

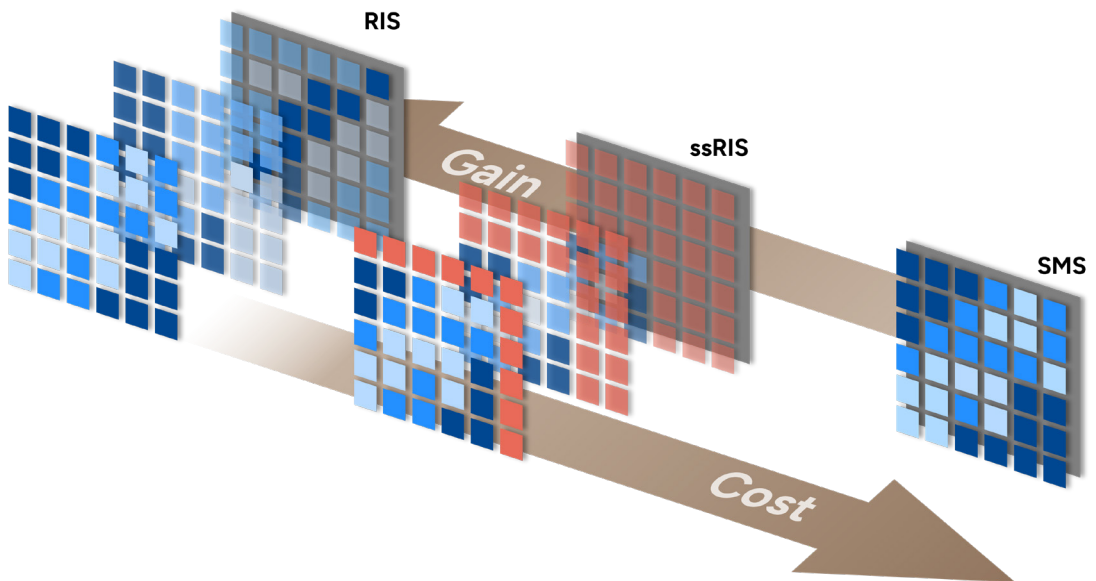


Licentiate Thesis in Information and Communication Technology

# Sustainable Metasurface-Assisted Indoor Wireless Communication System Design

ZHENYU LI

KTH ROYAL INSTITUTE OF TECHNOLOGY



# Sustainable Metasurface-Assisted Indoor Wireless Communication System Design

ZHENYU LI

Academic Dissertation which, with due permission of the KTH Royal Institute of Technology, is submitted for public defence for the Degree of Licentiate of Engineering on Friday the 17th of April 2026, at 10:00 a.m. in Harry Nyquist, Malvinas Väg 10, Kungliga Tekniska Högskolan, Stockholm.

Licentiate Thesis in Information and Communication Technology  
KTH Royal Institute of Technology  
Stockholm, Sweden 2026

© Zhenyu Li

TRITA-EECS-AVL-2026:26  
ISBN 978-91-8106-569-5

Printed by: Universitetservice US-AB, Sweden 2026

# Abstract

The densification of wireless networks toward fifth- and sixth-generation standards has intensified the demand for reliable high-throughput connectivity in Indoor Dense Spaces (IDS), such as aircraft cabins, metro wagons, and stadiums. Although Millimetre-Wave (mmWave) communication offers the spectral resources needed to meet this demand, its sensitivity to propagation loss and blockages severely limits its performance, particularly in IDS. Metasurfaces have emerged as a promising means of extending mmWave coverage through manipulating the propagation environment. Advanced investigations have been conducted on metasurface-featured system performance enhancement. However, the operating cost, which is a practical and critical concern of metasurface deployment, has received insufficient attention in the literature. Deploying a reconfigurable metasurface in practice requires cabling, power supply, and control infrastructure, costs that represent a real barrier to scalable deployment, particularly in indoor environments like IDS, where infrastructure installation is physically limited or tightly regulated.

This thesis investigates the design of sustainable metasurface-assisted indoor wireless communication systems, placing operating cost alongside performance as a primary design criterion. The work examines different types of metasurfaces that differ in the metasurface gain they provide and the operating cost they incur. By identifying and verifying an optimal design choice among these alternatives, this thesis advances a sustainable metasurface-assisted system that addresses the performance-cost dilemma inherent to IDS deployments.

The first contribution studies the trade-off between operating cost and performance enhancement by optimizing a mixed Static Metasurface (SMS) and Reconfigurable Intelligent Surface (RIS) deployment in an mmWave IDS. Using a Feasible Point Pursuit Successive Convex Approximation (FPP-SCA)-based iterative algorithm, the results reveal a diminishing-returns relationship. While replacing two SMSs with RISs already yields a 13 Mbps gain, increasing the RIS count beyond 16 out of 22 surfaces produces less than 1 Mbps of additional gain, confirming that full reconfigurability is unnecessary

and motivating a more cost-effective middle-ground solution. The second contribution proposes and evaluates a Self-Sustainable RIS (ssRIS)-assisted mmWave system for IDS, where ssRIS achieves self-sustainability through power harvesting via a codebook-based element splitting scheme, eliminating the need for cabling and external power. A two-stage iterative algorithm jointly optimizes phase shifts, User Equipment (UE)-to-ssRIS associations, and time allocation. The results show that ssRIS outperforms SMS by up to 19.8 Mbps in compact environments, confirming a favorable position within the gain-cost trade-off, with coverage advantages diminishing as deployment distances grow. The third contribution conducts a feasibility study of ssRIS across diverse scenarios, analyzing how element count scales with transmit power, data rate demands, and outage constraints under Element Split (ES) and Time Split (TS) schemes. TS benefits from stronger channel hardening under moderate conditions, but scales exponentially with harvesting difficulty, whereas ES scales only linearly, offering greater robustness in challenging environments. Together, these findings provide actionable guidance for practical ssRIS deployment.

### **Keywords**

Millimeter wave communication, indoor dense space, reconfigurable intelligent surfaces, static metasurfaces, self-sustainable reconfigurable intelligent surfaces, energy harvesting, operating cost, successive convex approximation.

# Sammanfattning

Förtätningen av trådlösa nätverk mot femte och sjättegenerationens standarder har intensifierat behovet av tillförlitlig höghastighetskommunikation i inomhusmiljöer med hög användartäthet (IDS), såsom flygplanskabiner, tunnelbanevagnar och arenor. Även om millimetervågskommunikation (mmWave) erbjuder de spektralresurser som krävs för att möta denna efterfrågan, begränsar dess känslighet för utbredningsförluster och blockeringar dess prestanda avsevärt, särskilt i IDS. Metaytor har framträtt som ett lovande verktyg för att utöka mmWave-täckning genom att manipulera utbredningsomgivningen. Avancerade undersökningar har genomförts avseende prestandaförbättring i metayta-baserade system. Driftskostnaden, som utgör ett praktiskt och kritiskt problem vid driftsättning av metaytor, har dock fått otillräcklig uppmärksamhet i litteraturen. Att i praktiken implementera omkonfigurerbara metaytor kräver kablage, strömförsörjning och styrsystem, vilket medför kostnader som utgör ett reellt hinder för skalbar implementering, särskilt i miljöer där fysiska begränsningar eller stränga regleringskrav försvårar infrastrukturinstallation.

Detta avhandlingsarbete undersöker utformningen av hållbara metayta-assisterade trådlösa kommunikationssystem inomhus, där driftskostnad jämföras med prestanda som ett primärt designkriterium. Arbetet undersöker olika typer av metaytor som skiljer sig åt i den metayteförstärkning de erbjuder och den driftskostnad de medför. Genom att identifiera och verifiera ett optimalt designval bland dessa alternativ bidrar denna avhandling till ett hållbart metayta-assisterat system som hanterar prestandakostnadsdilemmat som är inneboende i IDS-miljöer.

Det första bidraget studerar avvägningen mellan den driftskostnad som följer med rekonfigurerbarheten och motsvarande prestandaförbättring, genom att optimera en blandad driftsättning av statisk metayta (SMS) och rekonfigurerbar intelligent yta (RIS) i ett mmWave IDS. Med hjälp av en algoritm baserad på successiv konvex approximation med genomförbar punktsökning (FPP-SCA) påvisar resultaten ett avtagande avkastningsförhållande. Redan ersättningen av två SMS med RIS ger en vinst på 13 Mbps, men att öka antalet RIS utöver 16 av 22 ytor ger mindre än 1 Mbps ytterligare

vinst, vilket bekräftar att full rekonfigurabilitet är onödig och motiverar sökandet efter en mer kostnadseffektiv mellanlösning. Det andra bidraget föreslår och utvärderar ett självhållbar rekonfigurerbar intelligent yta (ssRIS)-assisterat mmWave-system för IDS, där ssRIS uppnår självhållbarhet genom energiinsamling via ett kodbruksbaserat elementdelningsschema, vilket eliminerar behovet av kablage och extern strömförsörjning. En tvåstegs iterativ algoritm optimerar gemensamt fasskift, användarutrustning (UE)-till-ssRIS-associeringar och tidsallokering. Resultaten visar att ssRIS överträffar SMS med upp till 19,8 Mbps i kompakta miljöer, vilket bekräftar en fördelaktig position inom prestandakostnadsavvägningen, medan täckningsfördelen minskar med ökande driftsättningsavstånd. Det tredje bidraget genomför en genomförbarhetsstudie av ssRIS i varierande scenarier, och analyserar hur elementantalet skalas med sändeffekt, datahastighetsrerv och avbrottsbegränsningar under elementdelning (ES)- och tidsdelning (TS)-scheman. TS gynnas av starkare kanalhårdning under måttliga förhållanden, men dess elementantal växer exponentiellt med insamlingssvårigheten, medan ES endast skalar linjärt, vilket ger större robusthet i utmanande miljöer. Sammantaget ger dessa resultat handlingsbara riktlinjer för praktisk driftsättning av ssRIS.

### **Nyckelord**

Millimetervågskommunikation, inomhusmiljöer med hög användartäthet, rekonfigurerbara intelligenta ytor, statiska metaytor, självhållbara rekonfigurerbara intelligenta ytor, energiinsamling, driftskostnad, successiv konvex approximation.

# Acknowledgment

The journey of my PhD feels like a road trip. I packed my ambitions like luggage I was not sure I could carry and set off the journey in 2023. There are days when I encounter problems that ignite something inside me, ideas bloom like spring flowers, and the world turns vivid. Interesting questions form naturally in my mind, and every conversation, every paper, every scribble on a whiteboard feels like it belongs to something larger. Yet spring doesn't hold long. There are also times when I feel like walking a narrow road through a dense summer forest, where the canopy closes overhead, and the end of the path is nowhere in sight. I would find myself one formula away from a solution, yet days and nights would pass, and the answer would not come. The work continued, but the light did not. Winter comes without announcement. A typo in a simulation script, quietly sitting there, shaping every result I had trusted. When I found it and fixed it, the numbers told a completely different story. I stood at the top of a snowy hill, the road I had come from buried under fresh snow, uncertain whether to step forward into the unknown or retrace steps that no longer existed. But there is also autumn. The acceptance notification, the moment when the results finally hold, when months of effort reveal themselves as something real and complete, like a harvest after a long growing season. Those moments do not erase the other seasons. They make them worth having.

This thesis would not exist without the people who walked different parts of this road with me. I would like to express my sincere gratitude to my main supervisor, Cicek Cavdar, and my co-supervisors, Emil Björnson and Özlem Tuğfe Demir, for their guidance, patience, and insight throughout this journey. I am also deeply grateful to my colleague and collaborator Ozan Alp Topal, whose partnership has shaped this work in ways that are difficult to fully account for.

I am sincerely grateful to Prof. Diana Pamela Moya Osorio for accepting the invitation to serve as opponent at this defence, and to Prof. Saikat Chatterjee for reviewing this thesis and offering feedback that made it sharper and stronger. My appreciation also goes to Ki Won Sung for serving as the examiner.

To my colleagues at KTH, Satya, Amna, Zinat, Irshad, Eren, Qichen, Peng, Zhe, Nasir, and everyone else in the COS department who shared an office, a lunch, a complaint, or a small celebration. The daily texture of this journey was made by you.

To Jiajie, thank you for your steady presence across all four seasons.

And lastly, to my family, all my honors belong to you.

长风破浪会有时，直挂云帆济沧海。

Sincerely,

Zhenyu Li

Stockholm, March 16, 2026

## List of included papers

**Contributions** This licentiate thesis consists of two parts. The first part provides an overview of the research field on which I focused and a summary of my contributions to it. The second part comprises the contributions that I have made to this research field through published and under-review works.

- Paper A** **Mixed Static and Reconfigurable Metasurface Deployment in Indoor Dense Spaces: How Much Reconfigurability is Needed?**, Zhenyu Li, Ozan Alp Topal, Özlem Tuğfe Demir, Emil Björnson, and Cicek Cavdar. In 2024 IEEE Wireless Communications and Networking Conference (WCNC), 2024
- Paper B** **Self-Sustainable Metasurface-Assisted mmWave Indoor Communication System**, Zhenyu Li, Ozan Alp Topal, Özlem Tuğfe Demir, Emil Björnson, and Cicek Cavdar. In the process of reviewing, submitted to IEEE Transactions on Wireless Communications, 2026
- Paper C** **Feasibility Study Regarding Self-Sustainable Reconfigurable Intelligent Surfaces**, Zhenyu Li, Ozan Alp Topal, Özlem Tuğfe Demir, Emil Björnson, and Cicek Cavdar. In IEEE Wireless Communications Letters, 2026



# Contents

<b>Abstract</b> . . . . .	v
<b>Sammanfattning</b> . . . . .	vii
<b>Acknowledgment</b> . . . . .	ix
<b>List of included papers</b> . . . . .	xi
<b>Contents</b> . . . . .	xiii
<b>List of Figures</b> . . . . .	xv
<b>List of Tables</b> . . . . .	xvii
<b>Acronyms</b> . . . . .	xix
<b>1 Introduction</b> . . . . .	1
1.1 Motivation . . . . .	1
1.2 Research Question . . . . .	2
1.3 Thesis Contributions . . . . .	3
1.4 Societal Value of the Thesis . . . . .	4
1.5 Organization . . . . .	4
<b>2 Background</b> . . . . .	7
2.1 Indoor Dense Spaces . . . . .	7
2.2 Millimeter-Wave . . . . .	8
2.3 Metasurfaces . . . . .	8
2.4 Metasurface-assisted mmWave IDS system modeling . . . . .	11
2.5 Successive Convex Approximation . . . . .	13
<b>3 Trade-off between operating cost and metasurface gain</b> . . . . .	17
3.1 Related Work . . . . .	18
3.2 Contributions . . . . .	18
3.3 System Model . . . . .	19
3.4 Problem Formulation . . . . .	21
3.5 Results . . . . .	24
3.6 Conclusion . . . . .	26
<b>4 Self-sustainable RIS, a balanced compromise within the trade-off</b> . . . . .	27
4.1 Related Work . . . . .	28

## CONTENTS

4.2	Contributions . . . . .	28
4.3	System Model . . . . .	29
4.4	Problem Formulation . . . . .	32
4.5	Selective Results . . . . .	39
4.6	Conclusion . . . . .	45
<b>5</b>	<b>Self-sustainable RIS feasibility analysis . . . . .</b>	<b>47</b>
5.1	Related Work . . . . .	48
5.2	Contributions . . . . .	48
5.3	System Model . . . . .	48
5.4	Problem Formulation . . . . .	52
5.5	Results . . . . .	54
5.6	Conclusion . . . . .	56
<b>6</b>	<b>Conclusions and future work . . . . .</b>	<b>57</b>
6.1	Concluding remarks . . . . .	57
6.2	Future Research Directions . . . . .	58
	<b>Bibliography . . . . .</b>	<b>61</b>

# List of Figures

2.2.1	The dilemma of the mmWave wireless communication system in IDS . . .	8
2.3.2	Illustration of the ES and TS ssRIS . . . . .	10
2.4.3	Illustration of the metasurface model in RT simulator . . . . .	12
2.4.4	The front view and the top view of the considered (partial) cabin. . . . .	12
3.3.1	Illustration of the RIS/SMS-assisted system model. . . . .	20
3.5.2	Objective value of the proposed algorithm with respect to each iteration	24
3.5.3	(a) Optimized minimum data rate under different $L_{\max}$ . (b) To reach the same minimum data rate when using 22 SMSs, at least 18 RISs are needed. . . . .	25
3.5.4	(a) Optimized SNR $\Gamma_k^*$ , (b) optimized allocated time portion $\tau_k^*$ , in the case where $L_{\max} = 8$ . The red boxes represent the location of the deployed RISs and the purple box represents the location of the BS. . . . .	25
4.3.1	Illustration of the ssRIS-assisted system model. . . . .	30
4.3.2	Illustration of the ssRIS with element splitting scheme and element working mode selected based on presets. . . . .	30
4.4.3	Flowchart of the two-stage iterative data rate optimization algorithm. . . . .	38
4.5.4	The cascaded BS-ssRIS-UE channel gain with respect to a ssRIS that is (a) far-away (b) near to the BS. The red rectangle represents the checked ssRIS, and the pink circle represents the BS. . . . .	39
4.5.5	Illustration of different coverage group designs. Coverage group consists of (a) 6 ssRISs, 6 UEs; (b) 4 ssRISs, 6 UEs; (c) 2 ssRISs, 6 UEs. . . . .	40
4.5.6	The upper bound of the achievable SNR under different scales of the coverage group. . . . .	40
4.5.7	(a) Averaged received power, harvested power, and consumed power under the searched-out reflecting area size with randomized phase shifts; (b) averaged searched-out reflecting area size and the distance between the ssRIS and the BS. . . . .	41
4.5.8	Convergence condition of applying Algorithm 3 concerning the inputs from coverage groups 1, 10, and 30. . . . .	42
4.5.9	(a) BS contribution for each UE; (b) ssRIS-UE association condition. . . . .	43
4.5.10	Self-sustainability condition of (a) ssRIS 12, close to the BS, $\bar{M}_{12} = 64$ ; (b) ssRIS 29, far from the BS, $\bar{M}_{29} = 169$ . . . . .	43
4.5.11	(a) Time allocation and the resulting SNR of the whole cabin; (b) intra-group time allocation and the optimized SNR of the coverage group that consists of UEs 13-18. . . . .	44
4.5.12	Illustrations of metasurface performance test. (a) Setups for different types of metasurface; (b) considered cabins of different sizes and their relative position. . . . .	45
4.5.13	Optimized data rate under the assistance of different types of metasurface and under different cell sizes. . . . .	45

LIST OF FIGURES

4.6.14 General comparison in terms of operating cost, metasurface gain, and system coverage among ssRIS, SMS, and RIS. . . . . 46

5.3.1 System model. The square and circle represent the position of the BS and the UE, respectively. . . . . 49

5.5.2 Element requirement across different system parameters. LoS and NLoS refer to the propagation conditions of the ssRIS-UE channel. (a) Impact of harvesting condition (via changing transmit power  $P$ ) with  $R_0 = 10$  Mbps; (b) Impact of data rate  $R_0$  with  $P = 0.1$  W. . . . . 55

5.5.3 Element requirement across different outage margins under the NLoS ssRIS-UE channel. (a)  $P = 0.1$  W and  $R_0 = 20$  Mbps; (b)  $P = 0.1$  W and  $R_0 = 15$  Mbps. . . . . 56

# List of Tables

2.4.1 Dielectric properties of materials at 28 GHz considered in the simulation. 13



# Acronyms

- ABS** Acrylonitrile Butadiene Styrene 12  
**AWGN** Additive White Gaussian Noise 49  
**BS** Base Station xv, xvi, 12, 18–20, 25, 26, 29–31, 39, 41–46, 48–54  
**CCCP** Concave-Convex Procedure 13  
**CDF** Cumulative Distribution Function 51  
**CLT** Central Limit Theorem 51  
**ES** Element Split vi, viii, xv, 9, 10, 47, 48, 51–56  
**FPP-SCA** Feasible Point Pursuit Successive Convex Approximation v, vii, 3, 13–15, 19, 23, 28, 36, 37, 41, 46  
**FSPL** Free-Space Path Loss 49  
**HaR** Harvesting-and-Reflecting 2, 5, 9, 10, 28, 29, 31, 47, 48, 56  
**IDS** Indoor Dense Spaces v–viii, 1–5, 7–9, 11, 17–19, 21, 24, 26–29, 32, 33, 38, 39, 43, 46, 47  
**IoT** Internet of Things 56  
**KKT** Karush-Kuhn-Tucker 14, 23  
**LoS** Line-of-Sight xvi, 4, 11, 30, 46, 48–56  
**LP** Linear Programming 34  
**MIP** Mixed Integer Programming 3, 19, 23, 28, 32, 33, 37  
**mmWave** Millimetre-Wave v–viii, 1–3, 5, 8, 9, 11, 18, 19, 26, 27, 29, 30, 33, 45  
**MRT** Maximum Ratio Transmission 20, 31, 33, 34, 41–43, 48, 50  
**NLoS** Non-Line-of-Sight xvi, 4, 48, 51, 53–56  
**PS** Power Splitting 51  
**PSD** Positive Semi-Definite 14, 36  
**QCQP** Quadratically Constrained Quadratic Program 14, 15  
**QoS** Quality of Service 47, 48  
**RIS** Reconfigurable Intelligent Surface v, vii, xv, xvi, 2, 3, 5, 9, 12, 17–29, 44–46, 48  
**RT** Ray Tracing xv, 7, 11, 12, 24, 26, 29  
**SCA** Successive Convex Approximation 4, 7, 13, 14  
**SDG** Sustainable Development Goals 4  
**SMS** Static Metasurface v–viii, xv, xvi, 2, 3, 5, 9, 17–22, 24–29, 44–46  
**SNR** Signal-to-Noise Ratio xv, 12, 20, 22, 25, 26, 28, 31, 33, 35, 40, 42–44, 48, 50, 51, 53, 54  
**SOC** Second-Order Cone 22, 35  
**ssRIS** Self-Sustainable RIS vi, viii, xv, xvi, 2–5, 7, 9–11, 14, 18, 26–33, 35, 38–56  
**TS** Time Split vi, viii, xv, 9, 10, 47, 48, 51–56  
**UE** User Equipment vi, viii, xv, xvi, 2–4, 11–13, 17–21, 24–29, 31–35, 38–40, 42–44, 46, 48–56  
**UPA** Uniform Planar Array 19, 29, 30, 48



# 1 Introduction

## 1.1 Motivation

The evolution toward 5G and beyond has fundamentally shifted the landscape of wireless communication deployment. Indoor environments and dense urban scenarios have become primary targets for network densification, driven by the recognition that a significant and growing share of wireless traffic originates and terminates in these scenarios. This shift places new demands on how wireless infrastructure is designed and deployed. Solutions that are lightweight and straightforward to implement are not merely preferable in such environments, they are often the only practically viable option. Extensive civil works, dedicated power supply lines, and complex maintenance procedures. Those that may be acceptable in macro-cell deployments have become serious barriers when the target environment is an aircraft cabin, a metro platform, or a dense office building. At the same time, the scale of next-generation networks is expected to grow exponentially, introducing a further dimension of pressure. As the number of connected devices grows and network densification accelerates, the scalability of any proposed solution becomes a first-order design concern. A solution that works at a small scale but imposes linearly growing infrastructure costs cannot realistically underpin a network of 6G ambitions. Taken together, the centrality of challenging indoor and dense deployment scenarios and the imperative of scalable, low-overhead solutions make the question of sustainable metasurface-assisted wireless communication design not only academically interesting but also a practical necessity.

Wireless communication systems are increasingly expected to deliver high throughput and reliable coverage in challenging environments. Indoor dense spaces (IDS), such as aircraft cabins, metro wagons, and stadiums, represent some of the most demanding deployment scenarios, where high user density and severe propagation conditions impose stringent requirements on the system. Millimeter wave (mmWave) frequencies have been identified as a promising means of meeting these throughput demands, owing to their abundant available spectrum. However, the same physical properties that make mmWave attractive also render it sensitive to significant path loss and blockage,

limiting its effective coverage in indoor environments. Addressing this coverage limitation while keeping deployment practical is a central challenge for next-generation wireless systems.

Metasurfaces have emerged as a compelling tool for shaping the wireless propagation environment, and have attracted significant research interest as a means of extending mmWave coverage. The appeal of metasurface-based solutions lies partly in their potential as lightweight, infrastructure-friendly alternatives to conventional active relays or dense access point deployments. However, this appeal rests on an assumption that is rarely examined critically in the literature, where the cost of operating the metasurface is regarded as negligible or irrelevant to the system design. In practice, operating a metasurface, particularly a reconfigurable one, incurs real costs like powering and cabling. These costs are not merely financial, they also represent a practical barrier to deployment, especially in scenarios where infrastructure is difficult or expensive to install.

This thesis adopts a different perspective. Rather than treating operating cost as a secondary concern, we place it at the center of the design problem. Intuitively, high-performing solutions tend to be costly to operate, while low-cost solutions tend to sacrifice performance. A fully reconfigurable metasurface offers adaptability and strong performance gains, but demands operating supporting infrastructure. A static metasurface minimizes operating cost, but provides only a fixed, non-adaptive channel modification. Between these two extremes lies a largely unexplored design space, one where the trade-off between performance and operating cost can be navigated more deliberately.

### 1.2 Research Question

In this thesis, we explore the optimal balance between the trade-off of performance and operating cost and investigate the design of a sustainable metasurface-assisted wireless communication system. The shallow-to-depth research questions are:

- RQ1:** How much reconfigurability is needed for a metasurface-assisted system to satisfy user requirements? Specifically, given a constraint on the number of deployable RISs, how can a mixed deployment of static metasurface (SMS) and reconfigurable intelligent surfaces (RIS) be optimally configured and placed to maximize the minimum data rate among all User Equipment (UE)s?
- RQ2:** How can a self-sustainable RIS (ssRIS)-assisted mmWave communication system in IDS be designed to maximize the minimum data rate of all UEs while satisfying the self-sustainability constraint? How can RIS, SMS, and ssRIS be positioned within the trade-off among operating cost, metasurface gain, and system coverage?
- RQ3:** How feasible is it to utilize ssRIS in various communication scenarios? Specifically, how does the element count required to achieve self-sustainability scale with varying transmit power, data rate requirements, and outage requirements for ssRIS implementing different Harvesting-and-Reflecting (HaR) schemes?

### 1.3 Thesis Contributions

To address **RQ1**, we investigate how a mixed deployment of SMSs and RISs can be optimized to deliver high data rates in an mmWave IDS. SMSs reflect incoming signals with fixed phase shifts, while RISs offer the flexibility to reconfigure their phase shifts dynamically. To characterize how this reconfigurability affects network performance, we propose an iterative algorithm based on the feasible point pursuit successive convex approximation (FPP-SCA) method, which jointly optimizes the surface types, their phase-shift configurations, and the time allocation across UEs, with the objective of maximizing the minimum data rate achieved in the network. The results of this study are presented in:

**Paper A:** Z. Li, O. A. Topal, Ö. T. Demir, E. Björnson and C. Cavdar, “Mixed Static and Reconfigurable Metasurface Deployment in Indoor Dense Spaces: How Much Reconfigurability is Needed?” 2024 IEEE Wireless Communications and Networking Conference (WCNC), Dubai, United Arab Emirates, 2024, pp. 1-6

My contributions in this work included system modeling, optimization problem formulation and convexification, numerical simulation implementation, results analysis, and manuscript preparation, with guidance from my supervisors and co-authors.

To address **RQ2**, we design and investigate the performance of an ssRIS-assisted system in IDS under the mmWave communication system. We present a large-scale ssRIS-assisted system designed with low computational complexity for optimization. In the system, the utilization of the ssRIS is simplified by considering a preset-based element splitting scheme for maintaining self-sustainability and the formation of coverage groups by manually associating ssRISs with the closest UEs. Building on this structure, we formulate a mixed-integer programming (MIP) problem that jointly optimizes the ssRIS–UE association, the phase-shift configurations of the ssRISs, and the time allocation across UEs, with the objective of maximizing the minimum achievable data rate in the network. To handle the non-convexity inherent in the problem, FPP-SCA is employed to convexify the constraints, upon which a two-stage iterative algorithm is developed. Finally, to assess the trade-off among metasurface gain, operating cost, and system coverage, we conduct a systematic comparison of the achievable data rates under ssRIS, RIS, and SMS deployments across indoor environments of varying sizes. The results of this study are presented in

**Paper B:** Z. Li, O. A. Topal, Ö. T. Demir, E. Björnson and C. Cavdar, “Self-Sustainable RIS-Assisted mmWave Indoor Communication System”, IEEE Transactions on Wireless Communications, under review.

My contributions in this work spanned the full research cycle, including system modeling, the derivation and convexification of the optimization framework, the development of heuristic simplifications to manage computational tractability, numerical simulation implementation, results analysis, and manuscript preparation, under the guidance of my supervisors and co-authors.

To address **RQ3**, we examine how element requirements scale with key system parameters, transmit power, data rate demands, and outage constraints under both line-of-sight (LoS) and non-line-of-sight (NLoS) ssRIS-to-UE channels. The results of this study are presented in

**Paper C:** Z. Li, O. Alp Topal, Ö. Tuğfe Demir, E. Björnson and C. Cavdar, “Feasibility Study Regarding Self-Sustainable Reconfigurable Intelligent Surfaces,” in *IEEE Wireless Communications Letters*, vol. 15, pp. 1045-1049, 2026

My contributions in this work included system modeling, optimization problem formulation, solution derivation, numerical simulation implementation, results analysis, and manuscript preparation, under the guidance of my supervisors and co-authors.

### 1.4 Societal Value of the Thesis

The research presented in this thesis carries societal relevance that extends beyond its immediate technical contributions. As wireless networks densify toward 5G and 6G, the aggregate energy and resource consumption of the wireless ecosystem becomes a first-order societal concern. The European Green Deal sets the objective of making the European Union climate neutral by 2050, and sustainability has been explicitly embedded as a core design principle in the European vision for 6G networks. Research that enables more self-sustaining wireless infrastructure directly supports this agenda.

This thesis contributes by advancing the understanding of how metasurface-assisted wireless systems can be designed with operating cost as a primary constraint. By demonstrating that ssRIS can navigate the trade-off between performance and operating cost, the work provides a technical foundation for deploying communication-enhancing solutions with reduced reliance on external power supply, low-overhead deployments that 6G demands. This direction also aligns with KTH’s institutional commitment to sustainable development and its systematic work toward the United Nations Sustainable Development Goals (SDG), particularly SDG 9, which calls for resilient and sustainable infrastructure that is not only technically capable but also practically viable to deploy at scale.

### 1.5 Organization

The thesis is organized into six chapters as follows:

- Chapter 1:** Introduces the motivation, research questions, thesis contributions, and societal value of the work.
- Chapter 2:** Establishes the technical background underlying the thesis, covering the characteristics and challenges of IDS, the different types of metasurfaces, ray tracing-based modeling of metasurface-assisted IDS, and fundamental Successive Convex Approximation (SCA) methods utilized in the included papers.

- Chapter 3:** Presents the optimization of a mixed SMS and RIS deployment, through which the trade-off between metasurface gain and operating cost is characterized.
- Chapter 4:** Presents the design and optimization of an ssRIS-assisted mmWave communication system in IDS, through which ssRIS is shown to achieve a favorable operating point within the gain-cost trade-off.
- Chapter 5:** Presents a fundamental feasibility study for utilizing ssRIS, examining how the element count requirement scales with varying system conditions for ssRIS implementing different HaR schemes.
- Chapter 6:** Concludes the thesis and outlines directions for future research.



## 2 Background

This chapter introduces the technical foundations underlying the research presented in this thesis. The discussion is organized around four interconnected topics. First, the characteristics of IDS are discussed, with particular attention to the dilemma of high system throughput demand and high blocking attenuation. To accurately capture these key features of IDS, ray tracing (RT) is briefly discussed. Second, metasurfaces are introduced as a means of shaping the wireless propagation environment, covering both their static form and their reconfigurable counterpart. Third, the concept of SS-RIS is presented, where energy harvesting capabilities are integrated into the surface itself, enabling autonomous operation. Finally, to tackle the complex optimization problem in the metasurface-assisted system, two Successive Convex Approximation (SCA) tricks are briefly introduced. Together, these topics establish the theoretical and practical context for the research questions addressed in the subsequent chapters.

### 2.1 Indoor Dense Spaces

Indoor dense spaces (IDS) refer to confined environments characterized by densely populated users and devices, such as aircraft cabins, metro wagons, and stadiums [4]. In these environments, a large volume of wireless traffic is generated within a spatially limited area, placing significant demand on the underlying communication infrastructure. In addition, another notable feature of the IDS is its richness in blocking obstacles, which creates challenges in maintaining wireless signal quality during propagation. Beyond throughput and coverage challenges, operating cost represents a critical practical consideration in IDS wireless system design. Here, **operating cost** refers broadly to the practical expenditures associated with deploying and maintaining active wireless infrastructure, including the power supply, cabling, and control signaling overhead required to sustain reconfigurable components in operation. Environmental regulations and space constraints in IDS mean that additional infrastructure, such as cabling, power supplies, or batteries, can increase maintenance efforts and even elevate safety risks, which is not preferred for IDS like an aircraft cabin. In summary, meeting

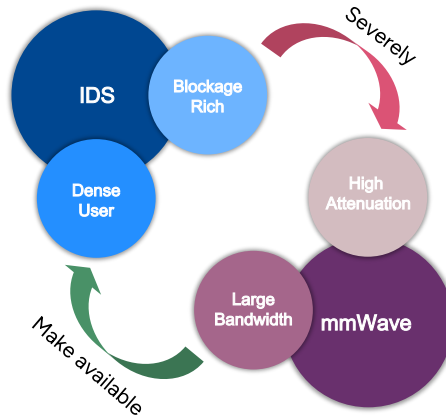


Figure 2.2.1: The dilemma of the mmWave wireless communication system in IDS

the service demands in IDS requires not only high aggregate throughput and reliable connectivity for each user, but also cost-effective and minimally intrusive deployment solutions.

## 2.2 Millimeter-Wave

Since arbitrarily increasing transmit power is impractical in environments dense with people and limited in space, Shannon’s theorem points naturally toward bandwidth expansion as the primary lever for capacity improvement, and this points to utilizing mmWave. Benefiting from its much larger available bandwidth [5], the mmWave system has been largely investigated for meeting the various requirements raised in the 5G era. However, implementing the mmWave system in IDS can encounter extra difficulties than a common indoor environment. As illustrated in Figure 2.2.1, the abundant bandwidth available at mmWave frequencies offers a promising means of serving dense users in IDS, yet the environment’s rich blockage characteristics further aggravate the inherently high propagation attenuation. This is the dilemma one has to tackle when designing wireless communication systems for IDS. One solution people have been proposing is to utilize metasurfaces.

## 2.3 Metasurfaces

Metasurfaces refer to artificial devices that can impose abrupt changes in the amplitude, phase, or polarization of incoming waves [6], and are thus widely applied for manipulating the wireless propagation environment to satisfy engineering needs. The desired variation of these electromagnetic properties is generally realized by carefully tuning the dielectric properties of the elements comprising the metasurface, through either active or passive approaches.

One of the most commonly considered active approaches to controlling the dielectric properties is to power the PIN diodes. Changing the voltage applied to one element

can result in different phase shifts in the reflected signal. One major benefit of this control mechanism is that it grants the feasibility to flexibly control the imposed phase shifts even during the operation of the metasurface. This capability is referred to as the metasurface's reconfigurability. Insignificant yet still non-negligible consumption, reconfigurability generally comes with the price of cabling, powering, and signaling [7]. On the other hand, reconfigurability also brings metasurfaces with the adaptivity to the dynamically changing wireless environment and the capability to perform complicated enhancement tasks. Reconfigurable intelligent surface (RIS) is the terminology that is generally used to describe the category of metasurface that has reconfigurability.

The wireless environment manipulation can also be achieved via passive methods. The abrupt changes to the electromagnetic properties that are injected into the incoming waves can be shaped by strategically designing the tilt, size, and materials of the elements. Benefiting from the passive nature of the control mechanism, metasurfaces that implement this are source-free, thus without requiring any additional power supply. However, it becomes incapable of reconfiguring the element while operating, and will impose a rather fixed wireless channel change after initialization. Without reconfigurability, we referred to this type of metasurface as the static metasurface (SMS).

As established in the discussion of IDS, wireless system design for such environments must address not only throughput and coverage demands but also operating cost constraints. In this context, RIS and SMS represent two extreme points in the design space. RIS maximizes adaptability and performance through reconfigurability, enabling dynamic response to changing channel conditions and sophisticated signal enhancement, but at the cost of continuous power consumption, cabling infrastructure, and control signaling overhead. SMS, by contrast, minimizes operating cost through its entirely passive operation, requiring no power supply or supporting infrastructure, but sacrifices adaptability, imposing only a fixed channel shifting that cannot respond to dynamic changes in the wireless environment. This trade-off between metasurface gain and operating cost naturally motivates the exploration of intermediate solutions.

### 2.3.1 Self-sustainable Reconfigurable Intelligent Surfaces

As discussed in [8, 9, 10, 11, 12, 13], a special RIS with the capability to perform energy harvesting to compensate for the power consumption required by its reconfigurability is investigated. In this thesis and its included works, we refer to this subcategory of RIS equipped with energy-harvesting capabilities as the self-sustainable reconfigurable intelligent surfaces (ssRIS), adopting the same terminology in [8, 12, 11]. With reconfigurability yet a much lower operating cost compared to RIS, this type of metasurface could potentially promise a better solution for the IDS mmWave system.

The harvesting-and-reflecting (HaR) scheme describes the operational logic of ssRIS. In general, the operational phases of any HaR schemes can be broadly separated into the harvesting phase and the reflecting phase. Different HaR schemes are distinguished by how resources are allocated across these two phases. For instance, the most commonly seen HaR schemes are the element splitting (ES) scheme and the time splitting (TS) scheme [13]. Figure 2.3.2 illustrates the ssRIS that implements ES and TS scheme separately. In the ES case, the harvesting and reflecting are performed simultaneously using different metasurface elements. In the TS case, the harvesting phase and reflect-

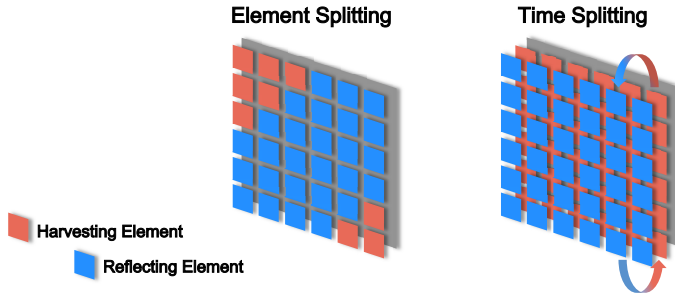


Figure 2.3.2: Illustration of the ES and TS ssRIS

ing phase happen orthogonally over time. During the harvesting phase, all elements will be connected to the harvesting circuit and perform energy harvesting, and during the reflecting phase, elements will be connected to the reflecting circuit that enables them to do the signal phase-shifting.

Self-sustainability requires that the energy harvested during the harvesting phase exceeds the energy consumed during the reflecting phase. In this thesis and its included works, the incident signal is assumed to be fully absorbed during harvesting and losslessly reflected during the reflecting phase. For the purpose of this general discussion, in this chapter, the received power at the ssRIS is generally denoted as  $P^{\text{Rc}}$ . The exact formulation varies across the papers presented in this thesis depending on the specific system models, HaR selection, and assumptions, as will be detailed in the relevant chapters. Due to the hardware imperfection and transforming efficiency, the effective harvested power will be lower than what is received. To capture this practical limitation, both linear and non-linear [14, 15] energy harvesting models are commonly employed in the literature. For the best of mathematical tractability, the linear harvesting model describes the harvesting efficiency as a constant coefficient  $\eta$  that is applied directly to the received power. Under this model, the harvested power  $P^{\text{Hr}}$  is given as

$$P^{\text{Hr}} = \eta P^{\text{Rc}}. \quad (2.3.1)$$

The non-linear model is more accurately capturing the non-linear characteristic of a real energy harvesting circuit [15]. Remain mathematically tractable while verified not to lose generality, the harvested power is calculated by referring to the non-linear energy harvesting model proposed in [14] as

$$P^{\text{Hr}} = \frac{q_1 P^{\text{Rc}}}{q_2 P^{\text{Rc}} + q_3}, \quad (2.3.2)$$

where  $q_1$ ,  $q_2$ , and  $q_3$  are fitted positive constants that are related to the efficiency of the energy harvesting hardware.

The power consumption required to maintain reconfigurability is determined by the control mechanism employed and is proportional to the number of elements operating

in the reflecting phase. Denoting the power consumption of a single reflective element as  $P_0$ , and assuming  $M$  elements contribute to redirect the incident wave, then the self-sustainability constraint is given as

$$P^{\text{Hr}} \geq P_0 M. \quad (2.3.3)$$

Given that ssRIS preserves reconfigurability while minimizing the operating cost when the self-sustainability constraint is satisfied, it can be a potential optimal within the trade-off between the metasurface gain and the operating cost.

## 2.4 Metasurface-assisted mmWave IDS system modeling

RT is a deterministic simulation technique used to model the propagation of electromagnetic waves in complex environments. By tracing the paths of individual rays as they undergo reflection, diffraction, and scattering upon interacting with physical objects and surfaces, RT produces accurate predictions of key channel parameters such as received signal strength, delay spread, and angle of arrival. Unlike statistical channel models, RT explicitly accounts for the geometry and material properties of the environment, making it particularly well-suited for indoor scenarios where the spatial layout and surface characteristics have a significant influence on propagation behavior.

This deterministic treatment is well-justified in IDS contexts. In environments such as aircraft cabins, UE mobility is typically restricted by limited space and regulatory constraints, with movement occurring at pedestrian speeds or slower. This restricted mobility results in an unusually long channel coherence time relative to typical communication frame durations. Combined with the geometry-dependent nature of mmWave links in IDS, channels are therefore modeled as deterministic and fixed in this thesis [4].

To accurately reflect the blockage richness property of the IDS, the RT simulations are considered in modeling the IDS and capturing realistic channel conditions. In this thesis and its included works, the RT simulation is performed using the commercial RT platform Wireless Insite [16].

Our prior work [4] conducted extensive mmWave channel modeling across different IDS layouts, including metro wagons and aircraft cabins. When mutual LoS connectivity is ensured through strategic device deployment, the underlying wireless channel characteristics, including angular spread, delay spread, and multipath structure, remain relatively consistent across IDS environments due to their shared properties of metallic enclosures, short propagation distances, and blockage richness. Given those, we consider a complete (in Paper B)/ partial (in Paper A) real aircraft cabin with 31 rows and 186 seats full of passengers to serve as a representative IDS environment. The performance trends, trade-off insights, and comparative advantages demonstrated in this work are expected to generalize to other IDS scenarios under appropriate deployment strategies.

Metasurface is generally not an integrated component in RT simulation platforms. To realistically mimic the behaviour of the metasurface in RT, a transceiver antenna array is utilized to represent the elements in the metasurface, with one antenna element

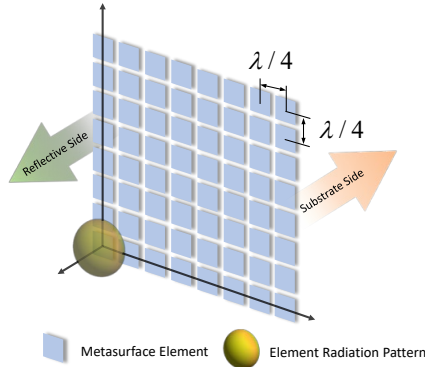


Figure 2.4.3: Illustration of the metasurface model in RT simulator

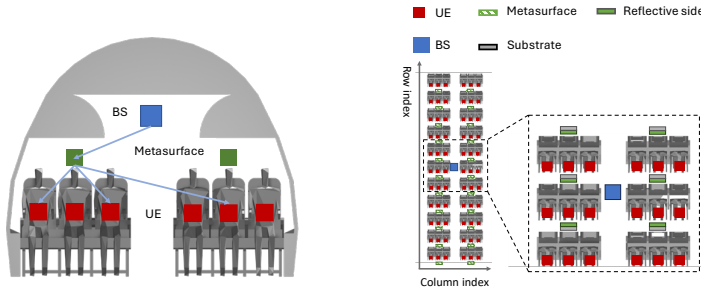


Figure 2.4.4: The front view and the top view of the considered (partial) cabin.

representing one metasurface element. As indicated in [7], the radiation pattern of the metasurface element is similar to that of a patch antenna. Therefore, the cosine antenna is considered to build the antenna array. Throughout the work presented in this thesis, the metasurface elements are uniformly spaced at  $\lambda/4$  spacing in both horizontal and vertical directions, where  $\lambda$  denotes the wavelength of the carrier signal. Later, a perfect wave absorber with the same size as the transceiver antenna array is added 1 cm behind the array since the RIS is not penetrable, and only one side of RIS reflects. We use the term reflective side to indicate the side without a perfect wave absorber. The detailed modeling is given in Figure 2.4.3.

The geometry of the considered airplane environment is shown in Fig. 2.4.4. The Base Station (BS) is placed in the middle of the cabin. For each passenger, a receiving node is placed to represent the UE held by the passenger. The material of the cabin shell is Acrylonitrile Butadiene Styrene (ABS), which is widely used in aircraft [17]. Windows and passenger seats are set to glass and nylon, respectively. Passengers are modeled by the same human skin model as in [18]. The dielectric properties of the materials we consider are given in Table 2.4.1. The metasurfaces are placed above the middle seats on both sides of the corridor in each row, with the reflective side perpendicular to the ground and facing the BS to provide the highest Signal-to-Noise Ratio (SNR) to the

UEs [19].

**Table 2.4.1:** Dielectric properties of materials at 28 GHz considered in the simulation.

Material	dielectric properties		Thickness (cm)
	$\epsilon$	$\sigma$	n/a
Skin [20]	19.3	30.40	0.1
ABS [21]	2.4	0.028	0.3
Nylon [22]	3.01	0.03	0.25
Glass [23]	6.27	0.15	0.3

## 2.5 Successive Convex Approximation

Many optimization problems that arise in wireless communication system design are inherently non-convex, making them intractable to solve directly. Successive convex approximation (SCA) [24] is a general algorithmic framework for addressing such problems by iteratively constructing and solving a sequence of convex approximations. At each iteration, a locally tight convex surrogate of the original non-convex objective or constraint is formed around the current iterate, and the resulting convex subproblem is solved to obtain an updated solution. Under appropriate regularity conditions, this iterative procedure is guaranteed to converge to a stationary point of the original problem.

A key strength of the SCA framework lies in its flexibility. Different surrogate construction strategies give rise to distinct algorithmic variants suited to different problem structures. Two such variants that are employed in the optimization algorithms developed in this thesis are the feasible point pursuit successive convex approximation (FPP-SCA) [25] method and the concave–convex procedure [26] (CCCP).

### 2.5.1 Concave Convex Procedure

A general instance of the SCA principle is the CCCP [26], which applies to virtually any energy minimization problem whose objective function has a bounded Hessian. The key idea is to decompose the differentiable objective  $f(\mathbf{x})$  into a convex and a concave part,

$$f(\mathbf{x}) = f_{\text{convex}}(\mathbf{x}) + f_{\text{concave}}(\mathbf{x}), \quad (2.5.4)$$

and at each iteration  $i$ , replace the concave part by its first-order Taylor expansion at the current point  $\mathbf{x}^{(i)}$ , which constitutes a global upper bound due to concavity. This yields a sequence of convex surrogate problems of the form

$$\mathbf{x}^{(i+1)} = \arg \min_{\mathbf{x}} f_{\text{convex}}(\mathbf{x}) + \nabla f_{\text{concave}}(\mathbf{x}^{(i)})(\mathbf{x} - \mathbf{x}^{(i)}), \quad (2.5.5)$$

which is equivalently characterized by the fixed-point condition  $\nabla f_{\text{convex}}(\mathbf{x}^{(i+1)}) = -\nabla f_{\text{concave}}(\mathbf{x}^{(i)})$ . Each iteration is guaranteed to monotonically decrease  $f(\mathbf{x})$ , and the procedure converges to a stationary point [26]. While CCCP provides a broadly

applicable framework, the specific structure of the QCQP problems arising in the SS-RIS-assisted system optimization calls for a more tailored treatment, motivating the use of FPP-SCA, which is introduced next.

## 2.5.2 Feasible Point Pursuit Successive Convex Approximation

FPP-SCA is a SCA-based algorithm for non-convex quadratically constrained quadratic programs (QCQPs). A central challenge with standard SCA is that it requires a feasible initial point, which is difficult to obtain in general. Approximating a constraint around an infeasible point may yield an empty feasible set, even when the original problem is feasible. FPP-SCA addresses this by augmenting the approximated constraints with non-negative slack variables, ensuring that the convex subproblem at each iteration always admits a feasible solution. A penalty term weighted by a large coefficient is added to the objective to discourage the use of these slacks, driving the iterates toward feasibility. Once a feasible point of the original non-convex problem is identified, convergence to a Karush-Kuhn-Tucker (KKT) point is thereafter ensured.

A QCQP can be generally expressed as

$$\mathbf{P2.1} \quad \underset{\mathbf{x} \in \mathbb{C}^n}{\text{minimize}} \quad \mathbf{x}^H \mathbf{A}_0 \mathbf{x} \quad (2.5.6a)$$

$$\text{s.t.} \quad \mathbf{x}^H \mathbf{A}_m \mathbf{x} \leq c_m, \quad m = 1, \dots, M \quad (2.5.6b)$$

where  $\mathbf{A}_0 \succeq 0$  is a positive semi-definite (PSD), and  $\mathbf{A}_m \in \mathbb{C}^{n \times n}$  are Hermitian matrices for all  $m \in \{1, \dots, M\}$ . For the case when  $\mathbf{A}_m \succeq 0$  for all  $m$ , the QCQP **P1.1** is a convex optimization problem that can be efficiently solved. For a more general case, where  $\mathbf{A}_m$  is not PSD, this problem becomes non-convex and NP-hard. Using eigen-decomposition, the matrix  $\mathbf{A}_m$  can be expressed as

$$\mathbf{A}_m = \mathbf{A}_m^{(+)} + \mathbf{A}_m^{(-)} \quad (2.5.7)$$

where  $\mathbf{A}_m^{(+)} \succeq 0$  and  $\mathbf{A}_m^{(-)} \preceq 0$ . Since  $\mathbf{A}_m^{(-)}$  is negative semi-definite, for any  $\mathbf{z}$  and  $\mathbf{x}$ ,  $(\mathbf{x} - \mathbf{z})^H \mathbf{A}_m^{(-)} (\mathbf{x} - \mathbf{z})$  holds. By formatting, it can be expressed as

$$\mathbf{x}^H \mathbf{A}_m^{(-)} \mathbf{x} \leq 2\Re\{\mathbf{z}^H \mathbf{A}_m^{(-)} \mathbf{x}\} - \mathbf{z}^H \mathbf{A}_m^{(-)} \mathbf{z}. \quad (2.5.8)$$

Therefore, using the linear restriction around the point  $\mathbf{z}$ , the  $m$ -th non-convex constraint can be approximated as

$$\begin{aligned} & \mathbf{x}^H \mathbf{A}_m \mathbf{x} \leq c_m \\ & = \mathbf{x}^H \mathbf{A}_m^{(+)} \mathbf{x} + \mathbf{x}^H \mathbf{A}_m^{(-)} \mathbf{x} \leq c_m \\ & = \mathbf{x}^H \mathbf{A}_m^{(+)} \mathbf{x} + 2\Re\{\mathbf{z}^H \mathbf{A}_m^{(-)} \mathbf{x}\} - \mathbf{z}^H \mathbf{A}_m^{(-)} \mathbf{z} - s_m \leq c_m \end{aligned} \quad (2.5.9)$$

which is convex with respect to the optimization variable  $\mathbf{x}$ . In (2.5.9),  $s_m$  is the slack variable introduced for avoiding the infeasibility issue that comes with the selection of  $\mathbf{z}$ . By penalizing  $s_m$  in the objective function, the optimization process searches for a

point that is feasible to the original problem **P2.1**. Together, after applying FPP-SCA, the tractable optimization problem is formatted as

$$\mathbf{P2.2} \underset{\mathbf{x}, \{s_m\}}{\text{minimize}} \quad \mathbf{x}^H \mathbf{A}_0 \mathbf{x} + \lambda \sum_{m=1}^M s_m \quad (2.5.10a)$$

$$\text{s.t.} \quad \mathbf{x}^H \mathbf{A}_m^{(+)} \mathbf{x} + 2\Re\{\mathbf{z}^H \mathbf{A}_m^{(-)} \mathbf{x}\} - \mathbf{z}^H \mathbf{A}_m^{(-)} \mathbf{z} - s_m \leq c_m \quad (2.5.10b)$$

$$s_m \geq 0 \quad (2.5.10c)$$

where  $\lambda$  is a positive penalty coefficient. The iterative algorithm that solves **P2.1** is given as

---

**Algorithm 1** FPP-SCA Algorithm

---

- 1: **Initialize:** Set iteration counter  $i = 0$ , and randomly initialize  $\mathbf{z}$ .
  - 2: **while** Until converge **do**
  - 3:      $i \leftarrow i + 1$
  - 4:     Solve **P2.2** and set  $\mathbf{z}$  to its solution  $\mathbf{x}^*$
  - 5: **end while**
- 

In general, the rate optimisation problem in a metasurface-assisted system can be formulated as a QCQP. In this thesis and all its included works, FPP-SCA is considered primary for handling the non-convexity.



### 3 Trade-off between operating cost and metasurface gain

While RIS offers superior adaptability over SMS by dynamically adjusting phase-shifts to serve individual users, this capability introduces non-negligible operating costs in the form of power supply, cabling infrastructure, and control signaling. Whether these costs are justified depends critically on the extent of the performance improvement provided by reconfigurability. While the direction of this improvement is intuitive, it is not easy to characterize its magnitude in practice.

This distinction carries significant implications for deployment design. If reconfigurability yields substantial and persistent gains, the deployment of full RIS would be warranted, despite the associated operational overheads. Conversely, if the marginal gain diminishes as the proportion of RISs increases, a denser deployment of SMSs, which are passive, require no operational infrastructure, and scale at a lower cost, may achieve the same performance at a lower cost. Quantifying this trade-off between operating cost and metasurface gain is therefore a practical prerequisite for principled deployment decisions in metasurface-assisted IDS.

This chapter presents the work in **Paper A**, which directly addresses research question **RQ1**:

How much reconfigurability is needed for a metasurface-assisted system to satisfy user requirements? Specifically, given a constraint on the number of deployable RISs, how can a mixed deployment of SMS and RIS be optimally configured and placed to maximize the minimum data rate among all UEs?

To this end, mixed deployments of SMSs and RISs in an aircraft cabin environment are investigated under varying deployment budgets. The results show that while replacing

SMSs with RISs consistently improves the minimum achievable data rate, the gain becomes negligible beyond a certain number of RISs, meaning that full reconfigurability is not always necessary to reach high performance. These findings establish the trade-off between operating cost and metasurface gain on concrete numerical grounds, and motivate the subsequent investigation into ssRIS as a principled middle-ground between the two extremes.

### 3.1 Related Work

Metasurface-assisted wireless communication system design has been widely investigated in recent years. Benefiting from the reconfigurability, RIS is widely considered for various communication scenarios. In [27], a RIS is used in an indoor corridor area. In the non-line-of-sight area, a 15-20 dB gain in received power is observed. In [28], the deployment of the BS and the RIS is jointly optimized to extend the coverage in an outdoor environment. In our previous work [19], we consider the deployment of the RISs to extend the mmWave coverage in an IDS. The main goal was to minimize the number of deployed RISs while satisfying the data rate requirements of UEs. In [29], the SMS has also been used to enhance the signal coverage of the mmWave network in an indoor environment. A comparison work on metasurface reconfigurability has been conducted in [30], where the elements of the metasurface are grouped and configured as a whole. Their results show that the less reconfigurable the metasurface is, the less of a gain it promises to bring. However, how large the advantage of RIS over SMS was not investigated, considering deployment costs in a real-deployment setup.

Except for the theoretical work featuring metasurface, prior research has also established the foundations of metasurface technology, with both RIS and SMS prototypes demonstrated in academic and industrial fields. As reported in [7, 31], RIS prototypes have been realized using low-cost materials, positioning them as a cost-effective alternative to deploying additional base stations and making dense deployment practically feasible. Nevertheless, these prototypes still require cabling and a power supply, which constitute small yet non-negligible operating costs. In contrast, [32] demonstrates a transparent SMS prototype whose phase-shifts are configured manually through mechanical means, eliminating the need for active supplement. Suggested by the authors, such surfaces could be widely and densely deployed, for instance, on window surfaces, without imposing significant installation constraints or environmental limitations. These prototypes have confirmed the need to understand how much reconfigurability is required in order to provide a solid foundation for utilizing them to assist real communication scenarios.

### 3.2 Contributions

The main contributions of this work are as follows:

- We tackle a novel mixed SMS and RIS deployment problem, which aims to maximize the minimum data rate among UEs by jointly choosing where to deploy SMSs and RISs from the candidate locations, determining the phase-shifts of the RISs and SMSs, and allocating time resources to the UEs.
- We relax the original non-convex problem by utilizing the feasible point pursuit

successive convex approximation (FPP-SCA) method after several mathematical manipulations. We propose an iterative algorithm to find a solution to the original problem, where at each iteration a mixed integer programming (MIP) problem is solved.

- The numerical results indicate that to reach a high data rate, it is not always necessary to introduce more reconfigurability. After some point, replacing more SMSs with RISs has a negligible increase in performance.

### 3.3 System Model

We consider a downlink mmWave communication system in which a single BS serves  $K$  UEs within an IDS. The BS is equipped with  $N$  antennas arranged as a  $\sqrt{N} \times \sqrt{N}$  uniform planar array (UPA), and each UE has a single antenna. The UEs are indexed by  $k \in \{1, 2, \dots, K\}$ . To eliminate inter-user interference, a time-division scheme is adopted, where  $\tau_k$  denotes the fraction of time allocated to UE  $k$ . The time allocations satisfy  $\sum_{k=1}^K \tau_k = 1$ , and each individual allocation is constrained to  $\tau_k \in [\tau_{\min}, 1]$ , where  $\tau_{\min} > 0$  is the minimum allowable time fraction.

To extend the coverage and enhance the downlink rates in the IDS,  $L$  metasurfaces are deployed and indexed by  $l \in \{1, 2, \dots, L\}$ . Each surface is composed of  $M$  elements arranged as a  $\sqrt{M} \times \sqrt{M}$  UPA. Two types of metasurfaces are considered in the deployment. The first type is a *static metasurface (SMS)*, whose phase-shifts are fixed at the time of manufacturing and cannot be altered during operation. Since no active components are involved, an SMS requires neither power supply, cabling, nor control signaling. The second type is a *reconfigurable intelligent surface (RIS)*, which dynamically adjusts its phase-shifts at each time slot to maximize the performance of the currently served UE. This adaptability, however, comes at the cost of power consumption, cabling, and control signaling overhead. To reduce the costs, we assume the system has a deployment budget that can support a maximum of  $L_{\max}$  RISs. Due to the inexpensive producing price of the metasurface and the source-free property of the SMS. We consider SMS could be densely deployed without introducing deployment cost. Therefore, the remaining  $L - L_{\max}$  surfaces are deployed as SMSs. To distinguish the two types, a binary variable  $\alpha_l \in \{0, 1\}$  is introduced as a surface type indicator, where  $\alpha_l = 1$  corresponds to a RIS and  $\alpha_l = 0$  corresponds to an SMS.

As illustrated in Fig. 3.3.1, the direct channel from the BS to UE  $k$  is denoted as  $\mathbf{h}_k \in \mathbb{C}^N$ . The channel from the BS to metasurface  $l$  and from metasurface  $l$  to UE  $k$  are denoted as  $\mathbf{G}_l \in \mathbb{C}^{M \times N}$  and  $\mathbf{g}_{l,k} \in \mathbb{C}^M$ , respectively. Given the quasi-static nature of the IDS environment and the geometry-dependent characteristics of mmWave propagation, all channels are modeled as deterministic and time-invariant [4]. The phase-shift configuration of RIS  $l$  when serving UE  $k$  is represented by the vector  $\phi_{l,k} \in \mathbb{C}^M$ , where the  $m$ -th entry  $\phi_{l,k,m}$  captures the complex response of the  $m$ -th element. Assuming lossless reflection, the unit-modulus constraint  $|\phi_{l,k,m}| = 1$  holds for all  $l, k$ , and  $m$ . Analogously, the phase-shift vector of SMS  $l$  is denoted as  $\theta_l \in \mathbb{C}^M$ , with  $|\theta_{l,m}| = 1$  for all  $l$  and  $m$ . The effective phase-shift vector of surface  $l$  when serving UE  $k$  is then expressed as

$$\psi_{l,k} = \alpha_l \phi_{l,k} + (1 - \alpha_l) \theta_l, \quad (3.3.1)$$

where  $\alpha_l$  selects the active phase-shift vector according to the surface type.

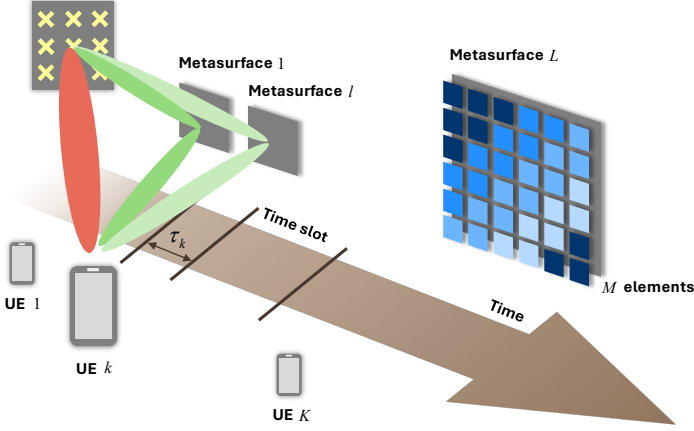


Figure 3.3.1: Illustration of the RIS/SMS-assisted system model.

To facilitate the problem formulation, the cascaded BS-metasurface-UE channel is aggregated as  $\mathbf{H}_{l,k} = \mathbf{G}_l^T \bar{\mathbf{G}}_{l,k} \in \mathbb{C}^{N \times M}$ , where  $\bar{\mathbf{G}}_{l,k} = \text{diag}(\mathbf{g}_{l,k})$ . The received signal at UE  $k$  is given by

$$\mathbf{y}_k = \left( \mathbf{h}_k + \sum_{l=1}^L \underbrace{\mathbf{G}_l^T \bar{\mathbf{G}}_{l,k}}_{\triangleq \mathbf{H}_{l,k}} \psi_{l,k} \right)^T \mathbf{w}_k x_k + n_k, \quad (3.3.2)$$

where  $x_k \in \mathbb{C}$  is the transmitted data symbol for UE  $k$  with transmit power  $\mathbb{E}\{|x_k|^2\} = P$  for all  $k$ . The unit-norm precoding vector is denoted as  $\mathbf{w}_k \in \mathbb{C}^N$ , and the receiver noise is modeled as  $n_k \sim \mathcal{CN}(0, BN_0)$ , where  $B$  is the communication bandwidth in Hz and  $N_0$  is the noise spectral density in W/Hz.

Since UEs are served without mutual interference under the orthogonal time-division scheme, the optimal precoding strategy at the BS is maximum ratio transmission (MRT). The MRT precoder for UE  $k$  is

$$\mathbf{w}_k = \frac{\left( \mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \psi_{l,k} \right)^*}{\left\| \mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \psi_{l,k} \right\|}. \quad (3.3.3)$$

The resulting signal-to-noise ratio (SNR) at UE  $k$  is

$$\Gamma_k = \frac{\left\| \mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \psi_{l,k} \right\|^2 P}{BN_0}, \quad (3.3.4)$$

which yields the achievable data rate as

$$R_k = \tau_k B \log_2(1 + \Gamma_k). \quad (3.3.5)$$

In the following section, a mixed SMS and RIS deployment optimization problem will be formulated.

### 3.4 Problem Formulation

To ensure fair service quality across all UEs, the optimization objective is to maximize the minimum achievable data rate in the IDS. This is accomplished by jointly determining the deployment type of each surface within the given budget, optimizing the phase-shift configurations of all SMSs and RISs, and allocating time resources among the UEs. The resulting mixed SMS and RIS deployment optimization problem is formulated as

$$\mathbf{P3.1} : \underset{\{\alpha_l, \tau_k, \phi_{l,k}, \theta_l\}}{\text{maximize}} \quad \min_{k \in \{1, \dots, K\}} R_k \quad (3.4.6a)$$

subject to:

$$\sum_{k=1}^K \tau_k \leq 1, \quad (3.4.6b)$$

$$\tau_k \geq \tau_{\min}, \quad \forall k, \quad (3.4.6c)$$

$$|\phi_{l,k,m}| = 1, \quad \forall l, k, m, \quad (3.4.6d)$$

$$|\theta_{l,m}| = 1, \quad \forall l, m, \quad (3.4.6e)$$

$$\alpha_l \in \{0, 1\}, \quad \forall l, \quad (3.4.6f)$$

$$\sum_{l=1}^L \alpha_l \leq L_{\max}. \quad (3.4.6g)$$

The objective function in (3.4.6a) maximizes the minimum achievable data rate across all UEs in the IDS. Constraints (3.4.6b) and (3.4.6c) enforce that the total allocated time fractions do not exceed unity and that each individual allocation remains above the minimum threshold  $\tau_{\min}$ , respectively. Constraints (3.4.6d) and (3.4.6e) impose the unit-modulus condition on the phase-shift vectors, reflecting the lossless reflection assumption. Constraint (3.4.6f) restricts the surface type indicator to binary values, and constraint (3.4.6g) ensures that the total number of deployed RISs does not exceed the deployment budget  $L_{\max}$ .

For ease of analysis, we introduce the variable  $R_{\min}$  that represents the minimum rate and recast **P3.1** as

$$\mathbf{P3.2} : \underset{\{\alpha_l, \tau_k, \phi_{l,k}, \theta_l, R_{\min}\}}{\text{maximize}} \quad R_{\min} \quad (3.4.7a)$$

subject to:

$$\tau_k B \log_2 \left( 1 + \frac{\|\mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \psi_{l,k}\|^2 P}{BN_0} \right) \geq R_{\min}, \quad \forall k, \quad (3.4.7b)$$

$$(3.4.6b), (3.4.6c), (3.4.6d), (3.4.6e), (3.4.6f), (3.4.6g).$$

The optimization problem is non-convex due to several sources of coupling and non-linearity. First, the binary variable  $\alpha_l$  is coupled with the complex variables  $\phi_{l,k}$  and

$\theta_l$  through the definition of  $\psi_{l,k}$  in (3.3.1). Second, the quadratic form inside the logarithm and the multiplicative coupling between the SNR term and  $\tau_k$  in (3.4.7b) introduce additional non-convexity. Third, the unit-modulus equality constraints (3.4.6d) and (3.4.6e) are non-linear and render the problem intractable in its original form. To address these challenges, a series of relaxations and convex approximations are introduced, and an iterative algorithm is proposed to obtain a solution to the original problem.

### 3.4.1 Mixed SMS and RIS deployment algorithm

To facilitate convex reformulation, the objective is modified by substituting  $R_{\min}$  with  $r_{\min} = \sqrt{R_{\min}}$  in **P3.2**. Since the square root is a monotonically increasing function, this substitution preserves the optimal solution while enabling the constraints to be expressed in a more tractable form.

An auxiliary variable  $d_k$  is introduced to decompose (3.4.7b) into two separate constraints

$$d_k \geq 2^{\frac{r_{\min}^2}{B\tau_k}}, \quad \forall k, \quad (3.4.8)$$

$$1 + \frac{\|\mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \psi_{l,k}\|^2 P}{BN_0} \geq d_k, \quad \forall k. \quad (3.4.9)$$

To further handle the quadratic-over-linear term in (3.4.8), a second auxiliary variable  $e_k$  is introduced, yielding

$$d_k \geq 2^{e_k}, \quad \forall k, \quad (3.4.10)$$

$$e_k \geq \frac{r_{\min}^2}{B\tau_k}, \quad \forall k, \quad (3.4.11)$$

where (3.4.10) is convex, and (3.4.11) can be equivalently cast as the second-order cone (SOC) constraint

$$e_k + \tau_k \geq \left\| \begin{bmatrix} \sqrt{2/B} r_{\min} & e_k & \tau_k \end{bmatrix} \right\|, \quad \forall k. \quad (3.4.12)$$

The coupling between the rate constraint and  $\tau_k$  has thereby been resolved. The remaining source of non-convexity lies in the quadratic norm term in (3.4.9). To address this, the coupling introduced by  $\psi_{l,k}$  is first decoupled through the auxiliary variable  $\mathbf{z}_{l,k}$ , defined via the constraints

$$\mathbf{z}_{l,k} = \phi_{l,k} + \theta_l, \quad \forall l, k, \quad (3.4.13)$$

$$|\phi_{l,k,m}| \leq \alpha_l, \quad \forall l, k, m, \quad (3.4.14)$$

$$|\theta_{l,m}| \leq 1 - \alpha_l, \quad \forall l, m. \quad (3.4.15)$$

When  $\alpha_l = 1$ , constraint (3.4.15) forces all entries of  $\theta_l$  to zero, effectively decoupling  $\alpha_l$  from the SMS phase-shifts. Conversely, when  $\alpha_l = 0$ , the RIS phase-shift vector  $\phi_{l,k}$  is nullified by constraint (3.4.14), canceling reconfigurability. In addition, the unit-modulus constraints are relaxed for tractability; at the conclusion of the algorithm,

all entries of  $\mathbf{z}_{l,k}$  are normalized to restore unit modulus. Substituting  $\mathbf{z}_{l,k}$  for  $\psi_{l,k}$  in (3.4.9) and expanding the norm-squared term yields

$$\begin{aligned} & \mathbf{z}_k^H \underbrace{\mathbf{H}_k^H \mathbf{H}_k}_{\triangleq \mathbf{A}_k} \mathbf{z}_k + 2\Re \left( \mathbf{z}_k^H \underbrace{\mathbf{H}_k^H \mathbf{h}_k}_{\triangleq \mathbf{b}_k} \right) + \underbrace{\mathbf{h}_k^H \mathbf{h}_k}_{\triangleq c_k} \\ & \geq \frac{BN_0(d_k - 1)}{P}, \quad \forall k, \end{aligned} \quad (3.4.16)$$

where  $\mathbf{H}_k = [\mathbf{H}_{1,k}, \dots, \mathbf{H}_{L,k}] \in \mathbb{C}^{N \times LM}$  and  $\mathbf{z}_k = [\mathbf{z}_{1,k}^T, \dots, \mathbf{z}_{L,k}^T]^T \in \mathbb{C}^{LM}$ . The constraint (3.4.16) remains non-convex due to the quadratic term  $\mathbf{z}_k^H \mathbf{A}_k \mathbf{z}_k$ . This is addressed by applying the FPP-SCA method [25]. Since  $\mathbf{A}_k$  is positive semi-definite, the following inequality holds for any arbitrary vector  $\zeta_k$ :

$$\mathbf{z}_k^H (-\mathbf{A}_k) \mathbf{z}_k \leq 2\Re (\zeta_k^H (-\mathbf{A}_k) \mathbf{z}_k) - \zeta_k^H (-\mathbf{A}_k) \zeta_k, \quad \forall k. \quad (3.4.17)$$

At each iteration  $i$ , setting  $\zeta_k = \mathbf{z}_k^{(i-1)}$  and introducing a non-negative slack variable  $s_k$ , the quadratic term in (3.4.16) is replaced by its affine approximation from (3.4.17), giving

$$\begin{aligned} & -2\Re \left( \left( \mathbf{z}_k^{(i-1)} \right)^H \mathbf{A}_k \mathbf{z}_k \right) + \left( \mathbf{z}_k^{(i-1)} \right)^H \mathbf{A}_k \mathbf{z}_k^{(i-1)} - 2\Re (\mathbf{z}_k^H \mathbf{b}_k) \\ & \leq s_k + c_k + \frac{BN_0}{P} (1 - d_k), \quad \forall k. \end{aligned} \quad (3.4.18)$$

The slack variables  $s_k$  are introduced to ensure feasibility at every iteration, as the successive convex approximation step may otherwise produce infeasible subproblems in early iterations. A large value of  $s_k$  at a given iteration indicates a significant violation of inequality (3.4.17), and a penalty term weighted by  $\Omega \gg 1$  is incorporated into the objective to drive these violations toward zero over successive iterations. At each iteration, the following penalized subproblem is solved optimally using a MIP solver:

$$\mathbf{P3.3} : \quad \underset{\{\alpha_l, \tau_k, \mathbf{z}_k, \phi_{l,k}, \theta_l, d_k, c_k, s_k, r_{\min}\}}{\text{maximize}} \quad r_{\min} - \Omega \sum_{k=1}^K s_k \quad (3.4.19a)$$

subject to:

$$(3.4.6b), (3.4.6c), (3.4.6f), (3.4.6g), (3.4.10),$$

$$(3.4.12), (3.4.13), (3.4.14), (3.4.15), (3.4.18)$$

$$s_k \geq 0, \quad \forall k. \quad (3.4.19b)$$

The steps of the proposed FPP-SCA algorithm are outlined in Algorithm 2. Note that after iterating, when the penalty term approaches zero, any feasible solution to **P3** will also be a feasible solution to **P1**. The objective function is bounded due to the limitation in the number of deployed RISs, transmit power, and time resources. Furthermore, since the solution improves the objective after each iteration, Algorithm 2 will converge to a KKT point of the original problem **P3.1** [25].

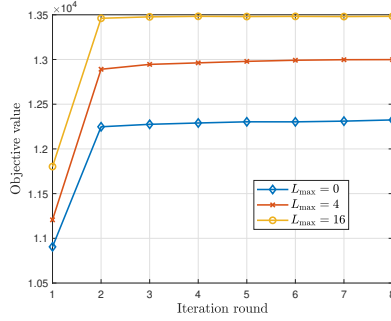


Figure 3.5.2: Objective value of the proposed algorithm with respect to each iteration

---

### Algorithm 2 Mixed SMS and RIS Deployment Algorithm

---

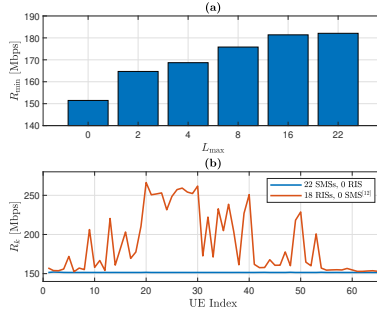
- 1: **Given:** The maximum number  $L_{\max}$  of RISs
  - 2: **Initialization:** Initialize  $\mathbf{z}_k^{(0)}$  randomly while keeping  $|\mathbf{z}_{l,k,m}^{(0)}| = 1, \forall l, k, m$ . Set the penalty coefficient  $\Omega \gg 1$ . Set the iteration counter to  $i = 0$ . Set the maximum iteration number to  $I$ .
  - 3: **while**  $i < I$  **do**
  - 4:      $i \leftarrow i + 1$
  - 5:     Solve **P3.3** and set  $\mathbf{z}_k^{(i)}, \forall k$  to its solution
  - 6: **end while**
  - 7:  $z_{l,k,m}^{(I)} \leftarrow z_{l,k,m}^{(I)} / |z_{l,k,m}^{(I)}|, \forall l, k, m$
  - 8: **Output:** The surface type indicator  $\alpha_l^{(I)}$ , phase-shift configurations  $\phi_{l,k}^{(I)}$  of RISs and  $\theta_l^{(I)}$  of SMSs, and allocated time portions  $\tau_k^{(I)}$
- 

## 3.5 Results

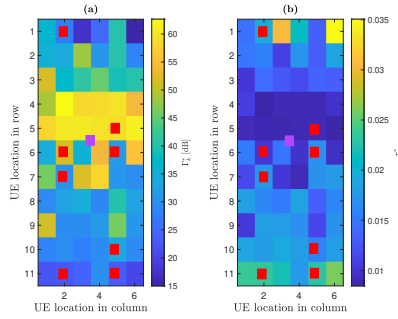
The IDS channel is obtained via RT simulation with the partial cabin considered as stated in Section 2.4. In total, 11 rows, 22 metasurfaces, and 66 seated passengers each holding one UE are considered.

Problem **P3.1** is solved using the proposed Algorithm 2 for deployment budgets  $L_{\max} \in \{0, 2, 4, 8, 16, 22\}$ . The case  $L_{\max} = 0$  corresponds to the scenario where operational constraints entirely prohibit RIS deployment, resulting in an all-SMS configuration. The case  $L_{\max} = 22$  represents an unconstrained deployment budget in which all surfaces can be RISs. The four intermediate cases involve mixed deployments of both SMSs and RISs.

The convergence behavior of the proposed algorithm is illustrated in Fig. 3.5.2, with the maximum iteration number set to  $I = 8$  and the penalty coefficient to  $\Omega = 100$ . The algorithm consistently converges within two to three iterations across all simulated deployment budgets.



**Figure 3.5.3:** (a) Optimized minimum data rate under different  $L_{\max}$ . (b) To reach the same minimum data rate when using 22 SMSs, at least 18 RISs are needed.



**Figure 3.5.4:** (a) Optimized SNR  $\Gamma_{k,r}^*$ , (b) optimized allocated time portion  $\tau_{k,r}^*$ , in the case where  $L_{\max} = 8$ . The red boxes represent the location of the deployed RISs and the purple box represents the location of the BS.

The optimized minimum data rate under varying  $L_{\max}$  is presented in Fig. 3.5.3(a). When only SMSs are deployed ( $L_{\max} = 0$ ), a minimum data rate of 152 Mbps is achieved. Replacing two SMSs with RISs already yields a 13 Mbps improvement, demonstrating that even limited reconfigurability is beneficial in this scenario. As  $L_{\max}$  increases further, the minimum data rate continues to improve. However, the marginal gain diminishes progressively. In particular, increasing  $L_{\max}$  from 16 to 22, which means replacing six additional SMSs with RISs, results in less than 1 Mbps of additional gain. This indicates that a deployment of 16 RISs already approaches the performance ceiling, beyond which further reconfigurability provides negligible benefit.

Fig. 3.5.3(b) compares the all-SMS case against the RIS-only deployment considered in [19], where the objective was to minimize the number of RISs required to satisfy a data rate threshold of 152 Mbps for all UEs. In that work, individual UE rates span a wide range of values due to the threshold-based formulation. The comparison reveals that deploying 22 SMSs achieves the same minimum data rate as deploying 18 RISs, suggesting that when the deployment budget is limited, densely deploying passive SMSs constitutes a viable and cost-effective alternative to RIS-based solutions.

To examine how the algorithm distributes resources across UEs, the optimized SNR and time allocation are shown in Fig. 3.5.4 for the case  $L_{\max} = 8$ . As seen in Fig. 3.5.4(a), UEs located near the BS in the center of the cabin benefit from strong direct links, resulting in higher post-optimization SNR values. For UEs further from the BS, severe mmWave blockage attenuation causes a rapid SNR degradation. The time allocation in Fig. 3.5.4(b) reflects a complementary pattern. UEs with lower SNR are compensated with a larger time portion, while those with higher SNR receive less, thereby equalizing the achievable data rates across all UEs.

### 3.6 Conclusion

This chapter addressed the research question **RQ1**. To answer this, a mixed SMS and RIS deployment problem was formulated and solved, jointly optimizing the surface types, phase-shift configurations, and time resource allocation within a given deployment budget. Channel coefficients were obtained from RT simulations of an aircraft cabin, providing a realistic basis for evaluating system performance.

The numerical results provide concrete evidence for the trade-off between operating cost and metasurface gain. The gap between an all-SMS deployment and an all-RIS deployment is substantial, confirming that reconfigurability does offer a significant performance advantage and that the trade-off is worth investigating. At the same time, this gain is not uniformly distributed across the deployment budget. The improvement in minimum data rate is most pronounced when the first few RISs are introduced, and diminishes progressively as more SMSs are replaced. Beyond a certain point, the marginal benefit of additional reconfigurability becomes negligible, meaning that the operating cost of further RIS deployment is no longer justified by the performance return. This suggests that a mixed deployment, rather than committing fully to either extreme, represents a more cost-effective strategy.

Furthermore, the numerical result also shows that the performance loss from abandoning reconfigurability can be partially compensated by increasing the density of passive metasurfaces. However, to reach a higher quality-of-service threshold, for a deployment budget-limited system, it is not always a practical guideline, as massive deployment SMS can still violate the environmental regulations bring unacceptable costs. This is especially true for environments like IDS. This motivates the subsequent investigation into ssRIS as a principled approach to identifying and operating at this optimal point.

## 4 Self-sustainable RIS, a balanced compromise within the trade-off

While densely deploying SMSs can partially compensate for the performance loss incurred by abandoning reconfigurability, this approach introduces practical concerns that are particularly acute in IDSs. In an aircraft cabin, for instance, a massive SMS deployment risks narrowing the already limited space available to passengers, potentially obstructing emergency pathways and raising safety concerns that may conflict with aviation regulations. These considerations point toward a source-free solution that achieves sufficiently high metasurface gain per device, such that service requirements can be met without resorting to excessive deployment density. As introduced in Section 2.3.1, ssRIS aligns closely with this objective by combining reconfigurability with self-sustainability through energy harvesting. However, ssRIS is not without its own challenges. The self-sustainability constraint imposes a strict condition on the energy balance of each surface, which can limit the available signal power for reflection and restrict feasible deployment locations. Furthermore, managing the association between ssRISs and UEs under varying channel conditions introduces considerable control complexity, which, if unaddressed, can result in prohibitive signaling overhead. How to design an effective ssRIS-assisted system that navigates these constraints while delivering meaningful performance gains over SMS and RIS thus constitutes the central problem investigated in this chapter.

This chapter presents the work in **Paper B**, which directly addresses research question **RQ2**:

How can an ssRIS-assisted mmWave communication system in IDS be designed to maximize the minimum data rate of all UEs while satisfying the self-sustainability constraint? How can RIS, SMS, and ssRIS be positioned within the trade-off among operating cost, metasurface gain, and system coverage?

This chapter proposes a low-complexity ssRIS-assisted system design that jointly optimizes ssRIS-UE association, phase shifts, and time resource allocation via a two-stage iterative algorithm, and positions ssRIS within the trade-off among metasurface gain, operating cost, and system coverage through comparison with RIS and SMS. Our numerical results have shown that ssRIS can provide an intermediate performance between SMS and RIS while minimizing the operating cost to fit the demand for IDS. Although its coverage range is limited by the self-sustainability, it has been verified to be a better compromise within the trade-off.

## 4.1 Related Work

Existing research on ssRIS primarily focuses on improving system performance under low power constraints. [8] maximizes the sum rate in a multi-user MISO downlink by jointly designing beamforming, phase shifts, and energy harvesting. [9] enhances wireless power transfer efficiency using an absorb-then-reflect scheme. [10] minimizes transmit power by jointly optimizing beamforming, phase shifts, and reflection amplitude while satisfying receiver SNR and ssRIS energy constraints. [13] introduces the energy-data rate outage probability metric and evaluates HaR protocols under different ssRIS placements to minimize this metric.

Despite these advances, the comparative performance of ssRIS against RIS and SMS has not been thoroughly investigated, and evaluations that account for the practical self-sustainability constraint in realistic deployment environments remain scarce. Furthermore, recent advances in RIS prototyping [33, 34] and explicit studies of metasurface energy harvesting [35] have established a solid foundation for ssRIS realization. Against this backdrop, a systematic investigation of ssRIS-assisted system design and comparison with other types of metasurface is both timely and necessary.

## 4.2 Contributions

The main contributions of this work are as follows:

- We present a large-scale ssRIS-assisted system designed with low computational complexity for optimization. In this system, the utilization of the ssRIS is simplified by following a pre-designed pattern and by considering the associations between the ssRISs and UEs based on the feasibility of achieving self-sustainability.
- We formulate a MIP optimization problem to maximize the minimum achievable data rate in the system by jointly determining the ssRIS-UE association, phase shifts of the ssRISs, and allocating time resources for each UE. To tackle the non-convexity inherent in the optimization problem, the FPP-SCA is employed to convexify the constraints. Moreover, we propose a two-stage iterative algorithm based on the convexified optimization problem.
- We investigate the feasibility and necessity of utilizing ssRIS by checking the ssRIS contribution, the ssRIS-UE association, and the self-sustainability condition for ssRISs in different positions.
- Extending Paper A [1], where we investigate the RIS- and SMS-assisted system, in this work, we focus on ssRIS, which is reconfigurable yet requires a much

lower operating cost than RIS. To assess the trade-offs between metasurface gain, operating costs, and system coverage, we conduct a thorough comparison of the resulting data rates using ssRIS, RIS, and SMS across indoor environments with varying sizes. Through our research, we provide a fresh perspective on better scenarios and methodologies for deploying ssRIS over other types of metasurfaces.

### 4.3 System Model

We consider the downlink of an mmWave communication system in which a single BS serves multiple UEs distributed across an IDS. The BS is equipped with  $N$  antennas arranged as a  $\sqrt{N} \times \sqrt{N}$  UPA, while each UE is equipped with a single antenna. The UEs are indexed by  $k \in \{1, 2, \dots, K\} = \mathcal{K}$ , where  $K$  denotes the total number of UEs. Due to the short and largely unobstructed propagation paths via the metasurface in an IDS, the resulting channels are of low rank, which favors orthogonal time-domain resource allocation as an effective interference mitigation strategy. Accordingly, each UE  $k$  is served during an exclusively allocated time portion  $\tau_k$ , with  $\sum_{k=1}^K \tau_k = 1$ . To account for hardware limitations on the phase-shift switching speed of the metasurface, the allocated time portion is bounded below by a minimum value  $\tau_{\min}$ , such that  $\tau_k \in [\tau_{\min}, 1]$  holds for all  $k$ , where  $0 < \tau_{\min} < 1$ .

To improve the downlink rates of the UEs,  $L$  ssRIS are deployed throughout the IDS, each comprising  $M$  elements arranged as a  $\sqrt{M} \times \sqrt{M}$  UPA. The element and surface indices are denoted as  $m \in \{1, 2, \dots, M\} = \mathcal{M}$  and  $l \in \{1, 2, \dots, L\} = \mathcal{L}$ , respectively. As illustrated in Figure 4.3.1, the direct channel from the BS to UE  $k$  is denoted as  $\mathbf{h}_k \in \mathbb{C}^N$ , while  $\mathbf{G}_l \in \mathbb{C}^{M \times N}$  and  $\mathbf{g}_{l,k} \in \mathbb{C}^M$  denote the BS-to-ssRIS and ssRIS-to-UE channels, respectively. Given the geometry-dependent nature of mmWave propagation and the restricted UE mobility in IDS environments, all channels are modeled as deterministic and fixed [36]. Realistic channel coefficients are subsequently obtained through RT simulations of a specific IDS environment.

To maintain self-sustainability, the ssRIS must operate under a harvest-and-reflect (HaR) protocol that governs how its elements are allocated between energy harvesting and signal reflection. Among the candidate HaR schemes, time splitting and element splitting have been shown to exhibit better resilience to varying propagation conditions [13]. In our setting, however, time resources are already shared among a densely populated UE set, and adopting time splitting would render the time slots available for phase-shift reconfiguration unrealistically short for practical hardware implementation. We therefore adopt the element splitting scheme, wherein each element independently operates in either reflecting or harvesting mode.

As illustrated in Figure 4.3.2, under the element splitting scheme, each element of the ssRIS selects its working mode between *reflecting* and *harvesting*. Reconfiguring element working modes and phase shifts on a per-UE basis would require highly accurate synchronization and fast hardware response, resulting in prohibitive control complexity in a densely populated IDS. To avoid this, we assume that the working mode configuration of each ssRIS remains fixed across all UE service intervals. The working mode of metasurface  $l$  is described by the indicator vector  $\beta_l \in \mathbb{B}^M$ , where the  $m$ -th entry  $\beta_{l,m} = 1$  indicates that element  $m$  operates in reflecting mode, and  $\beta_{l,m} = 0$  indicates

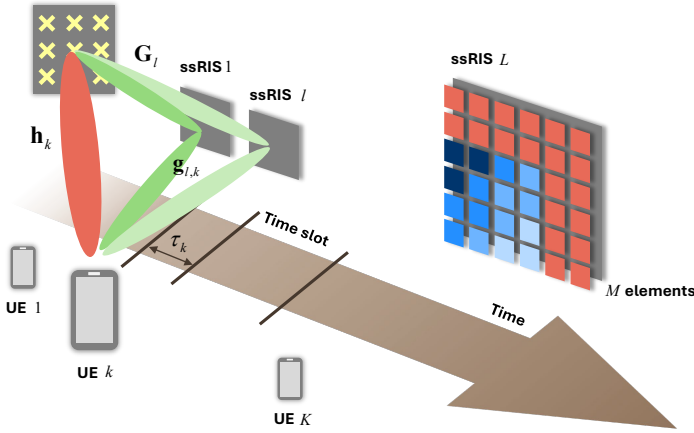


Figure 4.3.1: Illustration of the ssRIS-assisted system model.

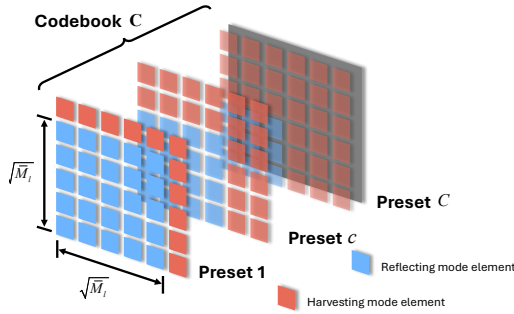


Figure 4.3.2: Illustration of the ssRIS with element splitting scheme and element working mode selected based on presets.

harvesting mode.

Since the ssRIS must be positioned in the LoS region of the BS to ensure sufficient received power for self-sustainability, and given that mmWave propagation is dominated by the LoS component under far-field conditions, the signal strength across different elements of the same ssRIS can be considered approximately uniform. Consequently, the reflection performance is primarily determined by the total number of reflecting elements rather than their specific arrangement. This observation motivates the use of a *preset*-based working mode configuration, in which  $\beta_l$  is selected from a predefined *codebook*  $\mathbf{C} \in \mathbb{B}^{M \times C}$  via a preset selector  $\xi_l \in \mathbb{B}^C$ , such that  $\beta_l = \mathbf{C}\xi_l$ . Each of the  $C$  columns of  $\mathbf{C}$  represents a distinct preset corresponding to a different number of reflecting elements. For the selected preset, the  $\bar{M}_l$  reflecting elements of ssRIS  $l$  are arranged contiguously in the form of a  $\sqrt{\bar{M}_l} \times \sqrt{\bar{M}_l}$  UPA, so that  $\sum_{m=1}^M \beta_{l,m} = \bar{M}_l$  holds for all  $l$ .

An element operating in reflecting mode is configured to apply a phase shift to the impinging signal, consuming a per-element power of  $P^{\text{Rf}}$ . An element in harvesting mode, by contrast, captures the energy of the impinging signal through its harvesting circuit, with no signal reflected back. The resulting response vector of metasurface  $l$  when serving UE  $k$  is given by  $\phi_{l,k} \in \mathbb{C}^M$  as

$$\phi_{l,k} = \text{diag}(\varphi_{l,k})\beta_l, \quad (4.3.1)$$

where  $\varphi_{l,k} \in \mathbb{C}^M$  is the phase-shift response vector. Neglecting reflection loss, the unit-modulus constraint  $|\varphi_{l,k,m}| = 1$  holds for all  $l, k, m$ , where  $\varphi_{l,k,m}$  denotes the  $m$ -th entry of  $\varphi_{l,k}$ .

For notational convenience, we define the aggregated BS-metasurface-UE channel from the BS via metasurface  $l$  to UE  $k$  as  $\mathbf{H}_{l,k} = \mathbf{G}_l^T \bar{\mathbf{G}}_{l,k} \in \mathbb{C}^{N \times M}$ , where  $\bar{\mathbf{G}}_{l,k} = \text{diag}(\mathbf{g}_{l,k})$ . The received signal at UE  $k$  is then expressed as

$$y_k = \left( \mathbf{h}_k + \sum_{l=1}^L \underbrace{\mathbf{G}_l^T \bar{\mathbf{G}}_{l,k}}_{\triangleq \mathbf{H}_{l,k}} \phi_{l,k} \right)^T \mathbf{w}_k x_k + n_k, \quad (4.3.2)$$

where  $x_k \in \mathbb{C}$  is the transmitted data symbol with power  $\mathbb{E}\{|x_k|^2\} = P$ ,  $\mathbf{w}_k \in \mathbb{C}^N$  is the unit-norm precoding vector, and  $n_k \sim \mathcal{CN}(0, BN_0)$  is the receiver noise with bandwidth  $B$  and noise spectral density  $N_0$ .

Under orthogonal UE scheduling, there is no inter-user interference, and MRT is the SNR-optimal precoding strategy. The MRT precoder for UE  $k$  is given by

$$\mathbf{w}_k = \frac{\left( \mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \phi_{l,k} \right)^*}{\left\| \mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \phi_{l,k} \right\|}, \quad (4.3.3)$$

which yields the received SNR

$$\Gamma_k = \frac{\left\| \mathbf{h}_k + \sum_{l=1}^L \mathbf{H}_{l,k} \phi_{l,k} \right\|^2 P}{BN_0}, \quad (4.3.4)$$

and the achievable data rate

$$R_k = \tau_k B \log_2(1 + \Gamma_k) \quad \text{bit/s}. \quad (4.3.5)$$

Following the element splitting-based HaR, to achieve self-sustainability, ssRIS should harvest power that is no less than what it consumes from the impinging radio waves to support its operation. The signals that are received by the harvesting elements of the metasurface  $l$  when the system is serving UE  $k$  can be expressed as

$$\mathbf{y}_{l,k}^{\text{Rc}} = \text{diag}(\mathbf{1}_M - \beta_l) \mathbf{G}_l \mathbf{w}_k x_k \quad (4.3.6)$$

and the power received by the metasurface  $l$  when the system is serving UE  $k$  is given as

$$P_{l,k}^{\text{Rc}} = \mathbb{E} \left\{ \|\mathbf{y}_{l,k}^{\text{Rc}}\|^2 \right\} = \sum_{m=1}^M (1 - \beta_{l,m}) \left| \bar{\mathbf{g}}_{l,m}^T \mathbf{w}_k \right|^2 P. \quad (4.3.7)$$

Here, the power received from the thermal noise is neglected. Moreover,  $\bar{\mathbf{g}}_{l,m}^T \in \mathbb{C}^{1 \times N}$  denotes the  $m$ -th row of  $\mathbf{G}_l$ . Note that due to the dependency between the selection of  $\phi_{l,k}$  and the beamforming vector  $\mathbf{w}_k$ , according to (4.3.7), the received power will also depend on the selection of phase shifts.

The harvested power follows the non-linear model (2.3.2) discussed in Section 2.3.1 for its accuracy in capturing the realistic hardware efficiency while maintaining a good mathematical tractability.

#### 4.4 Problem Formulation

To ensure fairness among all UEs, the system is designed to maximize the minimum achievable data rate in the IDS. This is accomplished by jointly tuning the phase shifts of the metasurfaces, selecting the preset configuration for each ssRIS, and allocating time resources across UEs. The resulting max-min rate optimization problem is formulated as

$$\mathbf{P4.1} : \text{maximize } \min_{\{\varphi_{l,k}, \xi_l, \tau_k\}_{k \in \mathcal{K}}} R_k \quad (4.4.8a)$$

$$\text{s.t. } \sum_{k=1}^K \tau_k \leq 1, \quad (4.4.8b)$$

$$\tau_k \geq \tau_{\min}, \quad \forall k, \quad (4.4.8c)$$

$$|\varphi_{l,k,m}| = 1, \quad \forall l, k, m, \quad (4.4.8d)$$

$$\xi_{l,c} \in \{0, 1\}, \quad \forall l, c, \quad (4.4.8e)$$

$$\sum_{c=1}^C \xi_{l,c} = 1, \quad \forall l, \quad (4.4.8f)$$

$$P_{l,k}^{\text{Hr}} \geq \sum_{m=1}^M \beta_{l,m} P^{\text{Rf}}, \quad \forall l, k. \quad (4.4.8g)$$

The objective in (4.4.8a) maximizes the minimum data rate achieved across all UEs. The constraints in (4.4.8b) and (4.4.8c) govern the time resource allocation, ensuring that the sum of all allocated time portions does not exceed one and that each individual portion remains no shorter than the minimum allowable value  $\tau_{\min}$ , respectively. The unit-modulus constraint in (4.4.8d) enforces lossless reflection at each metasurface element. The constraints in (4.4.8e) and (4.4.8f) jointly regulate the preset selection, requiring the selector to be binary and ensuring that exactly one preset is active per ssRIS. Finally, the self-sustainability of each ssRIS is enforced by the constraints in (4.4.8g). The presence of binary variables renders **P4.1** a MIP problem.

#### 4.4.1 Simplifying the utilization of ssRIS

Although the non-convexity of **P4.1** presents its own challenges, the more fundamental obstacle to solving it directly is computational complexity. The number of binary variables in **P4.1**, such as the preset selectors  $\xi_{l,c}$ , scales with the number of ssRISs and UEs in the network, and the worst-case complexity of MIP grows exponentially with the number of such variables [37]. To address this, we decompose **P4.1** into smaller subproblems through two heuristic simplifications.

Given the high propagation attenuation of mmWave signals and the obstacle-rich nature of the IDS, an ssRIS can only meaningfully serve UEs in its surrounding area. Moreover, due to the highly directive nature of MRT beamforming and the rank-limited nature of IDS, an ssRIS serving an angularly separate UE is unlikely to harvest sufficient energy to maintain self-sustainability. We therefore introduce a binary ssRIS-UE association indicator  $\alpha_{l,k} \in \mathbb{B}$ , where  $\alpha_{l,k} = 1$  indicates that ssRIS  $l$  is assigned to serve UE  $k$ . Rather than optimizing all possible associations jointly, we first partition the network into *coverage groups* by manually assigning each ssRIS to the UEs in its nearby area. The fine-grained ssRIS-UE association within each group is then determined as part of the joint optimization. Coverage group  $i \in \{1, \dots, I\}$  contains the metasurface and UE index sets  $\mathcal{L}_i \subseteq \mathcal{L}$  and  $\mathcal{K}_i \subseteq \mathcal{K}$ , with  $L_i$  surfaces and  $K_i$  UEs, respectively. Note that the same ssRIS may appear in multiple coverage groups, so the sets  $\mathcal{L}_i$  and  $\mathcal{L}_j$  are not necessarily disjoint for  $i \neq j$ . The SNR of UE  $k$  in coverage group  $i$  is then expressed as

$$\tilde{\Gamma}_k = \frac{\|\mathbf{h}_k + \sum_{l \in \mathcal{L}_i} \alpha_{l,k} \mathbf{H}_{l,k} \phi_{l,k}\|^2 P}{BN_0}. \quad (4.4.9)$$

Since the self-sustainability of each ssRIS depends on the MRT beamformer, which in turn depends on the phase shifts, jointly optimizing the preset selectors and phase shifts is computationally complicated. To decouple these, we determine the preset for each ssRIS separately by evaluating self-sustainability under a randomized phase-shift assumption. Specifically, the MRT beamformer  $\hat{\mathbf{w}}_k$  is computed with randomized ssRIS responses  $\hat{\phi}_{l,k} = e^{-j\hat{\vartheta}_{l,k}}$ , where the entries of  $\hat{\vartheta}_{l,k}$  are drawn uniformly from  $(0, 2\pi]$ . Under this beamformer, the maximum feasible reflecting element count  $M_l$  for each ssRIS is identified from the codebook as the largest preset that satisfies the self-sustainability constraint while serving the UEs in its coverage group.

#### 4.4.2 Joint data rate optimization algorithm

With  $\mathcal{L}_i$ ,  $\mathcal{K}_i$ , and  $\beta_l$  treated as known quantities, **P4.1** is decomposed into two subproblems. Subproblem **P4.2.1** maximizes the minimum data rate within each coverage group by jointly optimizing the phase shifts, ssRIS-UE association, and intra-group time allocation  $\tilde{\tau}_k$ . When  $\alpha_{l,k} = 0$ , the self-sustainability constraint is automatically satisfied, as unassigned ssRISs consume no power. Denoting the intra-group data rate

as  $\tilde{R}_k(\phi_{l,k}, \alpha_{l,k}, \tilde{\tau}_k) = \tilde{\tau}_k B \log_2(1 + \tilde{\Gamma}_k)$ , subproblem **P4.2.1** is formulated as

$$\mathbf{P4.2.1} : \text{maximize } \min_{\{\varphi_{l,k}, \alpha_{l,k}, \tilde{\tau}_k\}_{k \in \mathcal{K}_i}} \tilde{R}_k \quad (4.4.10a)$$

$$\text{s.t. } \sum_{k \in \mathcal{K}_i} \tilde{\tau}_k \leq 1, \quad (4.4.10b)$$

$$\tilde{\tau}_k \geq \tau_{\min}, \quad \forall k \in \mathcal{K}_i, \quad (4.4.10c)$$

$$|\varphi_{l,k,m}| = 1, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{K}_i, \forall m, \quad (4.4.10d)$$

$$\alpha_{l,k} \in \{0, 1\}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{K}_i, \quad (4.4.10e)$$

$$P_{l,k}^{\text{Hr}} \geq \alpha_{l,k} \bar{M}_l P^{\text{Rf}}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{K}_i. \quad (4.4.10f)$$

Subproblem **P4.2.2** then allocates time resources system-wide to maximize the minimum data rate across all UEs. Given the solutions  $\phi_{l,k}^*$ ,  $\alpha_{l,k}^*$ , and  $\tilde{\tau}_k^*$  obtained from **P4.2.1** for each coverage group, the system-wide data rate for each UE is expressed as  $R_k = \hat{\tau}_k \tilde{R}_k(\phi_{l,k}^*, \alpha_{l,k}^*, \tilde{\tau}_k^*)$ , where  $\hat{\tau}_k \in \mathbb{R}$  is the system-wide time allocation. Subproblem **P4.2.2** is then given as

$$\mathbf{P4.2.2} : \text{maximize } \min_{\{\hat{\tau}_k\}} R_k \quad (4.4.11a)$$

$$\text{s.t. } \hat{\tau}_k \tilde{\tau}_k^* \geq \tau_{\min}, \quad \forall k, \quad (4.4.11b)$$

$$\sum_{k \in \mathcal{K}} \hat{\tau}_k \tilde{\tau}_k^* \leq 1. \quad (4.4.11c)$$

Unlike **P4.2.1**, subproblem **P4.2.2** is a linear programming (LP) problem and can be solved efficiently by any standard convex solver.

Problem **P4.2.1** is non-convex due to multiple sources of coupling. The binary association indicator  $\alpha_{l,k}$  is coupled with the complex phase-shift variables  $\varphi_{l,k}$  in the expression of  $\tilde{R}_k$ , while the quadratic form inside the logarithm and its coupling with  $\tilde{\tau}_k$  further preclude convexity. In the self-sustainability constraint (4.4.10f), the MRT precoder introduces an additional non-convex coupling between  $\alpha_{l,k}$  and  $\phi_{l,k}$ . To render **P4.2.1** tractable, we introduce the auxiliary variable  $\tilde{R}_i^{\min}$  representing the minimum intra-group data rate and recast it as

$$\mathbf{P4.3} : \text{maximize } \tilde{R}_i^{\min} \quad (4.4.12a)$$

$$\{\varphi_{l,k}, \alpha_{l,k}, \tilde{\tau}_k, \tilde{R}_i^{\min}\}$$

$$\text{s.t. } \tilde{R}_k(\phi_{l,k}, \alpha_{l,k}, \tilde{\tau}_k) \geq \tilde{R}_i^{\min}, \quad \forall k \in \mathcal{K}_i, \quad (4.4.12b)$$

$$(4.4.10b), (4.4.10c), (4.4.10d), (4.4.10e), (4.4.10f). \quad (4.4.12c)$$

To decouple  $\alpha_{l,k}$  from  $\phi_{l,k}$ , we introduce an auxiliary variable  $\mathbf{z}_{l,k} \in \mathbb{C}^M$  to replace the product  $\alpha_{l,k} \phi_{l,k}$ , subject to

$$\left| [\mathbf{z}_{l,k}]_m \right| \leq \alpha_{l,k}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{K}_i, \forall m. \quad (4.4.13)$$

When  $\alpha_{l,k} = 1$ , each entry of  $\mathbf{z}_{l,k}$  is bounded within the unit circle, and the solution is empirically observed to lie on the unit circle at convergence, thereby satisfying the unit-modulus constraint (4.4.10d). When  $\alpha_{l,k} = 0$ , the constraint forces  $\mathbf{z}_{l,k} = \mathbf{0}_M$ , effectively canceling the contribution of ssRIS  $l$  to UE  $k$ .

We next address the non-convexities in (4.4.12b). Substituting  $\tilde{\tau}_i^{\min} = \sqrt{\tilde{R}_i^{\min}}$  in place of  $\tilde{R}_i^{\min}$ , which is valid since the square root is monotonically increasing, and reformulating (4.4.12b) in terms of SNR yields

$$1 + \frac{\|\mathbf{h}_k + \sum_{l \in \mathcal{L}_i} \mathbf{H}_{l,k} \mathbf{z}_{l,k}\|^2 P}{BN_0} \geq 2 \frac{(\tilde{\tau}_i^{\min})^2}{B\tilde{\tau}_k}, \quad \forall k \in \mathcal{K}_i. \quad (4.4.14)$$

Since both sides of (4.4.14) contain coupled variables, we introduce an auxiliary variable  $e_k \in \mathbb{R}$  to separate them as

$$e_k \geq 2 \frac{(\tilde{\tau}_i^{\min})^2}{B\tilde{\tau}_k}, \quad \forall k \in \mathcal{K}_i, \quad (4.4.15)$$

$$1 + \frac{\|\mathbf{h}_k + \sum_{l \in \mathcal{L}_i} \mathbf{H}_{l,k} \mathbf{z}_{l,k}\|^2 P}{BN_0} \geq e_k, \quad \forall k \in \mathcal{K}_i. \quad (4.4.16)$$

To address the quadratic-over-linear term in (4.4.15), a further auxiliary variable  $f_k \in \mathbb{R}$  is introduced, splitting the constraint into

$$e_k \geq 2f_k, \quad \forall k \in \mathcal{K}_i, \quad (4.4.17)$$

$$f_k \geq \frac{(\tilde{\tau}_i^{\min})^2}{B\tilde{\tau}_k}, \quad \forall k \in \mathcal{K}_i. \quad (4.4.18)$$

The constraint (4.4.17) is a convex exponential cone constraint. The constraint (4.4.18) can be equivalently expressed as

$$f_k^2 + 2f_k\tilde{\tau}_k + \tilde{\tau}_k^2 \geq \left( \sqrt{\frac{2}{B}} \tilde{\tau}_i^{\min} \right)^2 + f_k^2 + \tilde{\tau}_k^2, \quad \forall k \in \mathcal{K}_i, \quad (4.4.19)$$

and further reformulated as a SOC constraint

$$f_k + \tilde{\tau}_k \geq \left\| \begin{bmatrix} \sqrt{2/B} \tilde{\tau}_i^{\min} & f_k & \tilde{\tau}_k \end{bmatrix} \right\|, \quad \forall k \in \mathcal{K}_i, \quad (4.4.20)$$

by which the coupling with  $\tilde{\tau}_k$  is fully resolved. The remaining non-convexity in (4.4.16) stems from the quadratic norm on the left-hand side. Expanding this term gives

$$\mathbf{z}_k^H \underbrace{\mathbf{H}_k^H \mathbf{H}_k}_{\triangleq \mathbf{A}_k} \mathbf{z}_k + 2\Re \left( \mathbf{z}_k^H \underbrace{\mathbf{H}_k^H \mathbf{h}_k}_{\triangleq \mathbf{b}_k} \right) + \underbrace{\mathbf{h}_k^H \mathbf{h}_k}_{\triangleq c_k} \geq \frac{BN_0(e_k - 1)}{P}, \quad \forall k \in \mathcal{K}_i, \quad (4.4.21)$$

where  $\mathbf{H}_k = [\mathbf{H}_{l_{i,1},k}, \dots, \mathbf{H}_{l_{i,L_i},k}] \in \mathbb{C}^{N \times L_i M}$ ,  $\mathbf{z}_k = [\mathbf{z}_{l_{i,1},k}^T, \dots, \mathbf{z}_{l_{i,L_i},k}^T]^T \in \mathbb{C}^{L_i M}$ , and  $l_{i,j}$  denotes the index of the  $j$ -th element in  $\mathcal{L}_i$ . Since  $\mathbf{A}_k$  is PSD, the FPP-SCA method [25] can be applied to obtain a local affine upper bound. For any vector  $\zeta_k$ ,

$$\mathbf{z}_k^H (-\mathbf{A}_k) \mathbf{z}_k \leq 2\Re(\zeta_k^H (-\mathbf{A}_k) \mathbf{z}_k) - \zeta_k^H (-\mathbf{A}_k) \zeta_k, \quad \forall k \in \mathcal{X}_i. \quad (4.4.22)$$

At each iteration  $\epsilon$ , substituting  $\zeta_k = \mathbf{z}_k^{(\epsilon-1)}$  and introducing a non-negative slack variable  $s_k$  to ensure feasibility throughout the iterative process, the quadratic term in (4.4.16) is replaced by its affine approximation as

$$\begin{aligned} & -2\Re\left(\left(\mathbf{z}_k^{(\epsilon-1)}\right)^H \mathbf{A}_k \mathbf{z}_k\right) + \left(\mathbf{z}_k^{(\epsilon-1)}\right)^H \mathbf{A}_k \mathbf{z}_k^{(\epsilon-1)} - 2\Re(\mathbf{z}_k^H \mathbf{b}_k) \\ & \leq s_k + c_k + \frac{BN_0}{P}(1 - e_k), \quad \forall k \in \mathcal{X}_i, \end{aligned} \quad (4.4.23)$$

$$s_k \geq 0, \quad \forall k \in \mathcal{X}_i. \quad (4.4.24)$$

To handle the non-convexity in the self-sustainability constraint (4.4.10f), we introduce an auxiliary variable  $u_{l,k} \in \mathbb{R}$  that is bounded between the consumed and harvested power, decomposing (4.4.10f) into

$$\alpha_{l,k} \bar{M}_l P^{\text{Rf}} \leq u_{l,k}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \quad (4.4.25)$$

$$\frac{q_1 \sum_{m=1}^M (1 - \beta_{l,m}) |\bar{\mathbf{g}}_{l,m}^T \mathbf{w}_k|^2 P}{q_2 \sum_{m=1}^M (1 - \beta_{l,m}) |\bar{\mathbf{g}}_{l,m}^T \mathbf{w}_k|^2 P + q_3} \geq u_{l,k}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i. \quad (4.4.26)$$

Since  $\mathbf{w}_k$  depends non-convexly on  $\mathbf{z}_{l,k}$ , the constraint (4.4.26) is not convex. To resolve this, we multiply both the numerator and denominator on the left-hand side of (4.4.26) by  $\|\mathbf{h}_k + \mathbf{H}_k \mathbf{z}_k\|^2$ , and introduce auxiliary variables  $t_{l,k,m} \in \mathbb{R}$  and  $d_k \in \mathbb{R}$  with

$$|\bar{\mathbf{g}}_{l,m}^H (\mathbf{h}_k + \mathbf{H}_k \mathbf{z}_k)|^2 P \geq t_{l,k,m}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \forall m, \quad (4.4.27)$$

$$\|\mathbf{h}_k + \mathbf{H}_k \mathbf{z}_k\|^2 \leq d_k, \quad \forall k \in \mathcal{X}_i, \quad (4.4.28)$$

$$t_{l,k,m} \leq \mathcal{M}(1 - \beta_{l,m}), \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \forall m, \quad (4.4.29)$$

where  $\mathcal{M}$  is a sufficiently large constant. The constraint (4.4.29) ensures that  $t_{l,k,m} = 0$  whenever  $\beta_{l,m} = 1$ , preventing reflecting elements from contributing to the harvested power computation. With these substitutions, (4.4.26) reduces to

$$\frac{q_1 \sum_{m=1}^M t_{l,k,m}}{q_2 \sum_{m=1}^M t_{l,k,m} + q_3 d_k} \geq u_{l,k}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i. \quad (4.4.30)$$

Residual non-convexities remain in (4.4.27) and (4.4.30) and are addressed in turn. The quadratic constraint (4.4.27) takes the same form as (4.4.16) and is handled analogously via FPP-SCA. Expanding the quadratic term gives

$$\begin{aligned} & \underbrace{\mathbf{h}_k^H \bar{\mathbf{g}}_{l,m} \bar{\mathbf{g}}_{l,m}^H \mathbf{h}_k}_{\triangleq \hat{c}_{l,k,m}} + 2\Re\left(\underbrace{\mathbf{z}_k^H \mathbf{H}_k^H \bar{\mathbf{g}}_{l,m} \bar{\mathbf{g}}_{l,m}^H \mathbf{h}_k}_{\triangleq \hat{\mathbf{b}}_{l,k,m}}\right) \\ & + \underbrace{\mathbf{z}_k^H \mathbf{H}_k^H \bar{\mathbf{g}}_{l,m} \bar{\mathbf{g}}_{l,m}^H \mathbf{H}_k \mathbf{z}_k}_{\triangleq \hat{\mathbf{A}}_{l,k,m}} \geq \frac{t_{l,k,m}}{P}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \forall m, \end{aligned} \quad (4.4.31)$$

which is then linearized via FPP-SCA with slack variable  $s'_{l,k,m} \geq 0$  as

$$-2\Re \left( \left( \mathbf{z}_k^{(\epsilon-1)} \right)^H \hat{\mathbf{A}}_{l,k,m} \mathbf{z}_k \right) + \left( \mathbf{z}_k^{(\epsilon-1)} \right)^H \hat{\mathbf{A}}_{l,k,m} \mathbf{z}_k^{(\epsilon-1)} - 2\Re \left( \mathbf{z}_k^H \hat{\mathbf{b}}_{l,k,m} \right) \leq s'_{l,k,m} + \hat{c}_{l,k,m} - \frac{t_{l,k,m}}{P}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \forall m, \quad (4.4.32)$$

$$s'_{l,k,m} \geq 0, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \forall m. \quad (4.4.33)$$

For the fractional constraint (4.4.30), we first reformulate it equivalently as

$$\left( u_{l,k} + q_2 \sum_{m=1}^M t_{l,k,m} + q_3 d_k \right)^2 \leq 2q_1 \sum_{m=1}^M t_{l,k,m} + u_{l,k}^2 + \left( q_2 \sum_{m=1}^M t_{l,k,m} + q_3 d_k \right)^2, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i. \quad (4.4.34)$$

Introducing  $v_{l,k} \in \mathbb{R}$  defined as

$$v_{l,k} = q_2 \sum_{m=1}^M t_{l,k,m} + q_3 d_k, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \quad (4.4.35)$$

the constraint (4.4.34) simplifies to

$$(u_{l,k} + v_{l,k})^2 \leq 2q_1 \sum_{m=1}^M t_{l,k,m} + u_{l,k}^2 + v_{l,k}^2, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i. \quad (4.4.36)$$

The non-convex quadratic terms  $u_{l,k}^2$  and  $v_{l,k}^2$  on the right-hand side are linearized via first-order Taylor expansion at the previous iterate, with slack variable  $s''_{l,k} \geq 0$  introduced for feasibility

$$(u_{l,k} + v_{l,k})^2 \leq 2q_1 \sum_{m=1}^M t_{l,k,m} + \left( u_{l,k}^{(\epsilon-1)} \right)^2 + 2u_{l,k}^{(\epsilon-1)} \left( u_{l,k} - u_{l,k}^{(\epsilon-1)} \right) + \left( v_{l,k}^{(\epsilon-1)} \right)^2 + 2v_{l,k}^{(\epsilon-1)} \left( v_{l,k} - v_{l,k}^{(\epsilon-1)} \right) + s''_{l,k}, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i, \quad (4.4.37)$$

$$s''_{l,k} \geq 0, \quad \forall l \in \mathcal{L}_i, \forall k \in \mathcal{X}_i. \quad (4.4.38)$$

With all non-convexities resolved except for the binary constraint (4.4.10e), which is handled by the MOSEK MIP solver via branch-and-bound [38], the fully solvable problem at each iteration is formulated as

$$\mathbf{P4.4} : \quad \begin{array}{llll} \text{maximize} & & \tilde{r}_i^{\min} - \bar{s}_i & \\ \left\{ \begin{array}{l} \mathbf{z}_{l,k} \\ f_k \\ v_{l,k} \\ \tilde{r}_i^{\min} \end{array} \right. & \begin{array}{llll} \alpha_{l,k} & \tilde{r}_k & e_k \\ u_{l,k} & t_{l,k,m} & d_k \\ s_k & s'_{l,k,m} & s''_{l,k} \end{array} & \end{array} \quad (4.4.39)$$

s.t. (4.4.10b), (4.4.10c), (4.4.10e), (4.4.13), (4.4.17), (4.4.20), (4.4.23), (4.4.24), (4.4.25), (4.4.28), (4.4.29), (4.4.32), (4.4.33), (4.4.37), (4.4.38).

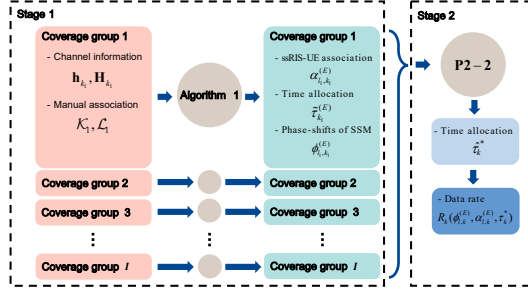


Figure 4.4.3: Flowchart of the two-stage iterative data rate optimization algorithm.

The penalty term  $\bar{s}_i$  aggregates the weighted slack variables introduced throughout the convexification process, driving the iterative solution toward feasibility of the original problem. With penalty coefficients  $\omega, \omega',$  and  $\omega''$  for  $s_k, s'_{l,k,m},$  and  $s''_{l,k}$  respectively, it is defined as

$$\bar{s}_i = \omega \sum_{k \in \mathcal{K}_i} s_k + \omega' \sum_{l \in \mathcal{L}_i} \sum_{k \in \mathcal{K}_i} \sum_{m=1}^M s'_{l,k,m} + \omega'' \sum_{l \in \mathcal{L}_i} \sum_{k \in \mathcal{K}_i} s''_{l,k}. \quad (4.4.40)$$

The complete two-stage iterative optimization algorithm is illustrated in Figure 4.4.3. In the first stage, Algorithm 3 is applied independently to each coverage group to obtain the optimized intra-group rates  $\tilde{R}_k$  for all  $k$ . In the second stage, these results are passed to **P4.2**, which allocates time resources system-wide to maximize the minimum data rate  $R_k$  across the entire network.

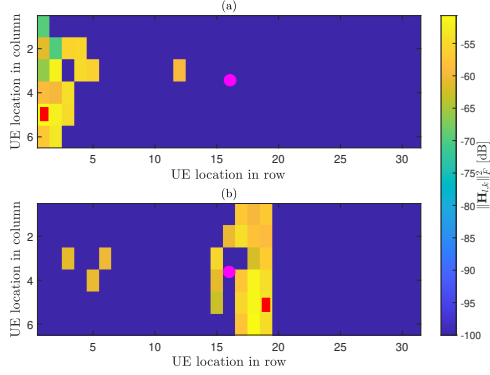
---

**Algorithm 3** Joint coverage group minimum data rate maximization algorithm

---

- 1: **Given:** The manual ssRIS-UE association  $\mathcal{K}_i, \mathcal{L}_i$ . The channel information of the coverage group  $\mathbf{h}_k, \mathbf{H}_k$
  - 2: **Initialization:** Initialize  $\mathbf{z}_k^{(0)}$  randomly while keeping  $|\mathbf{z}_{l,k,m}^{(0)}| = 1, \forall l, k, m$ . Set the iteration counter to  $\epsilon = 0$ . Set the maximum iteration number to  $E$ .
  - 3: **while**  $\epsilon < E$  **do**
  - 4:      $\epsilon \leftarrow \epsilon + 1$
  - 5:     Solve **P4.4** and set  $\mathbf{z}_k^{(\epsilon)}, \forall k \in \mathcal{K}_i$  to its solution
  - 6: **end while**
  - 7: **Output:** The ssRIS-UE association  $\alpha_{l,k}^{(E)}$ , phase-shift configurations  $\phi_{l,k}^{(E)}$ , and allocated time portions  $\tau_k^{(E)}$
- 

In the considered IDS, the restricted mobility of passengers results in channel and beam coherence times that substantially exceed typical communication frame durations, allowing the ssRIS configuration to remain valid over extended periods without loss of optimality. Under this framework, Algorithm 3 is executed either during the initial system setup or at periodic intervals dictated by the channel coherence time. Once



**Figure 4.5.4:** The cascaded BS-ssRIS-UE channel gain with respect to a ssRIS that is (a) far-away (b) near to the BS. The red rectangle represents the checked ssRIS, and the pink circle represents the BS.

configured, the ssRIS applies a fixed phase-shift pattern and cycles through UE associations according to the optimized time allocation schedule provided by the BS, with no need for continuous real-time reconfiguration. Since all phase-shift configurations are prescribed by the BS, the ssRIS requires no onboard channel estimation or signal processing, substantially reducing both control signaling overhead and hardware complexity at the surface and supporting the practical feasibility of large-scale deployment in IDS.

## 4.5 Selective Results

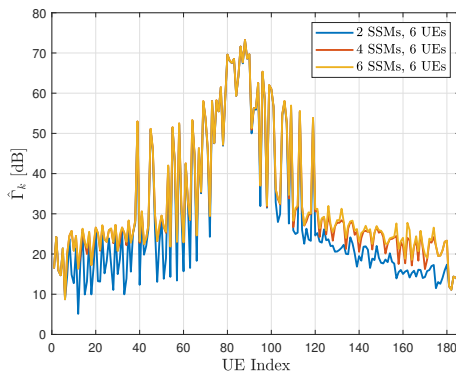
### 4.5.1 Coverage group design

To characterize the spatial coverage of an ssRIS in the considered environment, the cascaded BS-ssRIS-UE channel  $\mathbf{H}_{l,k}$  is evaluated for all ssRISs against every UE in the network. Owing to the high geometric symmetry of the IDS, all ssRISs exhibit similar coverage patterns regardless of their position. As shown in Figure 4.5.4, approximately three rows of UEs in front of the reflective side are effectively covered, consistently across both near and far ssRIS locations. Based on this observation, coverage groups are formed by associating one row of UEs with their nearest one, two, or three rows of ssRISs, with each group indexed by the row location of the UEs. The UE row is always placed at the reflective side of the associated ssRISs. For instance, as illustrated in Figure 4.5.5, when two rows of ssRISs are grouped with one row of UEs, the UEs at row  $i$  are associated with the four ssRISs at rows  $i$  and  $i - 1$  to form coverage group  $i$ .

To determine the appropriate coverage group size, we consider an optimistic upper bound in which the self-sustainability constraints are relaxed and all ssRIS elements are used exclusively for reflection, with  $\alpha_{l,k} = 1$  for all  $l$  and  $k$ . Under this assumption,



**Figure 4.5.5:** Illustration of different coverage group designs. Coverage group consists of (a) 6 ssRISs, 6 UEs; (b) 4 ssRISs, 6 UEs; (c) 2 ssRISs, 6 UEs.



**Figure 4.5.6:** The upper bound of the achievable SNR under different scales of the coverage group.

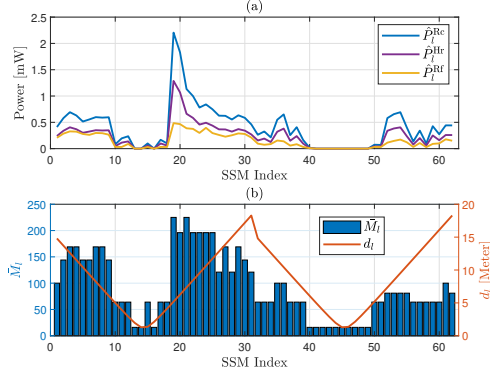
an upper bound on the achievable SNR is given by

$$\frac{\|\mathbf{h}_k + \mathbf{H}_k \phi_k\|^2 P}{BN_0} \leq \frac{(\|\mathbf{h}_k\| + \|\mathbf{H}_k\|_F)^2 P}{BN_0} = \hat{\Gamma}_k. \quad (4.5.41)$$

Figure 4.5.6 plots  $\hat{\Gamma}_k$  for coverage groups comprising 2, 4, and 6 nearest ssRISs per UE row. The results confirm that the dominant indirect signal contribution originates from ssRISs within the same row, and that the marginal benefit of incorporating additional ssRISs diminishes beyond four. Accordingly, each coverage group is designed to consist of 6 UEs paired with their 4 nearest ssRISs, corresponding to one row of UEs and two rows of ssRISs.

## 4.5.2 Preset design

Following the procedure described in Section 4.4.1, the preset for each ssRIS is selected from codebook  $\mathbf{C}$ , which comprises  $C = 10$  presets with 16, 64, 81, 100, 121, 144, 169, 196, 225, and 256 reflecting elements. The selection is performed over 10 Monte Carlo



**Figure 4.5.7:** (a) Averaged received power, harvested power, and consumed power under the searched-out reflecting area size with randomized phase shifts; (b) averaged searched-out reflecting area size and the distance between the ssRIS and the BS.

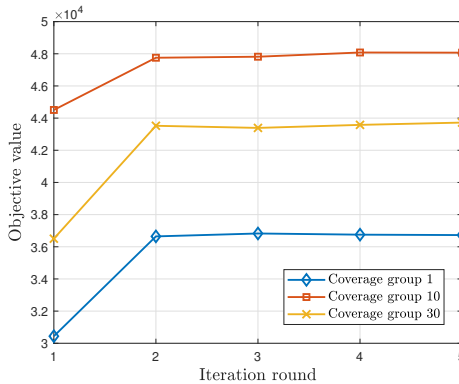
simulation rounds per ssRIS. For evaluation purposes,  $\hat{P}_l^{\text{Rc}}$ ,  $\hat{P}_l^{\text{Hr}}$ , and  $\hat{P}_l^{\text{Rf}}$  denote the average minimum received, harvested, and consumed power for ssRIS  $l$ , respectively.

Figure 4.5.7 shows the resulting reflecting area sizes and the corresponding power levels for all self-sustaining ssRISs. As shown in Figure 4.5.7(b), the selected reflecting area  $\bar{M}_l$  follows a non-monotonic pattern along the cabin. It is smallest near the BS, increases toward an intermediate peak, and then decreases again at the far edge. This behavior is explained by the interplay between MRT beamforming and path loss. Near the BS, MRT concentrates transmit power toward the stronger direct channel, leaving little power incident on the ssRIS and constraining  $\bar{M}_l$  to a small value. At the far edge of the cabin, the increased path loss reduces the total received power at the ssRIS, again limiting the supportable  $\bar{M}_l$ . At intermediate locations, the direct channel is weaker, causing MRT to distribute power more broadly, while path loss remains moderate, together enabling the largest feasible reflecting areas.

### 4.5.3 Convergence and complexity analysis

With the coverage groups and presets determined, the two-stage iterative algorithm is applied with a maximum of  $E = 5$  iteration rounds. The penalty coefficients are set to  $\omega = 100$ ,  $\omega' = 10^7$ , and  $\omega'' = 10^7$  based on empirical testing to ensure constraint satisfaction throughout the optimization.

Directly solving **P4.1** via branch-and-bound incurs a worst-case complexity of  $\mathcal{O}(2^{LC})$  [37], which is intractable at the considered simulation scale where  $C \times L = 620$  binary variables are involved. The proposed decomposition reduces this to  $\mathcal{O}(IE2^{L_i K_i})$  across all  $I$  coverage groups, with  $L_i \times K_i = 24$  binary variables per subproblem, rendering the optimization computationally feasible. Within each subproblem, FPP-SCA guarantees monotonic improvement of the continuous variables; however, due to the joint optimization of discrete associations and continuous phase shifts, theoretical convergence to a stationary point cannot be formally established. Empirically, as illustrated



**Figure 4.5.8:** Convergence condition of applying Algorithm 3 concerning the inputs from coverage groups 1, 10, and 30.

in Figure 4.5.8, the algorithm exhibits consistent objective improvement across coverage groups and converges within approximately two to three iterations, regardless of the specific coverage group considered.

#### 4.5.4 Feasibility and necessity of utilizing ssRIS

Due to propagation attenuation, the power harvested by an ssRIS may be insufficient to sustain its operation at certain locations. Furthermore, the contribution of an ssRIS to the received SNR varies depending on its deployment location relative to the BS and the served UEs. To quantify the latter, we define a contribution indicator  $\eta_k \in \mathbb{R}$  as the ratio of the direct-link SNR to the total received SNR, where  $\Gamma_k^{\text{BS}} = \|\mathbf{h}_k\|^2 P / (BN_0)$  denotes the SNR achievable from the BS alone under MRT beamforming

$$\eta_k = \frac{\Gamma_k^{\text{BS}}}{\Gamma_k}. \quad (4.5.42)$$

As the direct component strengthens,  $\eta_k$  approaches 1, indicating that the ssRIS contributes marginally to the received SNR. Conversely,  $\eta_k = 0$  indicates that all received signal power is delivered via the ssRIS, with no direct BS connectivity.

Figure 4.5.9(a) shows the spatial distribution of  $\eta_k$  across the cabin. UEs within the six rows closest to the BS almost universally exhibit  $\eta_k \geq 0.8$ , indicating that MRT beamforming in that region is dominated by the strong direct channel, leaving little room for ssRIS contribution. This is consistent with the observations reported in [39]. Figure 4.5.9(b) shows the number of UEs associated with each ssRIS. ssRISs near the cabin edge are associated with substantially more UEs, whereas those close to the BS serve comparatively few. Together, these results suggest that deploying ssRISs in the vicinity of the BS offers limited benefit, as the direct link already dominates and fewer UEs require indirect assistance in that region.

Regarding self-sustainability, all ssRISs in the system are verified to satisfy the harvesting constraint after optimization. Figure 4.5.10 illustrates a snapshot comparison

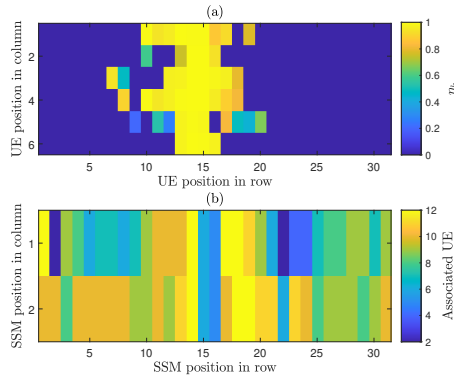


Figure 4.5.9: (a) BS contribution for each UE; (b) ssRIS-UE association condition.

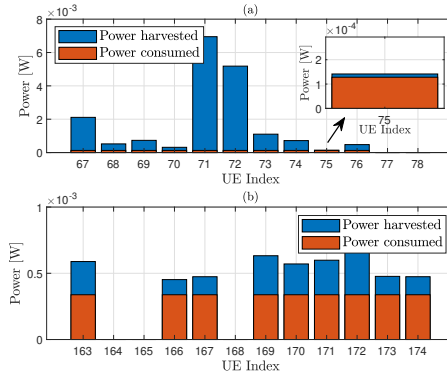
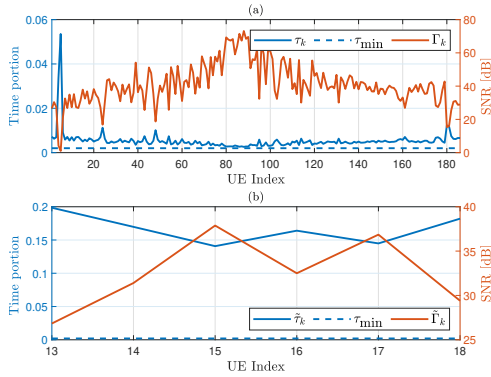


Figure 4.5.10: Self-sustainability condition of (a) ssRIS 12, close to the BS,  $\bar{M}_{12} = 64$ ; (b) ssRIS 29, far from the BS,  $\bar{M}_{29} = 169$ .

between ssRIS 12, located near the BS in the middle of the cabin, and ssRIS 29, positioned at the far edge. Both surfaces harvest more power than they consume across all assigned UEs in their coverage groups, confirming self-sustainability. However, the unassigned UEs within each coverage group reveal that the energy balance is contingent on the specific UE being served—an important practical consideration in low channel-rank regimes such as IDS. Notably, despite its shorter distance to the BS, ssRIS 12 harvests less power than ssRIS 29, as MRT beamforming concentrates transmit power toward the stronger direct channel in that region, reducing the power incident on the ssRIS.

The time resource allocation is illustrated in Figure 4.5.11. Figure 4.5.11(a) confirms that all UEs achieve an SNR exceeding 0 dB, guaranteeing network-wide connectivity. As expected, SNR decreases with UE-BS distance, and the time allocator compensates by assigning proportionally more resources to UEs with weaker channels. UEs near



**Figure 4.5.11:** (a) Time allocation and the resulting SNR of the whole cabin; (b) intra-group time allocation and the optimized SNR of the coverage group that consists of UEs 13-18.

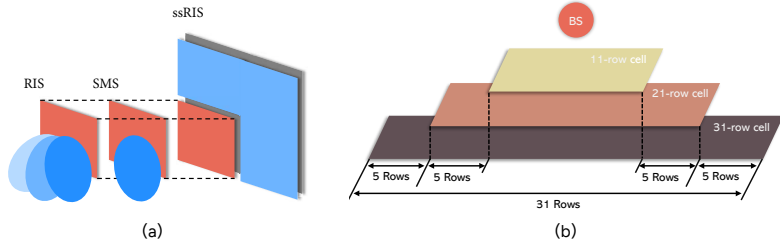
the BS receive the least time allocation, yet their portions remain above the minimum threshold  $\tau_{\min}$ . Figure 4.5.11(b) confirms that this inverse SNR-time relationship is maintained consistently within individual coverage groups through Algorithm 3.

#### 4.5.5 Achievable data rate

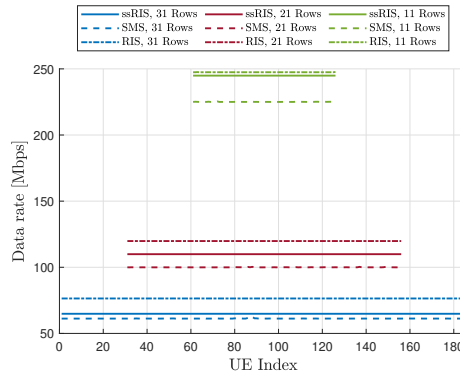
To position ssRIS, SMS, and RIS within the trade-off space among metasurface gain, system coverage, and operating cost, the minimum data rate achieved under each metasurface type is compared. To ensure a fair comparison, the RIS and SMS are configured with the same reflecting area as the reflective part of the ssRIS and placed at identical locations, as illustrated in Figure 4.5.12a. Under this setup, the channel coefficients seen by all three metasurface types are identical. Since RIS and SMS are not subject to self-sustainability constraints, they are operated under the system model of Figure 4.3.1, where all surfaces jointly serve every UE. The performance of the RIS- and SMS-assisted systems is evaluated using the algorithm from Paper A [1], which jointly optimizes phase shifts and time resource allocation to maximize the minimum data rate.

To examine how the trade-offs evolve with deployment scale, the comparison is conducted across three cabin sizes consisting of 11, 21, and 31 rows of UEs and ssRISs, respectively, as illustrated in Figure 4.5.12b. In all cases, the BS is placed at the center of the cabin.

The results are presented in Figure 4.5.13. Under identical channel conditions, RIS consistently outperforms SMS owing to its reconfigurability, with an average per-UE performance gap of 22.4, 19.8, and 15.2 Mbps for the 11-, 21-, and 31-row cabins, respectively. As a self-sustainable alternative, ssRIS achieves performance intermediate between SMS and RIS: its reconfigurability yields gains over SMS, while the self-sustainability constraint introduces a performance penalty relative to RIS. The performance advantage of ssRIS over SMS amounts to 19.8, 9.9, and 3.6 Mbps for the 11-, 21-, and 31-row cabins, respectively. Notably, in the 11-row cabin, the short BS-to-ssRIS



**Figure 4.5.12:** Illustrations of metasurface performance test. (a) Setups for different types of metasurface; (b) considered cabins of different sizes and their relative position.



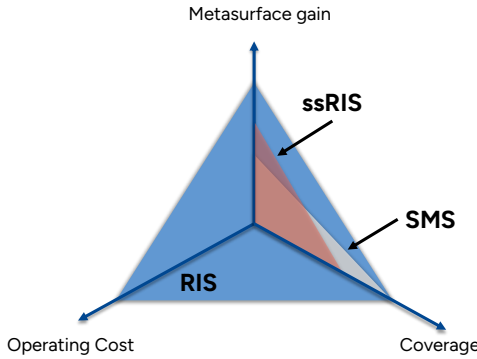
**Figure 4.5.13:** Optimized data rate under the assistance of different types of metasurface and under different cell sizes.

distance makes self-sustainability considerably easier to satisfy, and the resulting performance closely approaches the RIS upper bound, confirming that the proposed simplifications introduce negligible degradation when the self-sustainability constraint is not tight.

As the cabin size increases, the advantage of ssRIS over SMS diminishes. The growing BS-to-ssRIS distance makes self-sustainability progressively harder to maintain, effectively limiting the coverage region within which ssRIS can operate. Beyond this region, ssRIS can no longer sustain its operation, whereas SMS, being entirely passive, remains deployable without restriction. Consequently, when system coverage is the primary requirement, SMS remains a viable and low-cost solution, particularly at the far end of the cabin where ssRIS feasibility degrades.

## 4.6 Conclusion

This chapter addressed the research question **RQ2**. To answer this, the design of an ssRIS-assisted indoor mmWave communication system is investigated, aiming to



**Figure 4.6.14:** General comparison in terms of operating cost, metasurface gain, and system coverage among ssRIS, SMS, and RIS.

achieve a favorable balance among metasurface gain, system coverage, and operating cost. To manage the computational complexity inherent in the joint optimization of phase shifts, ssRIS-UE association, and time resource allocation, a preset-based element splitting scheme and coverage group formation are introduced, and a two-stage iterative algorithm based on FPP-SCA is proposed to maximize the minimum achievable data rate across all UEs.

The results reveal that ssRIS is most effective when deployed in the LoS region of the BS, where sufficient energy can be harvested to maintain self-sustainability, and is most beneficial for supporting UEs with weak direct BS connections in spatially compact environments. Compared to SMS, ssRIS delivers higher data rates owing to its reconfigurability. Compared to RIS, it incurs a performance penalty due to the self-sustainability constraint. The key limitation of ssRIS is its restricted coverage. As the deployment environment grows larger and BS-to-ssRIS distances increase, maintaining self-sustainability becomes progressively more challenging, and SMS remains a practical low-operating-cost alternative for coverage at the far end of the cabin. Overall, small IDSs such as aircraft cabins represent the most favorable deployment scenario for ssRIS, where the self-sustainability constraint is less restrictive, and the performance-cost trade-off is most advantageous. Figure 4.6.14 summarizes the qualitative performance comparisons.

While this chapter demonstrates that ssRIS offers a favorable trade-off in small, space-constrained IDSs, the observed coverage limitation points toward a broader question. ssRIS extracts power from the same wireless signals it received. Any reduction in received power therefore directly threatens its self-sustainability. Such reductions can arise from increased propagation distance, limited system bandwidth, or higher path loss in other deployment environments. In each case, the harvested power may prove insufficient to sustain reconfigurability while simultaneously delivering acceptable data rates. Having verified that ssRIS can potentially be a better compromise in the metasurface trade-off space, the natural next step is to examine how broadly this holds across different communication scenarios and system parameters, which is the subject of the following chapter.

## 5 Self-sustainable RIS feasibility analysis

The preceding chapter established that ssRIS can offer a favorable position in the meta-surface trade-off space, particularly in small, space-constrained IDS where the compactness of the deployment area supports self-sustainability. However, coverage limitations observed in that setting reveal a more fundamental concern about the feasibility of maintaining self-sustainability. Any factor that reduces the power incident on the surface directly threatens its ability to sustain reconfigurability while simultaneously meeting data rate requirements. This tension between harvesting sufficiency and communication performance raises a natural and practically important question: how feasible is it to utilize ssRIS?

The number of elements is the critical design parameter through which this question must be examined. As beamforming gain scales quadratically with element count, a larger surface enables stronger signal amplification and thus improved Quality of Service (QoS) satisfaction. At the same time, each additional element increases power consumption, placing greater demand on the harvesting mechanism and challenging the self-sustainability requirement. Beyond the operational dimension, larger element counts also escalate manufacturing costs and control complexity, imposing practical limits on implementation. The minimum number of elements required to satisfy a given QoS target under a self-sustainability constraint therefore serves as a direct indicator of feasibility.

This chapter presents the work in **Paper C**, which directly addresses research question **RQ3**:

How feasible is it to utilize ssRIS in various communication scenarios? Specifically, how does the element count required to achieve self-sustainability scale with varying transmit power, data rate requirements, and outage requirements for ssRIS implementing different HaR schemes?

This chapter assesses ssRIS feasibility under the ES and TS HaR schemes by analyzing

how element requirements scale with transmit power, data rate demands, and outage constraints under both LoS and NLoS ssRIS-to-UE channels. The results reveal that TS demonstrates better channel hardening, maintaining more stable element requirements across varying outage margins, yet its element count scales exponentially with harvesting difficulty and data rate. ES, by contrast, exhibits only linear growth under worsening harvesting conditions, offering greater robustness in demanding scenarios. These findings establish TS as the preferable scheme in indoor environments where reliable connectivity is prioritized, while ES is better suited for challenging outdoor conditions where operational robustness is paramount.

## 5.1 Related Work

This work is mainly referred to [13]. As introduced in [13], ssRISs offer reconfigurability like RIS while harvesting energy from incoming electromagnetic waves to power their operation. This capability makes ssRIS a promising solution for cost-effective, high-throughput metasurface-assisted systems. Also investigated in the same paper, different HaR schemes show different performance across various communication scenarios. However, the element count requirements for achieving specific QoS targets remain underexplored.

## 5.2 Contributions

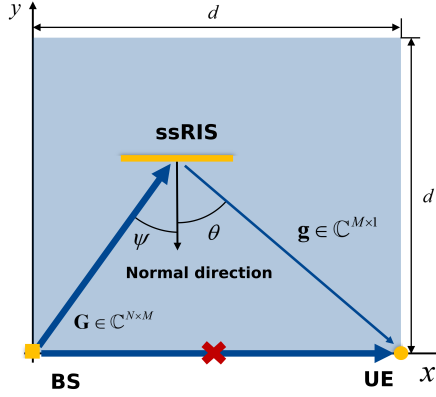
The main contributions of this work are as follows:

- For a single-UE MISO system, we prove that when the ssRIS-to-BS channel is LoS-dominated, MRT precoding guarantees both optimal energy harvesting and optimal SNR at the UE when there is no direct link between the BS and UE.
- Our analytical derivations together with numerical results reveal that different HaR schemes exhibit varying feasibility under different harvesting conditions and data rate requirements. The TS scheme enhances ssRIS feasibility when harvesting conditions are favorable and data rate requirements are low. Additionally, the TS scheme shows better feasibility in terms of maintaining a low outage. Conversely, in more challenging scenarios with poor harvesting conditions or high data rate demands, the ES scheme better maximizes ssRIS feasibility.

## 5.3 System Model

As illustrated in Fig. 5.3.1, the considered system comprises a BS, a UE, and an ssRIS positioned within a square region of side length  $d$ , with the BS and UE located at opposite corners. The BS is equipped with  $N$  antennas arranged in a UPA, the ssRIS consists of  $M$  elements also in a UPA configuration, and the UE is equipped with a single isotropic antenna. All devices are coplanar, with the ssRIS surface oriented parallel to the BS plane.

Since metasurface gains diminish significantly in the presence of a direct BS-UE LoS link [1, 39], the UE is assumed to have no direct link to the BS due to blockage. Furthermore, since the ssRIS relies on harvesting energy from incident electromagnetic waves,



**Figure 5.3.1:** System model. The square and circle represent the position of the BS and the UE, respectively.

it must be deployed within the LoS region of the BS to ensure sufficient power collection for self-sustainability. Accordingly, the BS-ssRIS channel  $\mathbf{G} \in \mathbb{C}^{N \times M}$  is modeled as a pure LoS channel

$$\mathbf{G} = \sqrt{\rho(\psi, d_{\text{SR}})} e^{-j\varphi_{\text{SR}}} \cdot \mathbf{a}_N(\psi) \mathbf{a}_M^T(\psi) \quad (5.3.1)$$

where  $\varphi_{\text{SR}}$  is the reference phase shift,  $d_{\text{SR}}$  is the propagation distance, and  $\psi$  is the azimuth angle of the ssRIS with respect to the BS. The free-space path loss (FSPL) is given by  $\rho(\psi, d) = \frac{\lambda^2}{(4\pi d)^2} G(\psi)$ , where  $\lambda$  is the carrier wavelength and  $G(\psi) = \pi \cos(\psi)$  for  $\psi \in [-\pi/2, \pi/2]$  denotes the ssRIS antenna gain under the cosine model [7, 40]. The array response vectors  $\mathbf{a}_M(\psi) \in \mathbb{C}^{M \times 1}$  and  $\mathbf{a}_N(\psi) \in \mathbb{C}^{N \times 1}$  have unit-modulus entries.

The channel from the ssRIS to the UE is denoted by  $\mathbf{g} \in \mathbb{C}^{M \times 1}$ . The received signal at the UE is expressed as

$$y = \phi^H \mathbf{H} \mathbf{w} s + n \quad (5.3.2)$$

where  $n$  denotes the additive white Gaussian noise (AWGN) with spectral density  $N_0$ , and the cascaded BS-ssRIS-UE channel is  $\mathbf{H} \in \mathbb{C}^{M \times N} = \text{diag}(\mathbf{g}) \mathbf{G}^H$ . The vector  $\phi \in \mathbb{C}^{M \times 1}$  represents the phase shifts applied by the ssRIS,  $\mathbf{w} \in \mathbb{C}^{N \times 1}$  is the unit-norm digital precoding vector at the BS, and  $s$  is the transmitted data stream satisfying  $\mathbb{E}\{|s|^2\} = P$ .

The power received by the ssRIS, which serves as the source for energy harvesting, is given by

$$\begin{aligned} P_{\text{Rc}} &= P \|\mathbf{G}^H \mathbf{w}\|^2 \\ &= P \rho(\psi, d_{\text{SR}}) (\mathbf{w}^H \mathbf{a}_N(\psi) \mathbf{a}_M^T(\psi) \mathbf{a}_M^*(\psi) \mathbf{a}_N^H(\psi) \mathbf{w}) \\ &\leq P N M \rho(\psi, d_{\text{SR}}) \end{aligned} \quad (5.3.3)$$

where the upper bound follows from the Cauchy-Schwarz inequality, and the contribution of AWGN to harvested power is neglected. It is worth noting that the precoding

vector  $\mathbf{w}$  influences not only the signal quality at the UE but also the self-sustainability of the ssRIS, as the inner product between  $\mathbf{a}_N(\psi)$  and  $\mathbf{w}$  in (5.3.3) directly determines the power available for harvesting.

Lossless reflection is assumed without loss of generality, such that the magnitude of each phase-shift entry satisfies  $|\phi_m| = 1$  for all  $m$ . MRT beamforming is employed at the BS, given by

$$\mathbf{w} = \frac{(\phi^H \mathbf{H})^H}{\|\phi^H \mathbf{H}\|}, \quad (5.3.4)$$

which maximizes the received SNR at the UE. Under the pure LoS assumption for the BS–ssRIS channel, the Gram matrix  $\mathbf{V} \in \mathbb{C}^{M \times M}$ , defined as  $\mathbf{V} = \mathbf{H}\mathbf{H}^H$ , has rank one. Consequently, the optimal phase-shift vector is given by  $\phi = \exp(j\angle \mathbf{u}_1)$ , where  $\mathbf{u}_1$  is the eigenvector of  $\mathbf{V}$  corresponding to its only non-zero eigenvalue  $\lambda_1$ . The resulting optimal SNR at the UE is

$$\Gamma^* = \frac{P\lambda_1 \|\phi^H \mathbf{u}_1\|^2}{BN_0} \quad (5.3.5)$$

where  $B$  is the system bandwidth. The  $\mathbf{u}_1$  and  $\lambda_1$  fully determine the system performance, and its specific form depends on the propagation conditions of the ssRIS-UE channel, as derived in the following subsections.

### 5.3.1 Data rate under different channel conditions

While the ssRIS must be positioned within the BS's LoS region to meet self-sustainability requirements, the ssRIS-UE channel need not be LoS. The data rate analysis is therefore conducted separately for each propagation condition.

For the LoS ssRIS-UE channel, under the assumption that the LoS component dominates the ssRIS-UE channel, the channel response is given as

$$\mathbf{g} = \sqrt{\rho(\theta, d_{\text{RD}})} e^{-j\varphi_{\text{RD}}} \mathbf{a}_M(\theta) \quad (5.3.6)$$

where  $\varphi_{\text{RD}}$  is the reference phase shift and  $\theta$  is the azimuth angle of the UE with respect to the ssRIS. Following a similar analysis as for  $\mathbf{G}$ , the Gram matrix  $\mathbf{V} = \text{diag}(\mathbf{g})\mathbf{G}^H\mathbf{G}\text{diag}(\mathbf{g})^H$  yields the dominant eigenvalue and eigenvector as

$$\mathbf{u}_1 = \frac{\bar{\mathbf{a}}_M}{\sqrt{\rho(\theta, d_{\text{RD}})M}}, \quad \lambda_1 = \rho_0 NM \quad (5.3.7)$$

where  $\bar{\mathbf{a}}_M = \text{diag}(\mathbf{g})\mathbf{a}_M^*(\psi)$  and  $\rho_0 \triangleq \rho(\psi, d_{\text{SR}})\rho(\theta, d_{\text{RD}})$ . This leads to the optimal beamforming vectors  $\phi = \sqrt{M}\mathbf{u}_1$  and  $\mathbf{w} = \frac{\mathbf{a}_N(\psi)}{\sqrt{N}}$ , yielding the achievable data rate

$$R = B \log_2 \left( 1 + \underbrace{P\rho_0/(BN_0)}_{\triangleq \Gamma_0} \cdot NM^2 \right) \quad (5.3.8)$$

and the optimal received power available for energy harvesting

$$P_{\text{Rc}} = PNM\rho(\psi, d_{\text{SR}}). \quad (5.3.9)$$

For the case where the ssRIS-UE channel is NLoS, the corresponding channel response is modeled as

$$\mathbf{g} = \sqrt{\rho(\theta, d_{\text{RD}})} \tilde{\mathbf{g}} \quad (5.3.10)$$

where  $\tilde{\mathbf{g}} \in \mathbb{C}^M$  is the fading vector. The Gram matrix  $\mathbf{V}$  then becomes

$$\begin{aligned} \mathbf{V} &= \rho(\theta, d_{\text{RD}}) \text{diag}(\tilde{\mathbf{g}}) \mathbf{G}^H \mathbf{G} \text{diag}(\tilde{\mathbf{g}})^H \\ &= \rho_0 N \|\tilde{\mathbf{a}}_M\|^2 \frac{\tilde{\mathbf{a}}_M \tilde{\mathbf{a}}_M^H}{\|\tilde{\mathbf{a}}_M\| \|\tilde{\mathbf{a}}_M\|} \end{aligned} \quad (5.3.11)$$

where  $\tilde{\mathbf{a}}_M = \text{diag}(\tilde{\mathbf{g}}) \mathbf{a}_M^*(\psi) \in \mathbb{C}^M$  is an auxiliary vector. The corresponding dominant eigenpair is

$$\mathbf{u}_1 = \frac{\tilde{\mathbf{a}}_M}{\|\tilde{\mathbf{a}}_M\|}, \quad \lambda_1 = \rho_0 N \sum_{m=1}^M |\tilde{g}_m|^2 \quad (5.3.12)$$

where  $\tilde{g}_m$  is the  $m$ -th entry of  $\tilde{\mathbf{g}}$ . Substituting into (5.3.5), the maximum SNR is

$$\Gamma^* = \Gamma_0 N \left( \sum_{m=1}^M |\tilde{a}_{M,m}| \right)^2 \quad (5.3.13)$$

where  $\tilde{a}_{M,m}$  denotes the  $m$ -th element of  $\tilde{\mathbf{a}}_M$ . From (5.3.12), the resulting optimal precoder is  $\mathbf{w} = \frac{\mathbf{a}_N(\psi)}{\sqrt{N}}$ , which simultaneously achieves maximum harvested power at the ssRIS.

Since practical hardware limitations prevent the ssRIS from reconfiguring at the rate of channel fluctuations, the outage probability is adopted as the performance metric for the NLoS scenario, defined as the probability that the instantaneous SNR falls below a threshold  $\gamma$ . Assuming uncorrelated Rayleigh fading  $\tilde{\mathbf{g}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ , the sum  $Y = \sum_{m=1}^M |\tilde{a}_{M,m}|$  can be approximated via the central limit theorem (CLT) as a truncated Gaussian random variable. The truncated Gaussian approximation is adopted since  $Y$  is strictly nonnegative, and has been shown to be accurate for  $M \geq 10$  in [41]. The corresponding cumulative distribution function (CDF) is

$$F_Y(y) = \begin{cases} 1 - CQ((y - \mu_Y)/\sigma_Y), & y \geq 0, \\ 0, & y < 0, \end{cases} \quad (5.3.14)$$

where  $\mu_Y = M\sqrt{\pi}/2$ ,  $\sigma_Y^2 = M(4 - \pi)/4$ , and  $C = 1/Q(-\mu_Y/\sigma_Y)$ , with  $Q(\cdot)$  denoting the Gaussian Q-function. The SNR outage probability is accordingly given by

$$P_{\text{out}}(\gamma) = \mathbb{P}\{\Gamma < \gamma\} = F_Y\left(\sqrt{\frac{\gamma}{N\Gamma_0}}\right). \quad (5.3.15)$$

### 5.3.2 Self-sustainability constraints under various HaR schemes

As analyzed in [13], for the assumed LoS BS-ssRIS channel, ES and Power Splitting (PS) exhibit equivalent performance, narrowing the analysis to ES and TS. Self-sustainability

in this work refers to the condition where the total harvested energy meets or exceeds the energy consumed during the reflecting phase.

In the ES scheme, harvesting and reflecting are performed simultaneously by dedicating different subsets of elements to each function. Specifically,  $M_{\text{Hr}}$  elements are connected to the harvesting circuit and fully absorb the incident signals, while  $M_{\text{Rf}}$  elements are connected to the reflecting circuit and inject phase shifts into the reflected signals. Correspondingly, the BS-ssRIS channel is partitioned into a harvesting-related sub-channel  $\mathbf{G}_{\text{Hr}} \in \mathbb{C}^{N \times M_{\text{Hr}}}$  and a reflecting-related sub-channel  $\mathbf{G}_{\text{Rf}} \in \mathbb{C}^{N \times M_{\text{Rf}}}$ . The self-sustainability constraint for the ES scheme is expressed as

$$\eta P \rho(\psi, d_{\text{SR}}) N M_{\text{Hr}} \geq M_{\text{Rf}} P_0 \quad (5.3.16)$$

where  $P_0$  denotes the power consumption per reflecting element and  $\eta$  is the energy harvesting efficiency.

In the TS scheme, the harvesting and reflecting phases occur orthogonally over time. During the harvesting phase, all elements are connected to the harvesting circuit and fully absorb the incident signals, while during the reflecting phase, all elements are switched to the reflecting circuit and apply phase shifts to the impinging signals. The fraction of time allocated to harvesting is denoted by  $\tau$ , with the remaining fraction  $1 - \tau$  dedicated to reflecting. The self-sustainability constraint for the TS scheme is expressed as

$$\tau \eta P \rho(\psi, d_{\text{SR}}) N M \geq (1 - \tau) M P_0 \quad (5.3.17)$$

where it can be observed that the element count  $M$  appears on both sides of the inequality and cancels out, indicating that the self-sustainability condition of the TS scheme is independent of the number of elements.

## 5.4 Problem Formulation

The feasibility of ssRIS deployment is quantified by formulating optimization problems that minimize the total number of elements required to satisfy a target data rate  $R_0$  while maintaining self-sustainability. A lower minimum element count under given operating conditions directly indicates higher feasibility, as the ssRIS can meet its performance targets with fewer resources.

For the ES scheme under a LoS ssRIS-UE channel, the element number optimization problem is formulated as

$$\mathbf{P5.1:} \underset{M_{\text{Hr}}, M_{\text{Rf}}}{\text{minimize}} \quad M_{\text{Hr}} + M_{\text{Rf}} \quad (5.4.18a)$$

$$\text{s.t.} \quad (5.3.16),$$

$$B \log_2 (1 + \Gamma_0 N M_{\text{Rf}}^2) \geq R_0 \quad (5.4.18b)$$

where constraint (5.3.16) ensures self-sustainability and constraint (5.4.18b) guarantees that the achievable data rate meets the target threshold. Since the total element count is unbounded from above, the optimum is always attainable, and both constraints

are satisfied with equality at the optimum. The optimal element counts are obtained analytically as

$$M_{\text{Rf}}^* = \sqrt{\left(2^{\frac{R_0}{B}} - 1\right) \frac{1}{N\Gamma_0}} \quad (5.4.19)$$

$$M_{\text{Hr}}^* = \frac{P_0}{\underbrace{\eta P N \rho(\psi, d_{\text{SR}})}_{\triangleq \alpha}} \sqrt{\left(2^{\frac{R_0}{B}} - 1\right) \frac{1}{N\Gamma_0}} \quad (5.4.20)$$

where the constant  $\alpha$  characterizes the harvesting condition. A smaller  $\alpha$  corresponds to more favorable harvesting, arising from higher transmit power, reduced propagation attenuation, or a larger antenna array at the BS.

For the TS scheme under a LoS ssRIS-UE channel, the optimization problem is formulated as

$$\mathbf{P5.2:} \underset{M, \tau}{\text{minimize}} \quad M \quad (5.4.21a)$$

$$\text{s.t.} \quad (5.3.17),$$

$$(1 - \tau)B \log_2(1 + \Gamma_0 N M^2) \geq R_0. \quad (5.4.21b)$$

Applying a similar analytical approach as for **P5.1**, the optimal time splitting ratio and element count are obtained as

$$\tau^* = \frac{P_0}{\eta P N \rho(\psi, d_{\text{SR}}) + P_0} = \frac{\alpha}{1 + \alpha} \quad (5.4.22)$$

$$M^* = \sqrt{\left(2^{\frac{R_0}{(1-\tau^*)B}} - 1\right) \frac{1}{N\Gamma_0}}. \quad (5.4.23)$$

When the ssRIS-UE channel experiences NLoS conditions, the deterministic data rate constraint is replaced by an outage probability constraint requiring  $P_{\text{out}}(\gamma) \leq \varepsilon$ , where  $\varepsilon$  is the tolerable outage margin. For the ES scheme, the resulting problem is

$$\mathbf{P5.3:} \underset{M_{\text{Hr}}, M_{\text{Rf}}}{\text{minimize}} \quad M_{\text{Rf}} + M_{\text{Hr}} \quad (5.4.24a)$$

$$\text{s.t.} \quad (5.3.16),$$

$$P_{\text{out}}(\gamma_0^{\text{ES}}) \leq \varepsilon \quad (5.4.24b)$$

where the SNR threshold is determined by the data rate requirement as  $\gamma_0^{\text{ES}} = 2^{R_0/B} - 1$ . Substituting  $\gamma_0^{\text{ES}}$  into (5.3.15), constraint (5.4.24b) expands to

$$\underbrace{(1 - \varepsilon)Q\left(-\sqrt{\frac{M_{\text{Rf}}\pi}{4 - \pi}}\right)}_{\triangleq f_1(M_{\text{Rf}})} \leq Q\left(2\sqrt{\frac{\gamma_0^{\text{ES}}}{N\Gamma_0 M_{\text{Rf}}(4 - \pi)}} - \sqrt{\frac{M_{\text{Rf}}\pi}{4 - \pi}}\right). \quad (5.4.25)$$

$$\underbrace{\hspace{15em}}_{\triangleq f_2(M_{\text{Rf}})}$$

Since  $f_2(0) - f_1(0) < 0$  and  $f_2(\infty) - f_1(\infty) = \varepsilon > 0$ , there exists at least one value of  $M_{\text{Rf}}$  satisfying the above inequality with equality. Denoting the minimum such root as  $\tilde{M}_{\text{Rf}}$ , which is computed numerically via the bisection method, the optimal solution is

$$M_{\text{Rf}}^* = \tilde{M}_{\text{Rf}}, \quad M_{\text{Hr}}^* = \alpha M_{\text{Rf}}^*. \quad (5.4.26)$$

For the TS scheme under NLoS conditions, the optimization problem is formulated as

$$\mathbf{P5.4:} \quad \underset{M, \tau}{\text{minimize}} \quad M \quad (5.4.27a)$$

$$\text{s.t.} \quad (5.3.17),$$

$$P_{\text{out}}(\gamma_0^{\text{TS}}) \leq \varepsilon \quad (5.4.27b)$$

where  $\gamma_0^{\text{TS}} = 2^{R_0/((1-\tau)B)} - 1$  is monotonically increasing with  $\tau$ , and thus the outage probability is also monotonically increasing with  $\tau$ . Consequently, the optimal time splitting ratio is given by (5.4.22), yielding the minimum feasible SNR threshold  $\tilde{\gamma}_0^{\text{TS}}$ . The corresponding minimum element count  $\tilde{M}$  is then obtained by substituting  $\gamma_0^{\text{ES}}$  and  $M_{\text{Rf}}$  with  $\tilde{\gamma}_0^{\text{TS}}$  and  $\tilde{M}$  in (5.4.25) and solving via the bisection method.

## 5.5 Results

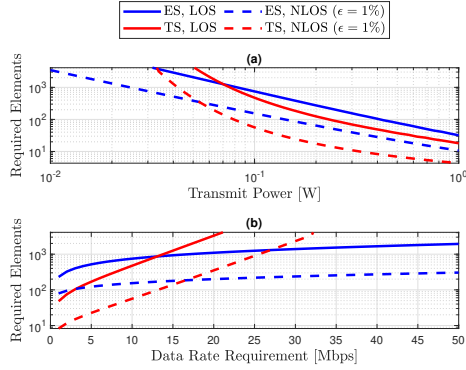
From the preceding analysis, the element requirement depends on three main factors: the data rate requirement  $R_0$ , the harvesting condition  $\alpha$ , and the outage margin  $\varepsilon$  for NLoS scenarios. This section investigates ssRIS feasibility by examining how the element requirement scales with each factor under the ES and TS schemes. Although  $\alpha$  can be improved by increasing transmit power  $P$ , expanding the antenna count  $N$ , or reducing the per-element power consumption  $P_0$  through hardware improvements, the analysis focuses on varying  $P$  as the most readily adjustable parameter in practical deployments. The impacts of  $R_0$  and  $\varepsilon$  are examined directly.

The simulation parameters are set as follows. The carrier frequency is  $f = 15$  GHz with a bandwidth of  $B = 50$  MHz. The number of BS antennas is  $N = 128$ , the energy harvesting efficiency is  $\eta = 0.65$ , and the per-element reflecting power consumption is  $P_0 = 2 \mu\text{W}$  [13]. For the NLoS ssRIS–UE channel, the outage margin is set to  $\varepsilon = 1\%$  as a conservative choice targeting high reliability. The side length of the deployment area is  $d = 50$  m.

As illustrated in Fig. 5.5.2, under the LoS ssRIS–UE channel, TS achieves better element feasibility with a lower element requirement at high  $P$  or low  $R_0$ . This advantage diminishes as  $P$  decreases or  $R_0$  increases, eventually reaching an intersection point beyond which TS becomes less feasible than ES. To understand this behavior, the total minimum element counts for ES and TS in the LoS case are expressed in closed form as

$$M_{\text{ES}} = (1 + \alpha) \sqrt{\left(2^{\frac{R_0}{B}} - 1\right) \frac{1}{N\Gamma_0}}, \quad (5.5.28)$$

$$M_{\text{TS}} = \sqrt{\left(2^{\frac{R_0(1+\alpha)}{B}} - 1\right) \frac{1}{N\Gamma_0}}. \quad (5.5.29)$$

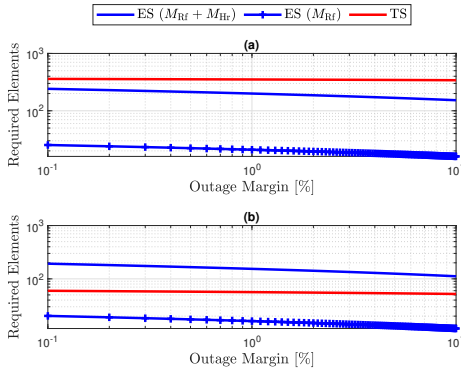


**Figure 5.5.2:** Element requirement across different system parameters. LoS and NLoS refer to the propagation conditions of the ssRIS-UE channel. (a) Impact of harvesting condition (via changing transmit power  $P$ ) with  $R_0 = 10$  Mbps; (b) Impact of data rate  $R_0$  with  $P = 0.1$  W.

Under favorable harvesting conditions where  $\alpha \rightarrow 0$ , both schemes converge to similar element requirements. However, as the harvesting condition deteriorates and  $\alpha \rightarrow \infty$ , the TS element requirement grows exponentially, whereas that of ES scales only linearly with  $\alpha$ . Consequently, TS requires fewer elements when harvesting conditions are favorable, but demands significantly more elements as conditions worsen. A similar divergence is observed with increasing  $R_0$ , since the data rate requirement enters the TS expression through the exponent scaled by  $\alpha$ , amplifying its effect relative to the ES case.

The feasibility characteristics in the NLoS case follow a similar pattern to those observed under LoS conditions, though with a generally more relaxed element requirement. This behavior can be understood through the structure of the outage probability constraint (5.4.25). Under favorable operating conditions, the root achieving equality in (5.4.25) lies in the central region of the Q-function, where both ES and TS require comparably small element counts to satisfy the communication constraint. In this regime, ES incurs additional harvesting elements to support its reflecting operations, resulting in a higher total element count than TS. Conversely, as operating conditions deteriorate and the root is pushed into the tail region of the Q-function, the outage constraint becomes exponentially more difficult to satisfy. Since  $\gamma_0^{\text{TS}}$  grows exponentially larger than  $\gamma_0^{\text{ES}}$  for the same  $R_0$ , the TS scheme faces a significantly more stringent outage constraint, and ES demonstrates higher feasibility in this regime.

To explicitly examine the impact of the outage margin, the element requirement is evaluated under system conditions near the intersection point identified in Fig. 5.5.2(a), with the corresponding results presented in Fig. 5.5.3. As the outage constraint tightens with decreasing  $\varepsilon$ , the root achieving equality in (5.4.25) is pushed further into the tail region of the Q-function, reducing the gap between the arguments on both sides and driving up the element requirement. The results in Fig. 5.5.3 confirm this trend, yet reveal that TS exhibits considerably less sensitivity to changes in the outage margin than ES. This can be attributed to the fact that, under the TS scheme, all elements



**Figure 5.3:** Element requirement across different outage margins under the NLoS ssRIS-UE channel. (a)  $P = 0.1$  W and  $R_0 = 20$  Mbps; (b)  $P = 0.1$  W and  $R_0 = 15$  Mbps.

participate in signal reflection during the reflecting phase, allowing a larger effective aperture to contribute to beamforming for the same data rate requirement. This yields an additional channel hardening gain that stabilizes the element requirement against tightening outage constraints.

## 5.6 Conclusion

This chapter examines the feasibility of ssRIS deployment under the ES and TS HaR schemes by analyzing how element requirements scale with harvesting conditions, data rate demands, and outage constraints across both LoS and NLoS ssRIS-UE channel scenarios.

The results reveal complementary feasibility characteristics between the two schemes. TS benefits from better channel hardening, maintaining more stable element requirements under varying outage margins. However, its element count scales exponentially with harvesting difficulty and data rate demands, limiting its practicality in challenging operating conditions. This makes TS well-suited for indoor deployments such as Internet of Things (IoT) and WiFi enhancement, where harvesting conditions are favorable and reliable connectivity is the primary objective. ES, by contrast, exhibits only linear sensitivity to worsening harvesting conditions and data rate requirements, offering greater robustness when self-sustainability is difficult to maintain. This positions ES as the more feasible choice for outdoor or resource-constrained deployments where operational robustness takes precedence.

## 6 Conclusions and future work

### 6.1 Concluding remarks

This thesis investigates the design of sustainable metasurface-assisted indoor wireless communication systems, with operating cost placed alongside performance as a primary design criterion. The central argument is that the trade-off between metasurface gain and operating cost is not a fixed constraint to be accepted, but a design space to be navigated deliberately. Three research questions are posed in Chapter ??, and the findings of this thesis provide concrete answers to each.

**RQ1** asked how much reconfigurability is needed for a metasurface-assisted system to satisfy user requirements, and how a mixed deployment of SMS and RIS can be optimally configured to maximize the minimum data rate. The results presented in Chapter 3 demonstrate a clear diminishing-returns relationship between reconfigurability and performance. In a 22-metasurface aircraft cabin deployment at 28 GHz, replacing just two SMSs with RISs already yields a 13 Mbps gain over the all-SMS baseline of 152 Mbps, confirming that even limited reconfigurability is immediately beneficial. However, increasing the RIS count from 16 to 22 produces less than 1 Mbps of additional gain, indicating that full reconfigurability is unnecessary. Notably, an all-SMS deployment of 22 surfaces achieves the same minimum data rate as a deployment of only 18 RISs, suggesting that densely deployed passive surfaces constitute a cost-effective alternative when the operating budget is constrained. This finding motivates the search for a middle-ground solution that retains meaningful reconfigurability without incurring the full operating overhead of RIS.

**RQ2** asked how an ssRIS-assisted mmWave system in IDS can be designed to maximize the minimum data rate while satisfying the self-sustainability constraint, and how ssRIS can be positioned within the gain-cost trade-off together with SMS and RIS. Chapter 4 addressed this by proposing a two-stage iterative algorithm based on FPP-SCA that jointly optimizes phase shifts, ssRIS-UE associations, and time allocation. The results confirm that ssRIS achieves a favorable position in the trade-off: it outper-

forms SMS by 3.6, 9.9, and 19.8 Mbps in cabins of 31, 21, and 11 rows, respectively, while eliminating the need for cabling and an external power supply. The performance advantage grows as the environment becomes more compact, where self-sustainability is easier to maintain due to shorter BS-to-ssRIS distances. The key limitation is coverage, as self-sustainability becomes increasingly difficult to maintain as BS-to-ssRIS distances grow, making SMS the more reliable choice in large-scale deployments where system coverage requirements are high.

**RQ3** asked how feasible it is to utilize ssRIS in diverse communication scenarios, and how element count requirements scale with varying system conditions under different HaR schemes. Chapter 5 provided an analytical and numerical feasibility characterization for the ES and TS schemes. Under favorable harvesting conditions, TS requires fewer elements and benefits from stronger channel hardening, exhibiting considerably less sensitivity to tightening outage constraints under NLOS ssRIS-UE channels. However, as harvesting conditions worsen or data rate demands increase, the TS element count grows exponentially, whereas ES scales only linearly with harvesting difficulty, offering greater robustness in demanding or resource-constrained deployments. These complementary profiles provide concrete guidance. TS is better suited for compact indoor environments where reliable connectivity is prioritized, while ES is the more robust choice where operational conditions are challenging.

Taken together, the three contributions form a coherent progression: from quantifying the boundary of the problem space, through designing a practical system-level solution, to establishing its fundamental feasibility limits. The thesis demonstrates that ssRIS represents a technically sound and practically deployable middle-ground between the extremes of SMS and RIS, and provides actionable guidance for sustainable metasurface selection and deployment in next-generation wireless systems.

## 6.2 Future Research Directions

The system model adopted in this thesis focuses on a single-cell indoor environment under idealized hardware assumptions. Extending the analysis to multi-cell or heterogeneous network scenarios, where interference between metasurface deployments across different cells must be managed, represents a natural and practically relevant next step. In parallel, incorporating hardware imperfections such as phase noise, and non-ideal energy harvesting circuits into the system model is particularly important for ssRIS, whose self-sustainability constraint is sensitive to deviations from ideal hardware behavior. These two directions together would substantially strengthen the practical applicability of the framework developed in this thesis.

Beyond data rate enhancement, ssRIS holds promise as a resilient backup solution for network infrastructure failures. Unlike conventional active relays or base stations, ssRIS operates without an external power supply or backhaul connectivity, making it inherently self-contained. In scenarios where power outages or infrastructure disruptions compromise the primary network, a pre-deployed ssRIS layer could maintain a baseline level of wireless coverage autonomously. Investigating the design and performance of ssRIS under such emergency or resilience-oriented use cases represents a compelling and underexplored research direction.

Finally, the utility of ssRIS need not be confined to communication throughput maximization. Integrated sensing and communication (ISAC) is an emerging paradigm for 6G, in which the same infrastructure simultaneously supports data transmission and environmental sensing. The passive and reconfigurable nature of ssRIS makes it a potentially attractive component in ISAC systems, yet its performance in joint sensing and communication tasks remains largely unexplored. Understanding the trade-offs introduced by the self-sustainability constraint in this broader context constitutes a promising avenue for future work.



## Bibliography

- [1] Zhenyu Li, Ozan Alp Topal, Özlem Tuğfe Demir, et al. “Mixed Static and Reconfigurable Metasurface Deployment in Indoor Dense Spaces: How Much Reconfigurability is Needed?” In: *2024 IEEE Wireless Communications and Networking Conference (WCNC)*. 2024, pp. 1–6. DOI: 10.1109/WCNC57260.2024.10570594.
- [2] Zhenyu Li, Ozan Alp Topal, Özlem Tuğfe Demir, et al. “Self-Sustainable Metasurface-Assisted mmWave Indoor Communication System”. In: *the process of reviewing, submitted to IEEE Transactions on Wireless Communications* (2026).
- [3] Zhenyu Li, Ozan Alp Topal, Özlem Tuğfe Demir, et al. “Feasibility Study Regarding Self-Sustainable Reconfigurable Intelligent Surfaces”. In: *IEEE Wireless Communications Letters* 15 (2026), pp. 1045–1049. DOI: 10.1109/LWC.2025.3647212.
- [4] Ozan Alp Topal, Zhenyu Li, Mustafa Ozger, et al. “Millimeter-Wave Channel Modeling and Coverage Analysis for Indoor Dense Spaces”. In: *IEEE Transactions on Vehicular Technology* 74.1 (2025), pp. 5–20. DOI: 10.1109/TVT.2024.3463193.
- [5] Theodore S Rappaport, George R MacCartney, Mathew K Samimi, et al. “Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design”. In: *IEEE transactions on Communications* 63.9 (2015), pp. 3029–3056.
- [6] Christopher L Holloway, Edward F Kuester, Joshua A Gordon, et al. “An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials”. In: *IEEE antennas and propagation magazine* 54.2 (2012), pp. 10–35.
- [7] Marco Di Renzo, Alessio Zappone, Merouane Debbah, et al. “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead”. In: *IEEE journal on selected areas in communications* 38.11 (2020), pp. 2450–2525.

## BIBLIOGRAPHY

- [8] Shaokang Hu, Zhiqiang Wei, Yuanxin Cai, et al. “Robust and secure sum-rate maximization for multiuser MISO downlink systems with self-sustainable IRS”. In: *IEEE Transactions on Communications* 69.10 (2021), pp. 7032–7049.
- [9] Yajun Cheng, Wei Peng, and Tao Jiang. “Self-sustainable RIS aided wireless power transfer scheme”. In: *IEEE Transactions on Vehicular Technology* 72.1 (2022), pp. 881–892.
- [10] Yuze Zou, Yusi Long, Shimin Gong, et al. “Robust Beamforming Optimization for Self-Sustainable Intelligent Reflecting Surface Assisted Wireless Networks”. In: *IEEE Transactions on Cognitive Communications and Networking* 8.2 (2022), pp. 856–870. DOI: 10.1109/TCCN.2021.3133839.
- [11] Yijin Pan, Kezhi Wang, Cunhua Pan, et al. “Self-sustainable reconfigurable intelligent surface aided simultaneous terahertz information and power transfer (STIPT)”. In: *IEEE Transactions on Wireless Communications* 21.7 (2022), pp. 5420–5434.
- [12] Friedemann Laue, Vahid Jamali, and Robert Schober. “Beam Training for Self-Sustainable RIS”. In: *2023 IEEE Globecom Workshops (GC Wkshps)*. 2023, pp. 1164–1169. DOI: 10.1109/GCWkshps58843.2023.10464537.
- [13] Dimitrios Tyrovolas, Sotiris A. Tegos, Vasilis K. Papanikolaou, et al. “Zero-Energy Reconfigurable Intelligent Surfaces (zeRIS)”. In: *IEEE Transactions on Wireless Communications* (2023), pp. 1–1. DOI: 10.1109/TWC.2023.3336956.
- [14] Yunfei Chen, Nan Zhao, and Mohamed-Slim Alouini. “Wireless Energy Harvesting Using Signals From Multiple Fading Channels”. In: *IEEE Transactions on Communications* 65.11 (2017), pp. 5027–5039. DOI: 10.1109/TCOMM.2017.2734665.
- [15] Elena Boshkovska, Derrick Wing Kwan Ng, Nikola Zlatanov, et al. “Practical non-linear energy harvesting model and resource allocation for SWIPT systems”. In: *IEEE Communications Letters* 19.12 (2015), pp. 2082–2085.
- [16] Remcom. *Wireless InSite Propagation Software*. <https://www.remcom.com/wireless-insite-propagation-software>.
- [17] Michael Peter, Wilhelm Keusgen, Andreas Kortke, et al. “Measurement and Analysis of the 60 GHz In-Vehicular Broadband Radio Channel”. In: *IEEE VTC-Fall*. 2007, pp. 834–838. DOI: 10.1109/VETEFC.2007.183.
- [18] Robert Felbecker, Wilhelm Keusgen, and Michael Peter. “Incabin millimeter wave propagation simulation in a wide-bodied aircraft using ray-tracing”. In: *IEEE VTC-Fall*. 2008, pp. 1–5.
- [19] Zhenyu Li, Ozan Alp Topal, Özlem Tuğfe Demir, et al. “mmWave Coverage Extension Using Reconfigurable Intelligent Surfaces in Indoor Dense Spaces”. In: *IEEE International Conference on Communications (ICC)*. 2023.
- [20] Ting Wu, Theodore S Rappaport, and Christopher M Collins. “The human body and millimeter-wave wireless communication systems: Interactions and implications”. In: *IEEE ICC*. 2015, pp. 2423–2429.
- [21] Rupinder Singh, Gurleen S Sandhu, Rosa Penna, et al. “Investigations for thermal and electrical conductivity of ABS-graphene blended prototypes”. In: *Materials* 10.8 (2017), p. 881.

- [22] Bill Riddle, James Baker-Jarvis, and Jerzy Krupka. “Complex permittivity measurements of common plastics over variable temperatures”. In: *IEEE Trans. Microw. Theory Techn.* 51.3 (2003), pp. 727–733.
- [23] P Series. “Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900 MHz to 100 GHz”. In: *Recommendation ITU-R* (2012), pp. 1238–7.
- [24] Meisam Razaviyayn. “Successive convex approximation: Analysis and applications”. PhD thesis. University of Minnesota, 2014.
- [25] Omar Mehanna, Kejun Huang, Balasubramanian Gopalakrishnan, et al. “Feasible Point Pursuit and Successive Approximation of Non-Convex QCQPs”. In: *IEEE Signal Processing Letters* 22.7 (2015), pp. 804–808. DOI: 10.1109/LSP.2014.2370033.
- [26] Alan L Yuille and Anand Rangarajan. “The concave-convex procedure”. In: *Neural computation* 15.4 (2003), pp. 915–936.
- [27] Enrico M Vitucci, Mattia Fabiani, and Vittorio Degli-Esposti. “Use of a Realistic Ray-Based Model for the Evaluation of Indoor RF Coverage Solutions Using Reconfigurable Intelligent Surfaces”. In: *Electronics* 12.5 (2023), p. 1173.
- [28] Ehsan Tohidi, Sven Haesloop, Lars Thiele, et al. “Near-Optimal LOS and Orientation Aware Intelligent Reflecting Surface Placement”. In: *arXiv preprint arXiv:2305.03451* (2023).
- [29] Chethan Kumar Anjinappa, Fatih Erden, and Ismail Güvenç. “Base station and passive reflectors placement for urban mmWave networks”. In: *IEEE Transactions on Vehicular Technology* 70.4 (2021), pp. 3525–3539.
- [30] Anders Enqvist, Özlem Tuğfe Demir, Cicek Cavdar, et al. “Optimizing Reconfigurable Intelligent Surfaces for Small Data Packets: A Subarray Approach”. In: *IEEE International Conference on Communications (ICC)*. 2022, pp. 2664–2669. DOI: 10.1109/ICC45855.2022.9839247.
- [31] Venkat Arun and Hari Balakrishnan. “RFocus: Beamforming Using Thousands of Passive Antennas.” In: *NSDI*. 2020, pp. 1047–1061.
- [32] NTT DOCOMO. “Docomo conducts world’s first successful trial of transparent dynamic metasurface”. In: (2020).
- [33] Roman Fara, Philippe Ratajczak, Dinh-Thuy Phan-Huy, et al. “A Prototype of Reconfigurable Intelligent Surface with Continuous Control of the Reflection Phase”. In: *IEEE Wireless Communications* 29.1 (2022), pp. 70–77. DOI: 10.1109/MWC.007.00345.
- [34] Yufei Zhao, Yuan Feng, Afkar Mohamed Ismail, et al. “2-Bit RIS Prototyping Enhancing Rapid-Response Space-Time Wavefront Manipulation for Wireless Communication: Experimental Studies”. In: *IEEE Open Journal of the Communications Society* 5 (2024), pp. 4885–4901. DOI: 10.1109/OJCOMS.2024.3439558.
- [35] Yiqing Wei, Junping Duan, Huihui Jing, et al. “A Multiband, Polarization-Controlled Metasurface Absorber for Electromagnetic Energy Harvesting and Wireless Power Transfer”. In: *IEEE Transactions on Microwave Theory and Techniques* 70.5 (2022), pp. 2861–2871. DOI: 10.1109/TMTT.2022.3155718.
- [36] Ozan Alp Topal, Mustafa Ozger, Dominic Schupke, et al. “mmWave Communications for Indoor Dense Spaces: Ray-Tracing Based Channel Characterization

## BIBLIOGRAPHY

- and Performance Comparison”. In: *IEEE International Conference on Communications (ICC)*. IEEE. 2022, pp. 516–521.
- [37] J Cole Smith and Z Caner Taskin. “A tutorial guide to mixed-integer programming models and solution techniques”. In: *Optimization in medicine and biology* (2008), pp. 521–548.
- [38] MOSEK ApS. *MOSEK - The Mathematical Optimization Software*. <https://www.mosek.com/>. 2024.
- [39] Yuanwei Liu, Xiao Liu, Xidong Mu, et al. “Reconfigurable Intelligent Surfaces: Principles and Opportunities”. In: *IEEE Communications Surveys & Tutorials* 23.3 (2021), pp. 1546–1577. DOI: 10.1109/COMST.2021.3077737.
- [40] Emil Björnson and Özlem Tuğfe Demir. *Introduction to multiple antenna communications and reconfigurable surfaces*. Now Publishers, Inc., 2024.
- [41] Neel Kanth Kundu and Matthew R McKay. “RIS-assisted MISO communication: Optimal beamformers and performance analysis”. In: *2020 IEEE Globecom Workshops (GC Workshops)*. IEEE. 2020, pp. 1–6.