) (2021) 3771–3786. doi:10.1109/TCOMM.2021.3062651.

[93] M. Najafi, V. Jamali, R. Schober, H. V. Poor, Physics-based modeling and scalable optimization of large intelligent reflecting surfaces, IEEE Transactions on Communications 69 (4) (2021) 2673–2691. doi:10.1109/TCOMM.2020.3047098.

[94] Z. Yang, W. Xu, C. Huang, J. Shi, M. Shikh-Bahaei, Beamforming design for multiuser transmission through reconfigurable intelligent surface, IEEE Transactions on Communications 69 (1) (2021) 589–601. doi:10.1109/TCOMM.2020.3028309.

[95] Z. Xie, W. Yi, X. Wu, Y. Liu, A. Nallanathan, Star-ris aided noma in multi-cell networks: A general analytical framework with gamma distributed channel modeling, IEEE Transactions on Communications (2022) 1–1doi:10.1109/TCOMM.2022.3186409.

[96] X. Mu, Y. Liu, L. Guo, J. Lin, R. Schober, Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications, IEEE Transactions on Wireless Communications 21 (5) (2021) 3083–3098.

[97] C. Wu, Y. Liu, X. Mu, X. Gu, O. A. Dobre, Coverage characterization of star-ris networks: Noma and oma, IEEE Communications Letters 25 (9) (2021) 3036–3040. doi:10.1109/LCOMM.2021.3091807.

[98] Z. Zhang, J. Chen, Y. Liu, Q. Wu, B. He, L. Yang, On the secrecy design of star-ris assisted uplink noma networks, IEEE Transactions on Wireless Communications (2022) 1–1doi:10.1109/TWC.2022.3190563.

[99] C. Wu, C. You, Y. Liu, X. Gu, Y. Cai, Channel estimation for star-ris-aided wireless communication, IEEE Communications Letters 26 (3) (2022) 652–656. doi:10.1109/LCOMM.2021.3139198.

[100] M. Ji, G. Caire, A. F. Molisch, Fundamental limits of caching in wireless d2d networks, IEEE Transactions on Information Theory 62 (2) (2016) 849–869. doi:10.1109/TIT.2015.2504556.

[101] W. Wang, L. Yang, A. Meng, Y. Zhan, D. W. K. Ng, Resource allocation for irs-aided jp-comp downlink cellular networks with underlaying d2d communications, IEEE Transactions on Wireless Communications 21 (6) (2022) 4295–4309. doi:10.1109/TWC.2021.3128711.

[102] G. C. Alexandropoulos, S. Samarakoon, M. Bennis, M. Debbah, Phase configuration learning in wireless networks with multiple reconfigurable intelligent surfaces, arXiv (2020) 1–6arXiv:2010.04376.

[103] G. Conway, L. Schott, A. Hirose, Measurement of Surface Reflection Coefficients via Multiple Reflection of Microwaves, Review of scientific instruments 65 (9) (1994) 2920–2928.

[104] A. Maltsev, R. Maslennikov, A. Sevastyanov, A. Khoryaev, A. Lomayev, Experimental Investigations of 60 GHz WLAN Systems in Office Environment, IEEE Journal on Selected Areas in Communications 27 (8) (2009) 1488–1499. doi:10.1109/JSAC.2009.091018.

[105] W. Tam, V. Tran, Propagation Modelling for Indoor Wireless Communication, Electronics & Communication Engineering Journal 7 (5) (1995) 221–228.

[106] W. Mei, R. Zhang, Cooperative Beam Routing for Multi-IRS Aided Communication, IEEE Wireless Communications Letters 10 (2) (2021) 426–430. arXiv:2010.13589, doi:10.1109/LWC.2020.3034370.

[107] S. Sun, M. Fu, Y. Shi, Y. Zhou, Towards Reconfigurable Intelligent Surfaces Powered Green Wireless Networks, arXiv (2020) 12–17arXiv:2005.01514.

[108] Z. Li, M. Hua, Q. Wang, Q. Song, Weighted sum-rate maximization for multi-irs aided cooperative transmission, IEEE Wireless Communications Letters 9 (10) (2020) 1620–1624. doi:10.1109/LWC.2020.2999356.

[109] M. A. Kishk, M. S. Alouini, Exploiting Randomly Located Blockages for Large-Scale Deployment of Intelligent Surfaces, IEEE Journal on Selected Areas in Communications 39 (4) (2021) 1043–1056. arXiv:2001.10766, doi:10.1109/JSAC.2020.3018808.

[110] W. Mei, R. Zhang, Performance Analysis and User Association Optimization for Wireless Network Aided by Multiple Intelligent Reflecting Surfaces, IEEE Transactions on Communications 69 (9) (2021) 6296–6312. doi:10.1109/TCOMM.2021.3087620.

[111] Z. Zhang, L. Dai, A Joint Precoding Framework for Wideband Reconfigurable Intelligent Surface-Aided Cell-Free Network, IEEE Transactions on Signal Processing 69 (2021) 4085–4101. doi:10.1109/TSP.2021.3088755.

[112] S. Huang, Y. Ye, M. Xiao, H. V. Poor, M. Skoglund, Decentralized Beamforming Design for Intelligent Reflecting Surface-Enhanced Cell-Free Networks, IEEE Wireless Communications Letters 10 (3) (2021) 673–677. doi:10.1109/LWC.2020.3045884.

[113] X. Hu, C. Zhong, Y. Zhang, X. Chen, Z. Zhang, Location information aided multiple intelligent reflecting surface systems, IEEE Transactions on Communications 68 (12) (2020) 7948–7962.

[114] W. Mei, B. Zheng, C. You, R. Zhang, Intelligent Reflecting Surface Aided Wireless Networks: From Single-reflection to Multi-reflection Design and Optimization, arXiv preprint arXiv:2109.13641 (2021).

[115] Q. Wan, J. Fang, Z. Chen, H. Li, Hybrid Precoding and Combining for Millimeter Wave/Sub-THz MIMO-OFDM Systems with Beam Squint Effects, IEEE Transactions on Vehicular Technology PP (c) (2021) 1–1. doi:10.1109/tvt.2021.3093095.

[116] T. Griesser, C. A. Balanis, K. Liu, RCS Analysis and Reduction for Lossy Dihedral Corner Reflectors, Proceedings of the IEEE 77 (5) (1989) 806–814. doi:10.1109/5.32071.

[117] A. Y. Modi, M. A. Alyahya, C. A. Balanis, C. R. Birtcher, Metasurface-Based Method for Broadband RCS Reduction of Dihedral Corner Reflectors With Multiple Bounces, IEEE Transactions on Antennas and Propagation 68 (3) (2020) 1436–1447. doi:10.1109/TAP.2019.2940494.

[118] H. Yan, H.-C. Yin, S. Li, L.-S. Li, 3-D Rotation Representation of Multiple Reflections and Parametric Model for Bistatic Scattering From Arbitrary Multiplate Structure, IEEE Transactions on Antennas and Propagation 67 (7) (2019) 4777–4791. doi:10.1109/TAP.2019.2911268.

[119] P. Cao, J. S. Thompson, H. Haas, Constant Modulus Shaped Beam Synthesis Via Convex Relaxation, IEEE Antennas and Wireless Propagation Letters 16 (2017) 617–620. doi:10.1109/LAWP.2016.2594141.

[120] X. Gao, L. Dai, S. Han, I. Chih-Lin, R. W. Heath, Energy-Efficient Hybrid Analog and Digital Precoding for MmWave MIMO Systems With Large Antenna Ar-

rays, IEEE Journal on Selected Areas in Communications 34 (4) (2016) 998–1009. arXiv:1507.04592, doi:10.1109/JSAC.2016.2549418.

[121] F. Sohrabi, W. Yu, Hybrid Analog and Digital Beamforming for mmWave OFDM Large-Scale Antenna Arrays, IEEE Journal on Selected Areas in Communications 35 (7) (2017) 1432–1443. doi:10.1109/JSAC.2017.2698958.

[122] O. Ozdogan, E. Bjornson, E. G. Larsson, Intelligent Reflecting Surfaces: Physics, Propagation, and Pathloss Modeling, IEEE Wireless Communications Letters 9 (5) (2020) 581–585. arXiv:1911.03359, doi:10.1109/LWC.2019.2960779.

[123] E. Björnson, L. Sanguinetti, Power Scaling Laws and Near-field Behaviors of Massive MIMO and Intelligent Reflecting Surfaces, IEEE Open Journal of the Communications Society 1 (2020) 1306–1324. doi:10.1109/OJCOMS.2020.3020925.

[124] W. Cai, H. Li, M. Li, Q. Liu, Practical Modeling and Beamforming for Intelligent Reflecting Surface Aided Wideband Systems, IEEE Communications Letters 24 (7) (2020) 1568–1571. arXiv:2006.01465, doi:10.1109/LCOMM.2020.2987322.

[125] E. Björnson, Optimizing a Binary Intelligent Reflecting Surface for OFDM Communications under Mutual Coupling 1 (3) (2021). arXiv:2106.04280.
URL `http://arxiv.org/abs/2106.04280`

[126] H. Li, W. Cai, Y. Liu, M. Li, Q. Liu, Q. Wu, Intelligent Reflecting Surface Enhanced Wideband MIMO-OFDM Communications: From Practical Model to Reflection Optimization, IEEE Transactions on Communications 69 (7) (2021) 4807–4820. arXiv:2007.14243, doi:10.1109/TCOMM.2021.3069860.

[127] S. Abeywickrama, R. Zhang, Q. Wu, C. Yuen, Intelligent Reflecting Surface: Practical Phase Shift Model and Beamforming Optimization, IEEE Transactions on Communications 68 (9) (2020) 5849–5863. arXiv:2002.10112, doi:10.1109/TCOMM.2020.3001125.

[128] L. Zhang, Z. X. Wang, R. W. Shao, J. L. Shen, X. Q. Chen, X. Wan, Q. Cheng, T. J. Cui, Dynamically realizing arbitrary multi-bit programmable phases using a 2-bit time-domain coding metasurface, IEEE Transactions on Antennas and Propagation 68 (4) (2019) 2984–2992.

[129] Y. Cao, T. Lv, W. Ni, Intelligent reflecting surface aided multi-user mmWave communications for coverage enhancement, IEEE International Symposium on Per-

sonal, Indoor and Mobile Radio Communications, PIMRC 2020-Augus (2020). arXiv:1910.02398, doi:10.1109/PIMRC48278.2020.9217160.

[130] E. Bjornson, L. Sanguinetti, Rayleigh Fading Modeling and Channel Hardening for Reconfigurable Intelligent Surfaces, IEEE Wireless Communications Letters (2020) 1–6arXiv:2009.04723, doi:10.1109/LWC.2020.3046107.

[131] T. Bai, R. W. Heath, Coverage and Rate Analysis for Millimeter-Wave Cellular Networks, IEEE Transactions on Wireless Communications 14 (2) (2015) 1100–1114. doi:10.1109/TWC.2014.2364267.

[132] R. Ord-Smith, Generation of Permutation Sequences: Part 1, The Computer Journal 13 (2) (1970) 152–155.

[133] L. Zhang, W. Liu, R. J. Langley, A Class of Constrained Adaptive Beamforming Algorithms Based On Uniform Linear Arrays, IEEE Transactions on Signal Processing 58 (7) (2010) 3916–3922.

[134] P. T. Boggs, J. W. Tolle, Sequential Quadratic Programming for Large-scale Nonlinear Optimization, Journal of Computational and Applied Mathematics 124 (1) (2000) 123 – 137, numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations. doi:https://doi.org/10.1016/S0377-0427(00)00429-5.

[135] P. Cao, J. S. Thompson, H. Haas, Constant modulus shaped beam synthesis via convex relaxation, IEEE Antennas and Wireless Propagation Letters 16 (2017) 617–620. doi:10.1109/LAWP.2016.2594141.

[136] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, Q. Cheng, Coding Metamaterials, Digital Metamaterials and Programmable Metamaterials, Light: Science & Applications 3 (10) (2014) e218.

[137] Y. M. Tsang, A. S. Poon, S. Addepalli, Coding The Beams: Improving Beamforming Training in Mmwave Communication System, in: 2011 IEEE Global Telecommunications Conference-GLOBECOM 2011, IEEE, 2011, pp. 1–6.

[138] H. Han, Y. Liu, L. Zhang, On half-power beamwidth of intelligent reflecting surface, IEEE Communications Letters (2020) 1–1doi:10.1109/LCOMM.2020.3046369.

[139] H. Guo, B. Makki, T. Svensson, A genetic algorithm-based beamforming approach for delay-constrained networks, in: 2017 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2017, pp. 1–7. doi:10.23919/WIOPT.2017.7959905.

[140] G. C. Trichopoulos, P. Theofanopoulos, B. Kashyap, A. Shekhawat, A. Modi, T. Osman, S. Kumar, A. Sengar, A. Chang, A. Alkhateeb, Design and evaluation of reconfigurable intelligent surfaces in real-world environment, IEEE Open Journal of the Communications Society 3 (2022) 462–474. doi:10.1109/OJCOMS.2022.3158310.

[141] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, 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 38 (8) (2020) 1809–1822. doi:10.1109/JSAC.2020.3000813.

[142] Q. Wu, R. Zhang, Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts, IEEE Transactions on Communications 68 (3) (2020) 1838–1851. doi:10.1109/TCOMM.2019.2958916.

[143] Y. Liu, L. Zhang, M. A. Imran, Multi-user Beamforming and Transmission Based on Intelligent Reflecting Surface, IEEE Transactions on Wireless Communications (2022) 1–1doi:10.1109/TWC.2022.3157808.

[144] Y. Liu, L. Zhang, B. Yang, W. Guo, M. A. Imran, Programmable wireless channel for multi-user mimo transmission using meta-surface, in: 2019 IEEE Global Communications Conference (GLOBECOM), 2019, pp. 1–6. doi:10.1109/GLOBECOM38437.2019.9013433.

[145] E. Björnson, Optimizing A Binary Intelligent Reflecting Surface for OFDM Communications under Mutual Coupling 1 (3) (2021). arXiv:2106.04280.
URL `http://arxiv.org/abs/2106.04280`

[146] K. Jain, J. Padhye, V. N. Padmanabhan, L. Qiu, Impact of Interference on Multihop Wireless Network Performance, Wireless Networks 11 (4) (2005) 471–487. doi:10.1007/s11276-005-1769-9.

[147] D. Dor, M. Tarsi, Graph Decomposition Is NP-Complete: A Complete Proof of Holyer's Conjecture, SIAM Journal on Computing 26 (4) (1997) 1166–1187.

[148] M. Tarsi, Decomposition of a complete multigraph into simple paths: nonbalanced handcuffed designs, Journal of Combinatorial Theory, Series A 34 (1) (1983) 60–70.

[149] Z. Wang, L. Liu, S. Cui, Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis, IEEE Transactions on Wireless Communications 19 (10) (2020) 6607–6620. doi:10.1109/TWC.2020.3004330.

[150] Q. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, M. Alouini, Intelligent reflecting surface-assisted multi-user miso communication: Channel estimation and beamforming design, IEEE Open Journal of the Communications Society 1 (2020) 661–680. doi:10.1109/OJCOMS.2020.2992791.

[151] S. Guo, S. Lv, H. Zhang, J. Ye, P. Zhang, Reflecting Modulation, IEEE Journal on Selected Areas in Communications 38 (11) (2020) 2548–2561. doi:10.1109/JSAC.2020.3007060.

[152] D. Mishra, E. G. Larsson, Passive Intelligent Surface Assisted MIMO Powered Sustainable IoT, in: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 8961–8965. doi:10.1109/ICASSP40776.2020.9053628.

[153] Q. Wu, R. Zhang, Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network, IEEE Communications Magazine 58 (1) (2020) 106–112. doi:10.1109/MCOM.001.1900107.

[154] Z. Yang, Y. Liu, X.-Y. Li, Beyond Trilateration: On the Localizability of Wireless Ad-hoc Networks, in: IEEE INFOCOM 2009, 2009, pp. 2392–2400. doi:10.1109/INFCOM.2009.5062166.

[155] S. Brahmbhatt, J. Gu, K. Kim, J. Hays, J. Kautz, Geometry-aware learning of maps for camera localization, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 2616–2625.

[156] G. Gradoni, M. Di Renzo, End-to-End Mutual Coupling Aware Communication Model for Reconfigurable Intelligent Surfaces: An Electromagnetic-Compliant Approach Based on Mutual Impedances, IEEE Wireless Communications Letters 10 (5) (2021) 938–942. arXiv:2009.02694, doi:10.1109/LWC.2021.3050826.

[157] T. Su, H. Ling, On modeling mutual coupling in antenna arrays using the coupling matrix, Microwave and Optical Technology Letters 28 (4) (2001) 231–237. doi:10.1002/1098-2760(20010220)28:4<231::AID-MOP1004>3.0.CO;2-P.

[158] X. Ge, R. Zi, H. Wang, J. Zhang, M. Jo, Multi-User Massive MIMO Communication Systems Based on Irregular Antenna Arrays, IEEE Transactions on Wireless Communications 15 (8) (2016) 5287–5301. arXiv:1604.04968, doi:10.1109/TWC.2016.2555911.

[159] T. Svantesson, A. Ranheim, Mutual coupling effects on the capacity of multielement antenna systems, ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings 4 (2001) 2485–2488. doi:10.1109/icassp.2001.940505.

[160] Y. Wang, H. Lu, H. Sun, Channel Estimation in IRS-Enhanced mmWave System With Super-Resolution Network 25 (8) (2021) 2599–2603.

[161] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, K. Chetty, A Survey on Fundamental Limits of Integrated Sensing and Communication, IEEE Communications Surveys and Tutorials 24 (2) (2022) 994–1034. doi:10.1109/COMST.2022.3149272.

[162] A. Hessel, J. Schmoys, D. Y. Tseng, Bragg-angle blazing of diffraction gratings, JOURNAL OF THE OPTICAL SOCIETY OF AMERICA 65 (April) (1975) 5.