

PROGRAMMABLE SMART RADIO ENVIRONMENTS:
FROM THEORY TO HARDWARE IMPLEMENTATION

KUN WOO CHO

A DISSERTATION

PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF
COMPUTER SCIENCE

ADVISER: KYLE JAMIESON

NOVEMBER 2024

© Copyright by Kun Woo Cho, 2024.

All rights reserved.

Abstract

Today’s wireless networks are evolving rapidly, experiencing an unprecedented surge in traffic volume, radio density, and spectral efficiency demands. This thesis addresses the critical challenges arising from this evolution of next-generation (NextG) wireless networks, focusing on three primary objectives: achieving high data rates, ultra-low latency, and massive connectivity.

To meet these diverse and demanding requirements, this thesis poses a central question: *Can we build a smarter radio environment controlled and learned by software, capable of self-configuring in real-time to meet different application needs?* Current approaches to handle uncontrolled wireless signals are end-to-end, but communication endpoints are limited in their ability to shape inherent propagation behavior. By focusing on changing the environment itself rather than endpoints, this thesis seeks to enhance key aspects of modern wireless networks.

Millimeter-wave technology enables multi-Gbps data rates, but its high-frequency signals are vulnerable to blockage, limiting its practical use. This thesis presents two innovative solutions to overcome this challenge. **mmWall** is a programmable smart surface, installed on buildings and composed of over 4,000 metamaterial elements. It can steer signals through the surface to extend outdoor mmWave signals indoors or reflect them to bypass obstacles. **Wall-Street** is a vehicle-mounted smart surface designed to provide robust mmWave connectivity in high-mobility environments, ensuring reliable communication even in dynamic scenarios. Extending our smart radio concepts to ultra-reliable, low-latency satellite networks, we introduce **Wall-E**, a dual-band smart surface that mitigates signal blockage by relaying full-duplex satellite-to-ground links, and **Monolith**, a smart surface that boosts data rates for inter-satellite communication. To address the growing overhead in massive Internet of Things (IoT) networks, we propose **CLCP**, a machine learning technique that predicts the radio environment to reduce communication overhead. This AI-driven approach complements our programmable surfaces, forming a comprehensive smart radio solution.

Given the highly complex nature of real-world systems, conceptual models alone are insufficient to fully explain them. Our solutions are implemented in physical hardware prototypes, integrated with existing network protocols, and rigorously tested through experimentation. This thesis thus offers a concrete answer to the above central question, laying the foundation for software-controlled smart radio environments in NextG wireless networks.

Acknowledgements

When beginning my Ph.D., I envisioned research as a solitary journey into the unknown, much like the astronaut protagonist in the movie *Gravity*, venturing alone through space to return to Earth. Looking back over the past six years, I've realized that what made this journey successful was the gravity of relationships. The support from mentors, collaborators, friends, and family served as a lifeline, keeping me on course, helping me bounce back from occasional failures, and anchoring my personal and professional growth. This journey would not have been possible without them.

I want to start by thanking my advisor, Kyle Jamieson, for his mentorship. It has been invaluable in shaping my growth as a researcher and as a person. Kyle's encouragement and advice were valuable during the challenging times when our prototype faced multiple failures and unexpected delays due to the pandemic. His persistent faith in me and the project's potential gave me the strength to persevere, allowing me to learn from failures and ultimately succeed. Kyle has also motivated me to expand my boundaries and grow into a researcher capable of making significant contributions. I am truly grateful for the opportunity to have worked under his supervision.

I thank my dissertation committee: Omid Abari, Andrew Appel, Ranveer Chandra, Yasaman Ghasempour, and Wyatt Lloyd, as well as my general exam committee, Amit Levy and Jennifer Rexford. Their insightful comments and feedback have significantly enhanced the quality of this dissertation. I'm particularly grateful to Yasaman for introducing me to relevant literature, and to Omid and Ranveer for providing alternative viewpoints and helping me better define the scope of my work. The collective expertise of the committee has been invaluable to this research. Omid, in particular, became a collaborator in the early years of my PhD, when I was still new to hardware research. His guidance was invaluable, particularly in teaching me research methods for hardware and mmWave technologies.

I thank Andrea Goldsmith for discussions and invaluable advice on navigating the early stages of an academic career. Her insights and guidance on career development have been instrumental in determining my next steps after graduation.

I thank my collaborators, Mohammad Mazaheri, Jeremy Gummeson, Marco Cominelli, Francesco Gringoli, Joerg Widmer, Ivan Seskar, and Prasanthi Maddala. Mohammad taught me so much, from electromagnetic simulations to PCB fabrications. His expertise contributed significantly to the success of my first hardware project. I greatly enjoyed working with the WINLAB team. Ivan not only advised my research but was also always available to consult on my career development. I will miss the memorable moments, like when we waited for the solar eclipse and got Cold Stone in Malaysia. I deeply thank Prasanthi for her extensive help with coding the COSMOS testbed. She was always there to help me, even when I needed services restarted late at night. Without her support, I wouldn't have been able to complete my final project. I also thank Jakub Kolodziejcki for 3D printing the winch cover, ensuring the safety of all my ten fingers. And a special thanks to Jennifer Shane for preventing my unintentional DDoS attack on the WINLAB network. I will miss the weekly cookie time at WINLAB, selecting Ivan's photos of wildlife animals.

I deeply thank my undergraduate advisor, Wenyao Xu, for his teaching and mentoring. I first met him when I was a sophomore majoring in chemical engineering. He was the first to see my potential and persuaded me to switch my major to computer science, marking the very beginning of my journey as a researcher. He taught me so much, from developing research ideas and writing papers to collaborating with peers and presenting my work to other scientists. I also thank my internship advisor, Cecilia Mascolo. Her support was invaluable during my graduate school application process. She not only provided insightful advice on my applications but also generously reached out to her professional network. I truly appreciate the profound impact her guidance has had on my academic journey. I also thank the Cambridge program director, Satpal Singh, for his mentoring and our intellectual conversations. Our weekly meetings, intended for progress reports, often extended into hours

of conversations covering various research topics, from fly behavior to bat echolocation. I appreciate Akanksha and Andrea's patience during these meetings, even though they often poked me under the table to make me stop talking.

I deeply thank the entire PAWS team: Zhenlin An, Xuyang Cao, Yunxiang Chi, Sai Srikar Kasi, Zhuqi Li, Minsung Kim, Waseem Ozan, Yao Peng, Longfei Shangguan, Abhishek Kumar Singh, Haoran Wan, Allen Welkie, Yaxiong Xie, Fan Yi, and Zijian Qin. Their regular discussions and feedback on my research, presentations, and talks have been invaluable. Throughout my time at Princeton, I've been extremely fortunate to be surrounded by such exceptional colleagues. I miss the time when Yaxiong, Zhuqi, Minsung, Srikar and I got a VIP card at Tiger Noodles, and how all of us ordered the same dish for nearly six months. I also thank the amazing CS staff, particularly Mitra Kelly for handling numerous equipment orders, Michael Estep for his assistance when I had to dispute with PCB manufacturers and equipment rental companies, and Nicki Mahler for guiding me through the graduation process. Also, I deeply thank Bert Harrop from the physics department for training me and significantly improving my soldering skills.

I thank my CS friends for friendship and memories. Yue Tan has been my travel mate to San Francisco, Seattle, and Vermont as well as a movie mate and roommate. I am grateful that we survived the deadly hiking trail at Olympic National Park where we couldn't see a sign of human existence for hours. Graduation from life could've happened earlier than our Ph.D. Srikar has been the person I've relied on the most in Princeton. We went through the entire grad school together, and his presence made the journey significantly more manageable and enjoyable. I also thank him for all our flat-earth jokes and for always saying yes to coffee breaks. Abhishek has been a friend who genuinely laughs at my jokes, making me proud of myself when I recall it before going to sleep. I also thank him for always saying yes to workouts at the Dillon gym and Indian foods in Edison. I will miss many sleepless horror game nights and Mango Mango bounds with Yue, Srikar, and Abhishek. I hope that they

take at least three months off and visit me in Houston to finish *Evil Within 2*. I also thank Sophia Yoo for hosting a memorable board game night and for our chimaek time.

I want to thank the two Jae sisters: Heejae Jang and Wonjae Lee. Although they left Princeton long ago, our daily texts have made it feel as if they were always by my side, providing me with huge emotional support. Heejae, a true New Yorker and a connoisseur of daily life with impeccable taste, never fails to give excellent recommendations. I'm thrilled that she has become a runner, and I hope we can run the New York Marathon together soon. Wonjae was my roommate during the pandemic, and our daily walks and chats along the D&R trail helped me maintain my mental balance during that challenging time. We three have so many memories together, one of the most memorable being when we built a gingerbread house at Heejae's place. I'm grateful that we were at least able to afford one house in Manhattan.

I'm grateful to my friends from UB for our friendship over the years. Doris Kwan, a West Village master, has kept me updated with NYC trends. During times when I felt overwhelmed in Princeton, she provided refreshing breaks by inviting me to K-pop concerts and hosting events in NYC like picnics in Central Park. I'm also thankful to her for introducing me to marathon running. I thank Angus Lam, a Williamsburg master, for taking me to many hipster coffee shops and cocktail bars in NYC and DC. I look forward to joining the silver town he plans to build on the west coast. I also thank Emily Walker for her random hip-checks. I miss sleeping over at her place in Long Island, and I hope to visit her in Portland soon.

I am grateful to have been part of the Korean graduate student & postdoc community at Princeton. I want to start by thanking Daniel Park, Jongbeom Park, and Kevin Suh because our Canada trip was by far the best trip I had during my Ph.D. I want to thank: Daniel for hosting us at his homes in Princeton and Canada, and for planning the entire trip, and for entertaining us by crying during his own defense; Jongbeom for always accepting my marathon bets and willingly buying me dinners; and Kevin for organizing so many events

and always being a yes man. I am looking forward for our ski trip even though no one knows when it will happen.

I'm grateful for my girls' night friends: Youn Kyoung Cho for her pterosaur-like screams at EDM festivals and her hip dances in Overwatch and real life; Esther Han for watching 20 movies with me, including 'Manchester by the Sea,' for her bravery to refuse tickets, and her counsel during difficult times; Jiwon Lee for paving a way through the crowds on the Brooklyn Bridge, ramen after EZoo, and our speakeasy crawls; Jinhee Park for her innate humor and karaoke at her home. As her #1 fan, I hope for her stand-up comedy in future; and Gawoon Shim for our night run in Philly and for her companionship during our anxious train ride back from NYC.

I'm grateful for my tennis and pomodoro friends: Junhwan Bae for our Gen Z debates and for dancing to 'Paradise Lost' at Sunnie's apartment; Sunnie Kim for hosting house parties with endless wine and her adventurous spirit in exploring new restaurants with me; Kyoungjoo Noh for his happiness and for introducing me to Texas Hold'em; and Yeji Park for our half-studying, half-chatting study sessions in Lakeside Commons and our croissant tour in NYC. I miss the countless weekend trips to Edison with Sunnie and Yeji, as well as our memorable Philly trip where we unintentionally competed in walking speed, ending up walking over 25K steps.

I'd like to thank my Equad friends: Yebin Chon for introducing me to archery; and Geon Hwi Kim for including me in his workout Instagram stories. I also appreciate Yebin, Jeremy, and Jenny for sharing free food announcements in their department.

I want to thank my S2O friends: Yoon Chang for the exciting ATV rides and his K-pop dances on our way back from Reunion; Roy Jang for our motivational weight loss bet and for carrying me back to Airbnb from S2O; and Ahn Young Jeong for our Nassau Hall bouldering and her efforts to enter me into S2O.

I want to thank my running and workout mates: Donghun Lee for coaching me for the tennis contest; Simon Park for training with me for the Princeton Half and pushing my

pace during the race; Seoyeon Yang for her candy surprises and for taking me to the Dillon gym even when I wore a leather jacket and jeans; and Sohye Yoon for joining me on my Oppenheimer run.

I also thank Jihye Jeon for our memorable Puerto Rico trip; Sanghak Jeon and Gyoonho Kong for being my neighbors, letting me play with Leo, and for our teamwork in Overwatch and Deep Rock Galactic; Jiyoung Kang for our Mamamoo conversations; Jiwoo Park for cooking memorable dinners at Dorsia; Yoonseo Zoh for our dry pot chat and her strict Yutnori rules; and baristas at Rojo's Roastery for always making me the best Americano.

I'm truly blessed to have connected with these individuals whose gravitational force has kept me tethered to Princeton. Princeton was so much fun because of them.

Above all, I am wholeheartedly thankful to my parents for their unconditional love and many sacrifices. My mom brought me and my brother to the US and raised us while pursuing her PhD. As a child, I couldn't fully understand the challenges, but now, completing my own PhD, I deeply appreciate her endurance and sacrifices. She has been a constant source of perseverance and academic inspiration for me. My dad, a resilient and positive presence in my life, stayed in Korea to support our family in the US. He consistently maintained an optimistic outlook, shielding our family from his concerns. I'm thankful for his behind-the-scenes hard work and comforting presence. I'm thankful to my brother, Jason Kun Wung Cho, for his thoughtfulness and the occasional gifts. Despite our childhood squabbles, I know that he is an incredibly caring person and have benefited greatly from it. I also thank the cutest boy, Khan Cho. We initially adopted him to provide a safe home, but he filled our house with joy. Our daily video calls brought me comfort and relief during my difficult times. I'm deeply thankful for his endurance in overcoming his past hardships and for his recovery and happiness. I also extend my heartfelt thanks to my grandparents, Sooja Kim and Jaeyoon Choi, my aunt, Hyunju Choi, and my late grandparents, Jongnim Ko and Seongjin Cho, for their enduring love and care. I would not have been able to complete this journey without my family, and I dedicate this dissertation to them.

To my parents, Hyunmi Choi and Sangbong Cho.

Contents

Abstract	3
Acknowledgements	5
List of Tables	17
List of Figures	18
1 Introduction	29
1.1 NextG Wireless Networks	30
1.2 Challenges in Realizing NextG Services	32
1.2.1 mmWave Networks	32
1.2.2 Satellite Networks	34
1.2.3 Massive IoT Networks	35
1.3 Thesis Contributions	36
1.3.1 Programmable Metamaterial Surfaces	37
1.3.2 Machine Learning for Wireless Networks	40
1.4 Statement on Publications	41
1.5 Funding Acknowledgements	42
I Pushing the Limits of mmWave Networks	43
2 A Steerable, Transflective Metasurface for mmWave Networks.	44
2.1 Related Works	47

2.2	Primer: Huygens Metamaterials	50
2.3	mmWall: Design and Analysis	51
2.3.1	Surface Hardware	51
2.3.2	Surface Control	57
2.3.3	Link Layer Design	60
2.4	Implementation	62
2.5	Evaluation	65
2.5.1	Methodology	65
2.5.2	In-situ Performance	67
2.5.3	Multi-armed Beams	72
2.5.4	Microbenchmarks	73
2.6	Discussion	77
2.7	Conclusion	79
2.8	Appendices	79
2.8.1	Appendix – Unit Cell Electromagnetic Analysis	79
2.8.2	Appendix – Path Loss Model	84
2.8.3	Appendix – Multi-armed Beam via Time Modulation	87
3	Metasurface-Enabled NextG mmWave for Roadside Networking	91
3.1	Primer: 5G Mobility Management	94
3.1.1	Initial Attachment and Tracking	94
3.1.2	5G SA Handover	95
3.2	Wall-Street: Design	98
3.2.1	Surface Hardware	98
3.2.2	Attachment and Tracking	100
3.2.3	Cell-to-Cell Handover	103
3.3	Implementation	109
3.3.1	Surface Hardware	109

3.3.2	COSMOS Testbed	110
3.4	Evaluation	112
3.4.1	Methodology	113
3.4.2	End-to-End Throughput and Delay	114
3.4.3	Microbenchmarks	119
3.5	Related Works	124
3.6	Conclusion	125
3.7	Appendix	125

II Reliable and Low-Latency Satellite Networks 127

4 Dual-band Reconfigurable Metasurfaces for Satellite Networking 128

4.1	Design Challenges	131
4.2	Failed Attempts: Design and Analysis	132
4.2.1	Meta-Atom Designs	133
4.2.2	Implementation	134
4.2.3	Near-Field VNA Measurements	135
4.3	Wall-E: Design and Analysis	137
4.3.1	Building the Surface: Meta-Atoms	137
4.3.2	Establishing a Surface-Satellite Link	141
4.3.3	Enhancing Satellite-Satellite Handover	143
4.4	Feasibility	144
4.5	Related Works and Discussion	147
4.6	Conclusion	148

5 Metasurface-Enabled Spatial Multiplexing for Rank-Deficient Satellite Networks 149

5.1	Primer: Orbital Angular Momentum	152
-----	--	-----

5.1.1	Is OAM a New Concept?	153
5.1.2	Understanding the Benefits of OAM	154
5.2	Monolith: Design and Analysis	157
5.2.1	Surface Design	158
5.2.2	Data Encoding and Decoding	161
5.3	Challenges and Discussion	162
5.4	Conclusion	164

III Intelligent Massive IoT Networks 165

6 Cross-Link Channel Prediction in Massive IoT Networks 166

6.1	Cross-Link Correlation Model	168
6.2	Primer: ML Background	171
6.3	CLCP Design	173
6.3.1	Cross-Link Prediction Model	176
6.3.2	Model Interpretability	180
6.3.3	Scheduling and Resource Allocation	182
6.4	Implementation	186
6.5	Evaluation	188
6.5.1	Experimental Methodology	188
6.5.2	End-to-End Performance	191
6.5.3	Microbenchmark	193
6.6	Discussion and Limitations	198
6.7	Supporting Applications	199
6.8	Related Works	201
6.9	Conclusion	204

7 Concluding Remarks 205

List of Tables

1.1	Major current satellite internet service providers and their primary frequency band allocations (GHz) [33].	35
2.1	mmWall: comparison to existing surface works.	47
2.2	mmWall design parameters.	83

List of Figures

1.1	Three distinct services supported by current 5G New Radio and beyond. . . .	30
1.2	Unique challenges of mmWave, satellite, and massive IoT networks.	32
1.3	Our solutions introduce programmable smart radio environments through various surface prototypes, including mmWall (24 GHz), in-vehicle Wall-Street (26 GHz), Wall-E (10 and 15 GHz Ku-band), and Monolith, as well as an ML-based OFDMA scheduling system.	37
2.1	mmWall re-focuses outdoor coverage indoors towards the user and potentially around obstacles, provides path diversity indoors by reflection, and splits an incoming beam for fast link establishment (<i>upper</i> : without mmWall; <i>lower</i> : with mmWall).	45
2.2	mmWall’s design converts an incident mmWave beam to a refracted (or reflected, not shown) beam via field discontinuities created by current in its resonators. <i>Inset</i> : magnetic meta-atoms are shown in front of the electric meta-atoms.	51
2.3	Unit cell response <i>v.</i> electric- and magnetic-side control voltages U_E and U_M — (a) : magnitude and (b) : phase. (c) : HFSS simulation (<i>left</i>) and near-field, real world VNA measurement (<i>right</i>)— arrows indicate control voltage pairs that yield a 360° phase shift of the incoming signal, with high transmission or reflection magnitude.	52

2.4	mmWall, prior Huygens unit cell designs (<i>top</i> : magnetic; <i>bottom</i> : electric side), and equivalent circuits. Green arrows indicate the oscillating current loops, and V_{in} indicates where the input voltages connect.	54
2.5	mmWall design parameter sensitivity analysis.	56
2.6	Biasing line designs: notable failed attempts include (a) straight microstrip, (b) coil inductor, and (c) radial stub. (d) mmWall uses an inner meander line for magnetic, and an outer meander line for electric meta-atoms.	57
2.7	mmWall’s refractive link establishment, angular reciprocity property, tracking, and reflective link establishment.	60
2.8	mmWall’s hardware implementation, transmissive (‘lens’) and indoor reflective (‘mirror’) evaluation scenarios. We placed mmWall at the same location for both scenarios.	63
2.9	mmWall’s <i>ribs</i> , comprised of our proposed meta-atom design fabricated on a Rogers printed circuit board.	64
2.10	mmWall’s FR4 holder/control board.	65
2.11	mmWall’s SNR improvement over the best NLoS environment path in two scenarios: (a) when both transmitter and receiver are located indoors (<i>upper</i> : transmitter facing mmWall perpendicularly; <i>lower</i> : transmitter facing 30° away from mmWall) and (b) when the transmitter is located outdoors (<i>left</i> : transmitting perpendicularly; <i>right</i> : transmitting at a 30° off-angle). We use the following notations: mmWall ★, transmitter ▲, receiver ○. ● indicates no signal.	68

2.12	The SNR improvement from the use of one or more mmWalls at various receiver locations in indoor-to-indoor (<i>upper</i>) and outdoor-to-indoor (<i>lower</i>) scenarios. SNRs collected from a given receiver location are plotted on the same y-axis value (<i>left</i> : CDFs of the best environment SNRs in black curves alongside the SNRs of mmWall links at the corresponding receiver location in rectangles. The maximum SNRs between two mmWalls placed in different locations are denoted with downward triangles; <i>center</i> : the best available SNRs with or without one or more mmWalls; <i>right</i> : the SNR gains attained with one or more mmWalls compared to the best environment path in various Rx locations).	70
2.13	The SNR improvement (multiple mmWalls) for dynamic links (<i>upper</i> : indoor-to-indoor; <i>lower</i> : outdoor-to-indoor scenarios). β is a blockage probability.	71
2.14	Evaluation of mmWall's multi-armed beams.	72
2.15	Microbenchmarks evaluating (<i>left to right</i> :) surface steerability, performance sensitivity of the incident wave angle, and angular reciprocity. Empirical points are denoted with markers, with simulation curves.	74
2.16	Microbenchmarks evaluating (<i>left to right</i> :) surface size and frequency bandwidth.	75
2.17	mmWall's phase coverage and consistency (VNA measurement) across different frequencies. The curves indicate the voltage pairs (U_M, U_E) that provide -180° to 180° phase shift with the step of 15° at 24.5 GHz. The phases are unwrapped across the mmWall's operating bandwidth.	76
2.18	Microbenchmark on meta-atom controllability and sensitivity with near-field, real world VNA measurements.	77
2.19	<i>Left</i> : C_{var} as the voltage applied to varactor changes, modeled with SPICE simulation; <i>Right</i> : mmWall element normalized beam pattern $F(\theta)$ simulated with HFSS and fitted function.	81

2.20	<i>Upper</i> : a set of voltage levels applied to the magnetic and electric meta-atoms U_M and U_E for transmissive steering; <i>Lower</i> : voltage levels applied for reflective steering.	83
2.21	Schematic illustration of multi-armed beam generation at $f_c - \Omega$ and $f_c + \Omega$ and multi-armed beam-steering. f_c is the carrier frequency, and Ω is the voltage modulation frequency.	87
2.22	Time-modulated multi-beam. Simulated beam efficiency versus frequency shift in log scale.	89
2.23	Simulated steering angle accuracy with time modulation. GT, SF, DF, and DB stand for groundtruth, single forward, double forward, and double backward, respectively.	90
3.1	Wall-Street’s design innovations electronically split, shape, and steer mmWave transmissions in real time to enable seamless roadside 5G New Radio networks (upper : without Wall-Street; lower : with Wall-Street).	92
3.2	5G SA beam acquisition and beam tracking.	95
3.3	5G SA Xn handover: (1) Preparation phase; (2) Execution phase; (3) Completion phase. Data communication is denoted with orange and yellow rectangles.	96
3.4	Wall-Street design overview: (a) Initial UE-RAN attachment (Section 3.2.2); (b) preparation using transfective surface power measurement to multiple cells (Section 3.2.3); (c) beam combining to enable make-before-break dual-cell connectivity (Section 3.2.3); (d) handover completion (Section 3.2.3).	97
3.5	VNA measured unit cell response: reflection and transmission through the surface.	99
3.6	Wall-Street’s multi-beam operational modes (<i>left to right</i>): bi-directional split and combination for handover preparation and uni-directional split and combination for handover execution.	99

3.7	Wall-Street process: beam attachment and measurement report process corresponding to Fig. 3.4(a) and Fig. 3.4(b), respectively. Beam attachment and measurement report require four SS bursts and one SS burst of five-millisecond duration.	100
3.8	Beam tracking: Wall-Street’s surface-enabled beam tracking leverages neighboring gNBs to add path diversity.	102
3.9	Wall-Street’s handover timeline: (1) Preparation, (2) Execution, and (3) Completion phases are indicated.	104
3.10	Measurement power translation: Inferring 3GPP-standard measurement report power readings based on Wall-Street’s reflective measurement report readings.	105
3.11	Wall-Street-COSMOS testbed integration: we integrate Wall-Street with COSMOS USRP-based gNBs and UE.	109
3.12	We illustrate a simplified control flow of Wall-Street-integrated COSMOS testbed.	110
3.13	Wall-Street’s hardware implementation and testbed: showing (a) individual hardware components, (b) indoor experimental UE-side setup, and (c) outdoor in-vehicle UE-side configuration.	111
3.14	Experimental gNB-side testbed setup and map: we deploy gNBs on the first floor of a lab facing an outdoor parking lot. The vehicle is approximately 20 to 30 m from the gNBs. The arrow indicates the direction that each gNB node is facing at 0°, and each node sweeps the beam from -60° to 60°.	113
3.15	Throughput and round trip time (RTT) achieved by four congestion control algorithms. The right and lower edge of the box represents the 10% percentile, respectively. The left and upper edge give the 90% percentiles. The intersection point of the horizontal and vertical error bar represents the median value.	115

3.16	Throughput and delay of UE 1 (<i>upper</i>) and UE 2 (<i>lower</i>) achieved without Wall-Street (<i>left</i>) and with Wall-Street (<i>right</i>) as the vehicle is moving along the same trajectory. Vertical dashed lines indicate the handover event.	116
3.17	Impact of speed: throughput, RTT, HO count (<i>left to right</i>) with different speed, ranging from 5 km/h to 15 km/h.	117
3.18	Measurement report impact: increase trend of sequence numbers during the entire 30s experiment (<i>left</i>) and handover preparation (<i>right</i>).	118
3.19	Make-before-break impact: packet reception rate of a link from each gNB without Wall-Street (<i>left</i>) and that of each gNB and of two links (packet duplication) from two gNB with Wall-Street along the same trajectory (<i>right</i>).	118
3.20	In-vehicle mmWave coverage. ●: no signal.	119
3.21	RSRP heatmap (<i>upper</i>) and corresponding gNB beam angles (<i>middle</i>) and UE/surface angles (<i>lower</i>) heatmaps (<i>left</i> : without Wall-Street; <i>right</i> : with Wall-Street and UE beam fixed to the surface). In-vehicle location of the UE is the same in both scenarios.	120
3.22	Multi-link operations. We use the following notations: gNB ▲, UE □, blockage ■.	121
3.23	Wall-Street's beam patterns over four operations. ×: empirical spectrum analyzer data; ■: benchmark data with the surface is set to 0V. The steering angles are highlighted in the background. 0 dB represents the spectrum analyzer noise floor (<i>upper</i> : $-40^\circ/40^\circ$; <i>middle</i> : $-40^\circ/55^\circ$; <i>lower</i> : $-68^\circ/45^\circ$). Red and blue markers are measured at 26 GHz and 26.1 GHz, respectively.	122
3.24	Power ratio adjustment with the steering angle of $-40^\circ/55^\circ$ (<i>upper</i> : beam combination; <i>middle</i> : bi-directional combination; <i>lower</i> : bi-directional split).	123
3.25	Wall-Street's unit cell design parameters. The pink box is an air box that defines the boundary conditions to simulate the unit cell.	126
4.1	Various use cases for a satellite smart surface.	129

4.2	Strawman (iii) features magnetic meta-atoms on the front side and electric meta-atoms on the back side (meta-atoms on left for uplink and meta-atoms on right for downlink). Green arrows indicate the direction of current oscillation. The design uses 1 oz copper and a Rogers 4003C dielectric substrate with a thickness of 0.813 mm. All meta-atoms are connected to a single ground meander line via through-hole vias, linking the front and back sides.	132
4.3	Hardware prototype of strawman (iii) design 1 and 2 (<i>upper</i> : ribs, comprised of magnetic and electric meta-atoms fabricated on a Rogers printed circuit board; <i>lower</i> : complete hardware prototypes).	134
4.4	Near-field testing with two-port Anritsu MS4647B Vector Network Analyzer (VNA). We supply different voltage levels to our prototypes using an EVAL-AD5370 DAC.	135
4.5	Simulated and VNA-measured Huygens pattern for Design 1 at downlink frequencies (10.8 to 11.2 GHz) and uplink frequencies (13.8 to 14.2 GHz).	136
4.6	Simulated and VNA-measured Huygens pattern for Design 2 at downlink frequencies (10.8 to 11.2 GHz) and uplink frequencies (13.8 to 14.2 GHz).	136
4.7	The magnetic and electric meta-atom design considerations (<i>top</i> : design schematics where the magnetic and electric meta-atom are colored in yellow and orange, respectively; <i>bottom</i> : transmission responses in magnitude $ S_{21} $). The electric and magnetic meta-atom inside a dotted black line are the designs selected for Wall-E.	138
4.8	Equivalent circuit and its transmission response in magnitude across frequencies and voltages. The downlink and uplink frequency regions are colored in grey.	140

4.9	Huygen’s transmissive and reflective pattern in magnitude and phase at downlink and uplink frequency with different voltages applied to electric meta-atom U_E and magnetic meta-atom U_M . The path denoted by the black dotted curve shows 360° , high amplitude phase coverage.	141
4.10	Wall-E’s bi-directional beam-steering in FDD communication. Due to its angular reciprocity, the steering angle of downlink θ_{DL} is equal to the steering angle of uplink θ_{UL}	142
4.11	An electric meta-atom’s surface currents and electric fields and a magnetic meta-atom’s surface currents and magnetic fields at 10 GHz and 15 GHz. . . .	143
4.12	The transmission and reflection efficiency as Wall-E steers the downlink and uplink beam.	144
4.13	The transmission efficiency of a beam splitted by Wall-E. The power split is even for (a) and uneven for (b).	144
4.14	Simulated SNRs of transmissive links with varying surface size, steering and incident angle compared to SNRs of a free space path and wall penetration without Wall-E.	146
5.1	Various use cases for Monolith. The sender (Alice) backscatters the signals, and the reader (Bob) decodes the signals. With Monolith, a metasurface artificially creates multiple orthogonal beams and spatially multiplexes them.	150
5.2	Mixed-state OAM generation using Monolith.	153
5.3	<i>Left</i> : Channel rank enhancement by OAM. <i>Right</i> : Capacity comparison between OAM and MIMO.	155
5.4	Schematic model of meta-atom and metasurface. Our meta-atom consists of several design parameters, including Δ , R_1 , R_2 , W_c , and L_c	158
5.5	<i>Left</i> : simulated magnitude; <i>right</i> : phase of reflected coefficient for the meta-atom across different voltage values from 0 to 10 V.	159

5.6	<i>Upper</i> : an applied voltage distribution; <i>lower</i> : phase distribution by applying the voltage distribution. X- and Y-axis represent the coordinate of each meta-atom with the meta-atom at the center being (0,0) mm.	160
5.7	<i>Upper</i> : a phase distribution of an OAM state from -2 to 2; <i>middle</i> : a magnitude of the generated OAM beams; <i>lower</i> : a phase of the generated OAM beams.	162
6.1	<i>Left</i> : Previous work on cross-band channel prediction infers a downlink channel at frequency f_2 using the uplink channel at frequency f_1 on the <i>same</i> link ([142, 14]). <i>Right</i> : CLCP infers the channel to Sensor 2 using channel measurements from Sensor 1.	166
6.2	<i>CLCP's mechanism</i> : Time of flight and angle of arrival for two nearby IoT sensors (<i>upper</i> and <i>lower</i> , respectively). While each sensor (link) has a distinct set of static wireless paths, their parameters both indicate reflections off the same moving object, highlighted in red dotted circles.	167
6.3	Cross-link correlation modeling under a multipath propagation scenario with two transmitters T_1 and T_2 , one receiver R , and a reflecting obstacle lying on the y-axis at X . Here, α and β are reflected path lengths.	170
6.4	<i>System overview for uplink transmission</i> : (1) an AP receives uplink traffic from multiple users simultaneously; (2) when channels become outdated, the AP extracts the path parameters of partial CSIs estimated from the latest OFDMA packet and predicts unobserved links using CLCP and unobserved bands using CBCP in a server; (3) Then, the AP schedules uplink traffic based on predicted CSIs and triggers the users. (4) Finally, the AP receives the scheduled packet.	174

6.5	CLCP ML model with N measured channels, each represented as a set of wireless path parameters $\{\theta_l, d_l, a_l, \phi_l\}_{l=0}^L$ with L paths estimated from measured channels. Each set of the parameters is served by a Single-view Encoder network E_i ($i \in [1, N]$) that compresses the measured wireless path information of its dedicated radio and outputs variational parameters μ_i and σ_i . The Multi-view Combiner integrates all variational parameters into μ and σ , based on which Single-view Decoder networks D_K generate a set of path parameters that are unobserved. If any input channel is not observed, CLCP drops the respective encoder network (E_2 , for example).	175
6.6	Multi-step training paradigm.	180
6.7	2D t-SNE visualization of the latent space of two adjacent clients before (<i>left</i>) and after training (<i>right</i>). Encoded CSI instances of Radio 1 and 2 are highlighted in red and blue, respectively.	181
6.8	A fully trained latent feature embedding alongside exemplary data points and their corresponding path parameters. The x-axis represents the Angle of Arrival (AoA), while the y-axis shows the Time of Flight (ToF). Data points that are positioned close to each other for each radio reveal similar path parameters.	182
6.9	802.11ax resource unit (RU) locations in a 40 MHz channel with 5 layers. RU locations in a 20 MHz channel with 4 layers are highlighted in red. The RU index at each layer increases from left to right. An intermediate 26-tone RU is denoted as C.	183
6.10	OFDMA transmissions of four users with an equal amount of data D in a $20MHz$ bandwidth channel. T_s is a scheduled transmission duration.	184

6.11	<i>Our scheduling and resource allocation algorithm:</i> (1) rate adaptation and bucketing; (2)-(3) UE selection and capacity comparison between the upper two layers; (4) UE selection and capacity comparison between previously selected RUs and the next layer.	185
6.12	CLCP preliminary experimental testbed floor plans and radio hardware with different operating bandwidths.	187
6.13	The variance of channel amplitude over time with and without human mobility for testbed (a) and (b). The channel amplitude is measured in AU as reported by the tool. Broadcom chipset’s maximum AU is 2000, while Atheros chipset’s maximum AU is 500.	190
6.14	End-to-end performance on throughput and power consumption: (a) aggregated throughput across time for every 500 ms, (b) throughput across users for 20, 40, 80, and 160 MHz bandwidth, and (c) device sleep time over the entire transmission duration and the total number of Target Wake Time (TWT) triggered on every user.	192
6.15	Microbenchmark on the prediction accuracy, runtime, overhead reduction, and scheduling performance under two scenarios: testbed (a) for cashierless store with 8 devices and testbed (b) for smart warehouse with 19 devices.	194
6.16	Aggregated throughput of joint OFDMA and MU-MIMO.	197
6.17	Overhead comparison across different protocols.	197

Chapter 1

Introduction

Wireless networks are undergoing a revolutionary transformation, scaling in spectral efficiency, radio count, and traffic volume at an unprecedented rate. At the forefront of this evolution are next-generation (NextG) networks, which aim to support three distinct services as shown in Fig. 1.1: (1) enhanced mobile broadband (eMBB), (2) ultra-reliable low latency communication (URLLC), and (3) massive machine type communication (mMTC) [1]. Each service poses unique challenges and requirements.

eMBB services cater to the ever-increasing demand for high-speed connectivity, aiming for peak download speeds of 10 Gbps. URLLC is focused on reliability and low latency, targeting a 1 ms end-to-end latency to support time-sensitive applications. mMTC, on the other hand, prioritizes uplink-centric communication of a massive number of low-rate and low-power devices — up to one million per square kilometer [16, 29].

To meet these diverse and demanding requirements, the 3rd Generation Partnership Project (3GPP) has introduced new types of wireless networks [3, 4, 2, 5]. This thesis delves into these networks, addressing the challenges each presents and proposing new solutions to fully satisfy the multifaceted needs of NextG networks.

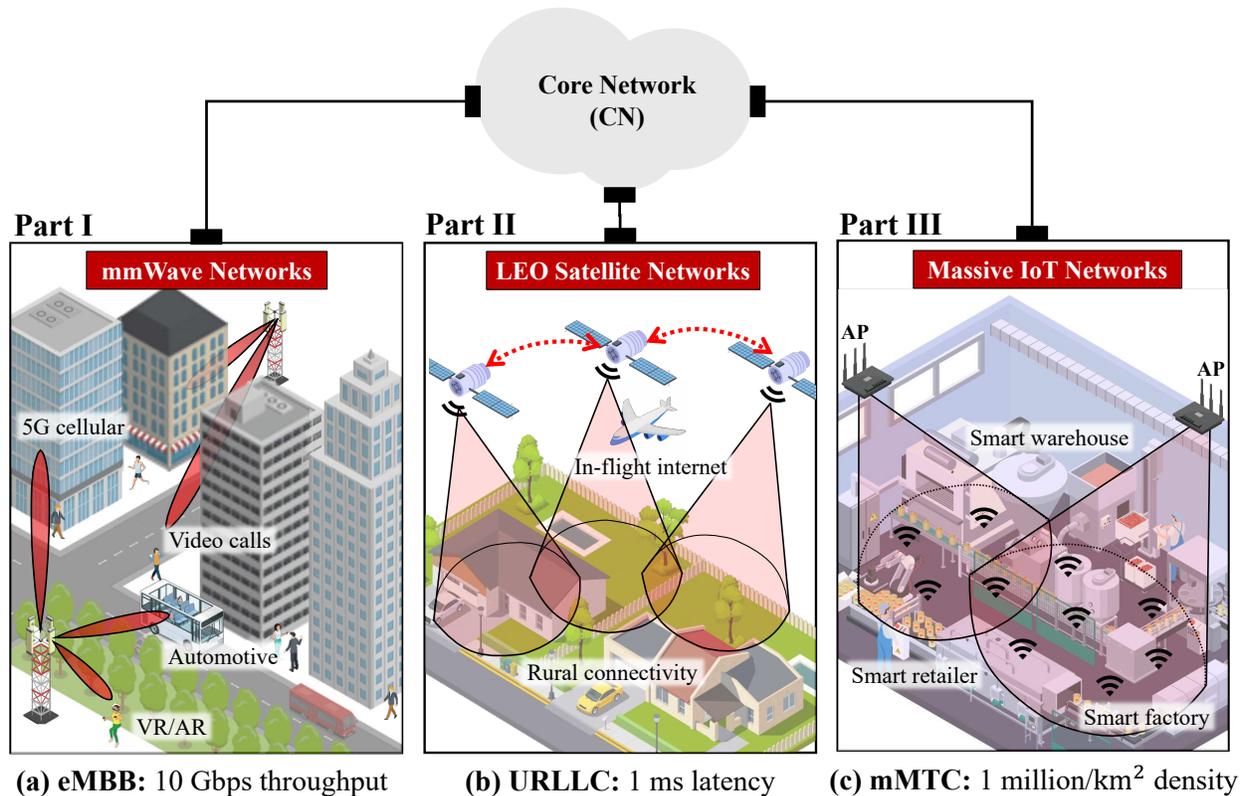


Figure 1.1: Three distinct services supported by current 5G New Radio and beyond.

1.1 NextG Wireless Networks

Millimeter-wave (mmWave) spectrum has emerged in the 5G/6G era as a key enabler towards realizing eMBB, fulfilling user demands for high spectral efficiency wireless networks. Higher carrier frequencies offer greater network capacity: for instance, the maximum carrier frequency of the 4G LTE band at 2.4 GHz provides an available spectrum bandwidth of only 100 MHz, while mmWave (above 24 GHz) can easily hold spectral bandwidths five to ten times greater, enabling multi-Gbit/sec data rates. Hence, mmWave spectrum enables a plethora of mobile applications that are currently infeasible due to their requirements for very high data rates, such as virtual reality, augmented reality, and seamless streaming of ultra high-definition video [115].

To support URLLC, 3GPP has defined a set of technical specifications, including the adoption of non-terrestrial networks (NTN) [2]. Among various NTN, Low-Earth Orbit

(LEO) satellites are considered as the most promising solution for 5G/6G URLLC services. Several well-known companies, such as SpaceX and OneWeb, have already begun deploying LEO satellite internet services, gaining over 1 million users [36]. LEO satellite networks offer three key advantages: (1) low-latency connectivity for vehicles, trains, and planes through inter-satellite link paths; (2) reliable connections through multiple communication paths and redundancy provided by hundreds or thousands of satellites; and (3) global coverage, including areas where terrestrial infrastructure is impractical or unavailable, with resilience to natural disasters. These characteristics make LEO satellite networks a promising technology for mission-critical applications such as intelligent transportation systems, drone control, and first responder communications [88, 119].

By enabling the connectivity of a vast number of Internet of Things (IoT) devices and sensors, mMTC is transforming industries and everyday life. A key enabler for massive IoT communication is known as Orthogonal Frequency-Division Multiple Access (OFDMA). OFDMA divides the frequency bandwidth into multiple subchannels, known as resource units (RUs) and allows simultaneous transmission from large numbers of IoT devices. This technology supports many compelling applications [54]: sensors in smart warehouses and smart-city contexts collect and transmit massive amounts of aggregate data around the clock. For instance, networks of video cameras in cashier-less stores demand large amounts of uplink traffic in a more spatially-concentrated pattern: large retailers worldwide have recently introduced cashier-less stores that facilitate purchases via hundreds of cameras streaming video to a nearby edge server, inferring the items the customer has placed into their basket as well as tabulating each customer's total when they leave the store.

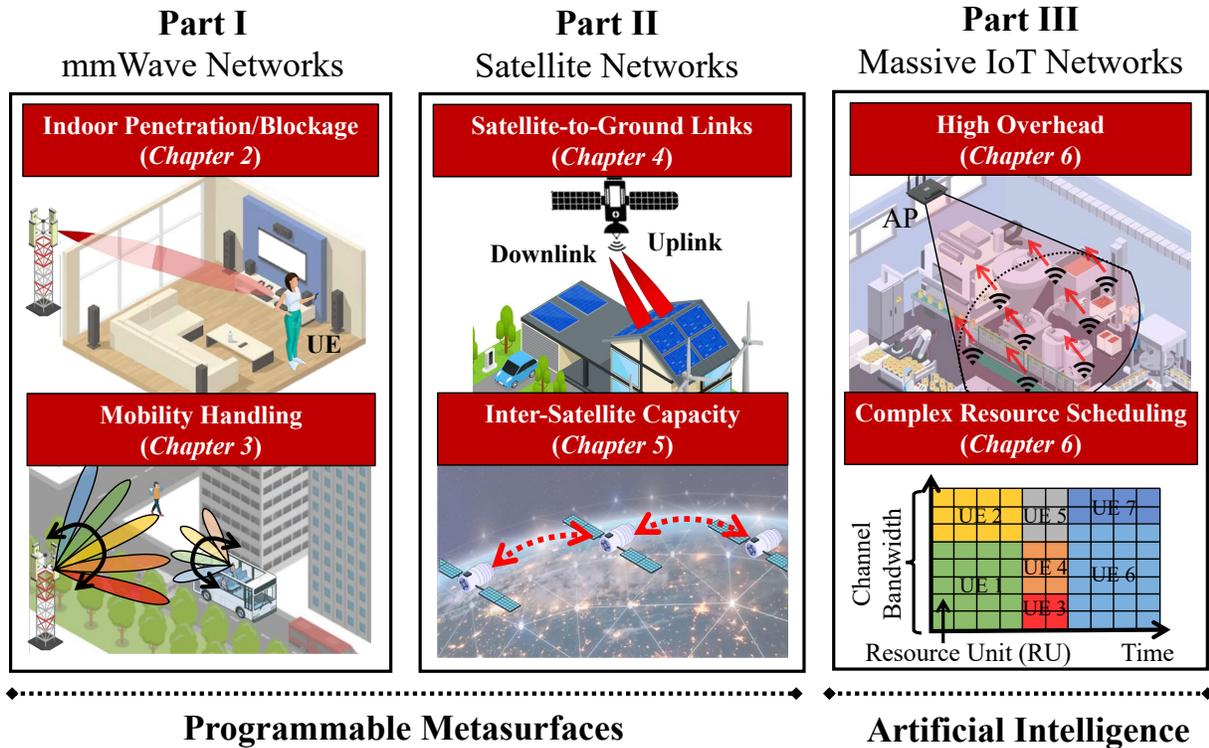


Figure 1.2: Unique challenges of mmWave, satellite, and massive IoT networks.

1.2 Challenges in Realizing NextG Services

While NextG networks are designed to provide diverse services with specific goals, their deployment in practice faces significant and unique challenges. This section delves into the challenges that each NextG network faces, as summarized in Fig. 1.2.

1.2.1 mmWave Networks

Blockage. One of the primary obstacles in mmWave communication is its susceptibility to blockage [92] and high path loss: due to its short wavelengths and limited diffraction capability, mmWave signals are easily obstructed by common obstacles such as buildings, foliage, and even the human body [137, 65, 169, 12]. mmWave technology faces significant headwinds in at least three blockage scenarios:

1. 5G outdoor coverage is difficult to bring indoors, as exterior building walls block mmWave signals, as do outdoor windows' tinted glass. Attenuation at 28 GHz is *ca.* 40 dB versus 4 dB through indoor glass [169], as outdoor metalized glass coatings attenuate by 25–50 dB per layer [146]. Currently, operators are forced to offload mmWave traffic onto lower frequencies or off their networks entirely (Wi-Fi) when users move indoors, incurring delay and application disruptions.
2. Indoors, people, furniture, and other clutter block mmWave, forcing data to flow over a much less reliable reflection path. Indeed, in an indoor measurement campaign at 28 GHz, MacCartney *et al.* observe a close-in best non-line of sight path loss exponent of *ca.* 3, with a normally-distributed additional loss with an 11 dB variance [93]. While the resulting temporary outages are common, highly demanding applications like VR/AR cannot tolerate these glitches.
3. Third, NextG cellular providers face challenges in adopting mmWave frequencies outdoors for primary service as well as wireless backhaul because mmWave signals are readily absorbed by foliage, and reflection off buildings is largely specular, constraining the angle of reflection to be equal to the angle of incidence. Measurements in New York City highlight this issue: 28 GHz data shows most links greater than 200 meters are in outage [12].

Mobility Handling. Due to high path loss and highly directional beams, mmWave networks necessitate a dense deployment of small cells to ensure adequate coverage, resulting in frequent handovers as users move between cells. Indeed, mmWave 5G New Radio (NR) experiences handovers (HOs) significantly more often than LTE or low band 5G NR [63, 102]. Moreover, the handover process in mmWave networks complicates the already-complex low-band handover [34, 159], including precise beam alignment between user equipment (UE) and base station (BS) [137], and requiring the UE to frequently measure the signal strength

of neighboring cells and report it to the serving BS, which then decides when and to which cell to hand over the UE.

High delays associated with mmWave handovers have significant impacts throughout the entire protocol stack, from the physical layer to the application layer, degrading quality of service (QoS) for end users, especially for latency-sensitive applications [63, 101]. To prepare a handover, the UE periodically suspends data communication with the serving BS to perform neighboring cell measurements. This process requires users to engage in an exhaustive search for the strongest signal from neighboring BSs, resulting in bufferbloat and slow growth of the congestion window [99]. Furthermore, the handover decision relies on these measurement reports, which may be outdated or lost due to the rapidly changing channel conditions and potential disconnections. This can trigger unnecessary handover ping-pongs and lead to transport protocol connection timeouts, impacting overall network performance.

1.2.2 Satellite Networks

Unreliable Satellite-to-Ground Links. LEO satellite networks use the Ku band (10.7–18 GHz), which has a longer wavelength (25–17 mm) than the higher frequency bands also in use, mitigating the impact of precipitation somewhat, yet also has a wavelength short enough to create narrow beams for highly directional communication to the ground. Since it has a short wavelength (25–17 mm) that experiences loss when traversing heavy walls, it requires a line-of-sight (LoS) or near-LoS (*i.e.*, traversing only through a low-loss material such as glass) path between the transmitter and receiver. Current systems are designed with a dish antenna that the user mounts outside the buildings, which communicates with the satellite in both the uplink and downlink directions. The dish antenna then communicates with the modem through a wire leading from the dish into the building to a modem, which then wirelessly communicates with the user. While this system reduces outage from transient blockage, this relay incurs delay and application disruptions.

Table 1.1: Major current satellite internet service providers and their primary frequency band allocations (GHz) [33].

	Starlink	OneWeb	TeleSat
Downlink	10.8–12.7	10.7–12.7	17.8–20.2
Uplink	14.0–14.5	12.8–14.5	27.5–30

Moreover, LEO satellite networks use different frequency sub-bands in the uplink (upper Ku band) and downlink (lower Ku band) directions, as Table 1.1 shows. This frequency division duplex (FDD) communication further complicates beam alignment between users and satellites, especially with both in motion and requiring frequent satellite handoffs.

Inter-satellite Capacity. Satellite systems are capacity-limited as they serve vast areas, typically with limited spectral resources and capacity per unit area. Moreover, the entire network shares inter-satellite links (ISLs), leaving ISL as a key network bottleneck. While a terrestrial cellular system increases capacity by simply adding more sectors or increasing the number of sites, satellite networks require more advanced antenna designs to boost the spectral efficiency. Traditionally, the multiple-input multiple-output (MIMO) system is used to boost the spectrum efficiency through independent parallel channels between multiple transmit and receive antennas. However, the capacity offered by spatial MIMO systems largely relies on multipath components in a rich scattering environment, and so the practical applicability of MIMO is challenging in *rank-deficient*, free-space satellite channels. Further issues surrounding the power consumption of such MIMO systems, including that of power-hungry radio frequency (RF) chains whose power requirements scale proportionally with the number of antennas, constrain the applicability in satellite networks even further.

1.2.3 Massive IoT Networks

High Communication Overhead. The increase in spectrum bandwidth has allowed a massive number of users to access the network simultaneously. To facilitate this, OFDMA divides the available spectrum into many narrowband channels, with each sub-channel as-

signed to different users. However, this system requires every user to report its channel measurements, hampering overall spectral efficiency. Measuring and propagating channel state information (CSI) to the access point (AP) scales with the product of the number of users, frequency bandwidth, antenna count, and frequency of measurement. Highly dynamic environments further exacerbate the problem, necessitating more frequent measurements due to rapidly changing channel conditions. With densely deployed IoT devices, the overhead of collecting CSI from all devices may thus deplete available radio resources [156].

Complex Resource Scheduling. Scheduling OFDMA resources is computationally challenging in their own right. Exhaustively searching all possible user and RU combinations is computationally infeasible, as the size of the search space increases exponentially with the number of users and RU granularity [124].

1.3 Thesis Contributions

To overcome these challenges, we introduce the concept of a programmable smart radio environment using two key technologies: programmable smart surfaces and artificial intelligence. At the physical layer, we have designed programmable smart surface systems and deployed them on buildings, introducing the concept of software-configurable radio environments. At the link/MAC layer, we have introduced a machine learning-assisted massive IoT system that predicts wireless channels and accordingly allocates networking resources. Real-world systems are often highly complex and cannot always be explained with conceptual models alone. In this thesis, we go beyond modeling systems by building and implementing our ideas into physical hardware systems, integrating them with existing network protocols, and verifying them through rigorous experimental evaluation as shown in Fig. 1.3.

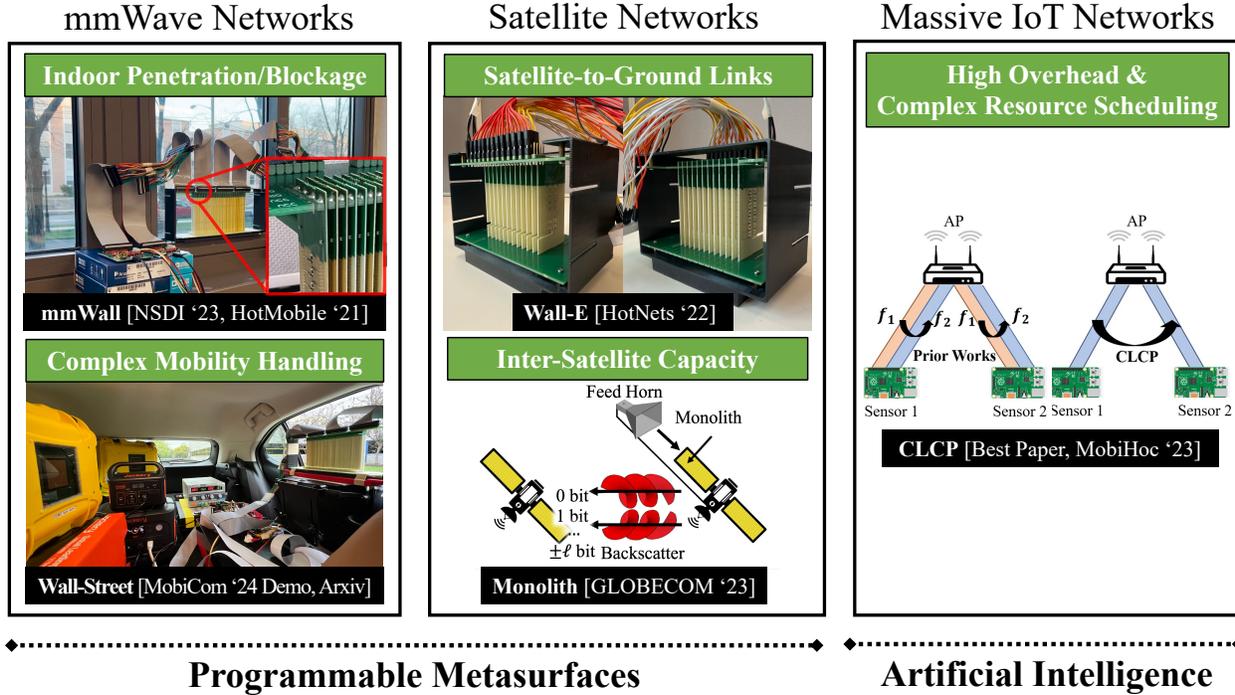


Figure 1.3: Our solutions introduce programmable smart radio environments through various surface prototypes, including mmWall (24 GHz), in-vehicle Wall-Street (26 GHz), Wall-E (10 and 15 GHz Ku-band), and Monolith, as well as an ML-based OFDMA scheduling system.

1.3.1 Programmable Metamaterial Surfaces

Prior wireless technologies have primarily focused on the endpoints themselves. However, these methods face fundamental limitations: they are constrained by the number of signal degrees of freedom available to endpoints, often rely on bulky or expensive RF chains, and are subject to environmental constraints that force endpoints to expend resources or increase in number.

Instead of relying on endpoints, we propose a fundamental shift: we can take full control of wireless signals by directly manipulating their electromagnetic properties in the environment. To achieve this, we built smart surfaces composed of thousands of meta-materials or meta-atoms, artificial composite materials engineered at a sub-wavelength scale. These devices fundamentally differ from conventional power-hungry wireless repeaters [7, 166, 74, 77], which merely receive signals, amplify, and re-transmit them. In contrast, each meta-atom directly

imparts a local phase shift as waves pass through it. Collectively, the surface can apply any fine-grained phase gradient required for reflecting, refracting, scattering, or combining the incoming beam. Our smart surfaces follow three design principles:

Principle 1: *Efficient Design:* We prioritize a compact and easily deployable form factor, coupled with low power consumption.

Principle 2: *Adaptive Functionality:* The surfaces can dynamically switch between multiple functions, rapidly time-multiplexing without requiring human intervention.

Principle 3: *Endpoint Optimization:* By providing active, sophisticated signal control at the surface level, we enable endpoints, such as base stations, routers, and hand-held devices, to become simpler, smaller, and more energy-efficient.

Below, I elaborate on our surfaces tailored to mmWave and satellite networks.

A Steerable, Transflective Metamaterial Surface for mmWave Networks

mmWall is the first electronically almost-360° steerable metamaterial surface that operates above 24 GHz and both refracts and reflects incoming mmWave transmissions. mmWall serves three key functionalities: (1) it brings and steers outdoor signals *through* the surface precisely to users indoors; (2) mmWall also reflects signals to bypass obstacles, ensuring reliable connections for outdoor users; (3) it conducts a rapid beam search by splitting and sweeping signals in multi-armed beams as a user moves. Integrated with the 5G New Radio link layer protocol, it autonomously switches functions without human intervention, programming the mmWave radio for the first time. Unlike conventional amplify-and-forward relays, mmWall operates without an RF chain or amplifier, drawing only a couple-of-hundred microwatts order of power. Our metamaterial design consists of 4,000 varactor-split ring resonator unit cells, miniaturized to under a millimeter in diameter for mmWave. Custom control circuitry drives each resonator, overcoming coupling challenges that arise at scale. We have fabricated a 10 cm by 20 cm mmWall prototype consisting of a 28 by 76 unit cell array and evaluated it in indoor, outdoor-to-indoor, and multi-beam scenarios. Indoors,

mmWall guarantees 91% of locations outage-free under 128-QAM mmWave data rates and boosts SNR by up to 15 dB. Outdoors, mmWall reduces the probability of complete link failure by a ratio of up to 40% under 0–80% path blockage and boosts SNR by up to 30 dB.

Smart Surface-Enabled 5G mmWave for Roadside Networking

Wall-Street is a smart surface installed on vehicles to enhance 5G mmWave connectivity for users inside. Wall-Street improves mobility management by (1) steering outdoor mmWave signals into the vehicle, ensuring coverage for all users; (2) enabling simultaneous serving cell data transfer and candidate handover cell measurement, allowing seamless handovers without service interruption; and (3) combining beams from old and new cells during a handover, realizing a make-before-break mechanism at mmWave for the first time. Through its flexible and diverse signal manipulation capabilities, Wall-Street provides uninterrupted high-speed connectivity for latency-sensitive applications in challenging mobile environments. We have implemented and integrated Wall-Street in the COSMOS testbed and evaluated its real-time performance with three gNBs and multiple mobile clients inside a surface-enabled vehicle, driving on a nearby road. In multi-UE scenarios, Wall-Street doubles the average TCP throughput and reduces delay by 30% over a baseline 5G Standalone handover protocol.

Dual-band Reconfigurable Metasurfaces for Satellite Networks

Wall-E explores how a dual-band, electronically tunable smart surface can enable dynamic beam alignment between the satellite and mobile users, making service possible in rural and remote areas. It is the first of its kind to target dual channels in the Ku radio frequency band with a novel dual Huygens resonator design that leverages radio reciprocity to allow our surface to simultaneously steer energy in the satellite uplink and downlink directions, and in both reflective and transmissive modes of operation. Our surface, Wall-E, is designed and evaluated in an electromagnetic simulator and demonstrates 94% transmission efficiency and a 85% reflection efficiency, with at most 6 dB power loss at steering angles over a 150 degree field of view for both transmission and reflection. With 75cm^2 surface, our link budget

calculations predict 4 dB and 24 dB improvement in the SNR of a link entering the window of a rural home in comparison to the free-space path and brick wall penetration, respectively.

OAM Metasurface for Rank-Deficient Satellite Networks

Monolith is a low-power, programmable metamaterial surface that enables MIMO multiplexing in rank deficient, free space wireless environments. Specifically, it leveraging ambient signals as the source of power and backscatters these signals by modulating them into several orthogonal beams, where each beam carries a unique Orbital Angular Momentum (OAM). Since Monolith itself generates multiple orthogonal beams, it eliminates the need for rich scattering environments for spatial multiplexing and thereby boosts the network capacity of inter-satellite links. Our results show that Monolith achieves an order of magnitude higher channel capacity than traditional spatial MIMO backscattering networks.

1.3.2 Machine Learning for Wireless Networks

This thesis also guides the way that machine learning algorithms will assist the physical and link/MAC layers of large-scale, ultra-dense wireless networks to perform the crucial tasks of radio resource management for massive IoT networks.

Cross-Link Channel Prediction (CLCP) in Massive IoT Networks

CLCP is a technique that leverages multi-view representation learning to predict the channel response of a large number of users, thereby reducing channel estimation overhead further than previously possible. A key insight is that modeling the radio environment in a shared wireless medium is feasible with sufficient background data. Inspired by how computer vision builds 3D models from photos taken at different angles, CLCP treats each channel observation from prior transmissions as a snapshot from a specific viewpoint, merging these views into a cohesive representation of the environment. This representation is then used to predict unobserved wireless links. CLCP’s design is highly practical, exploiting existing transmissions rather than dedicated channel sounding or extra pilot signals. Lastly, we propose an efficient OFDMA scheduling algorithm that jointly optimizes both channel conditions

and user data requirements. We evaluate CLCP in two large-scale indoor scenarios involving both line-of-sight and non-line-of-sight transmissions with up to 144 different 802.11ax users, demonstrating a 2x throughput gain over baseline and a 30% throughput gain over existing prediction algorithms.

1.4 Statement on Publications

As a lead author, I have designed, implemented, and evaluated all projects.

Chapter 2 revises a previous publication [27, 26]:

- **mmWall: A Steerable, Transflective Metamaterial Surface for NextG mmWave Networks**, Kun Woo Cho, Mohammad H. Mazaheri, Jeremy Gummeson, Omid Abari, and Kyle Jamieson, In 20th USENIX Symposium on Networked Systems Design and Implementation (USENIX NSDI '23)¹.
- **mmWall: A Reconfigurable Metamaterial Surface for mmWave Networks**, Kun Woo Cho, Mohammad H. Mazaheri, Jeremy Gummeson, Omid Abari, and Kyle Jamieson, In Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications (ACM HotMobile '21)¹.

Chapter 3 revises a previous publication [25]:

- **Wall-Street: Smart Surface-enabled 5G mmWave for Roadside Networking**, Kun Woo Cho, Prasanthi Maddala, Ivan Seskar, and Kyle Jamieson, arXiv preprint arXiv:2405.06754, 2024².
- **Demo: Metasurface-Enabled NextG mmWave for Roadside Networking**, Kun Woo Cho, Prasanthi Maddala, Ivan Seskar, and Kyle Jamieson, In Proceedings

¹Mohammad H. Mazaheri, Jeremy Gummeson, Omid Abari, and Kyle Jamieson supervised the project.

²Prasanthi Maddala and Ivan Seskar helped integrating Wall-Street into the COSMOS testbed and implementing new PHY features. Ivan Seskar and Kyle Jamieson supervised the project.

of the 30th Annual International Conference on Mobile Computing And Networking (ACM MobiCom '24).

Chapter 4 revises a previous publication [23]: **Towards Dual-band Reconfigurable Metasurfaces for Satellite Networking**, Kun Woo Cho, Yasaman Ghasempour, and Kyle Jamieson, In Proceedings of the 21st ACM Workshop on Hot Topics in Networks (ACM HotNets '22)³.

Chapter 5 revises a previous publication [24]: **A Low-power OAM Metasurface for Rank-deficient Wireless Environments**, Kun Woo Cho, Srikar Kasi, and Kyle Jamieson, In IEEE Global Communications Conference (IEEE GLOBECOM '23)⁴.

Chapter 6 revises a previous publication [22]: **Scalable Multi-Modal Learning for Cross-link Channel Prediction in Massive IoT Networks**, Kun Woo Cho, Marco Cominelli, Francesco Gringoli, Joerg Widmer, and Kyle Jamieson, In Proceedings of the 24th International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Computing (ACM MobiHoc '23)⁵ **Best Paper Award**.

1.5 Funding Acknowledgements

My dissertation research is based upon work supported by SEAS Project X Fund and the National Science Foundation under grant no. CNS-2148271 and is supported in part by funds from federal agency and industry partners as specified in the Resilient & Intelligent NextG Systems (RINGS) program.

³Yasaman Ghasempour and Kyle Jamieson supervised the project.

⁴Srikar Kasi simulated channel rank enhancement and capacity comparison between OAM and MIMO. Kyle Jamieson supervised the project.

⁵Marco Cominelli collected data using 802.11ax devices. Francesco Gringoli, Joerg Widmer, and Kyle Jamieson supervised the project.

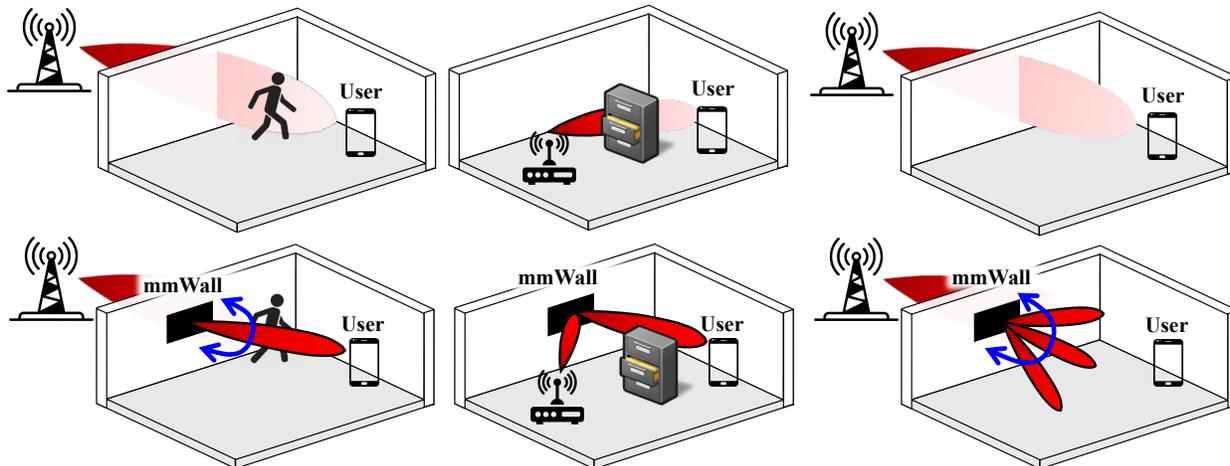
Part I

Pushing the Limits of mmWave Networks

Chapter 2

A Steerable, Transflective Metasurface for mmWave Networks.

This chapter describes the design and implementation of *mmWall*, an electronically reconfigurable surface that addresses three use cases shown in Figure 2.1. Like much prior work (§2.1), mmWall leverages *metamaterials*, artificial composite materials engineered at a sub-wavelength scale to exhibit unique electromagnetic properties that do not exist in naturally occurring materials [67]. But mmWall is the first practical work to our knowledge to use a specific class of metamaterials capable of refracting incoming radiation with (theoretically) no loss: *Huygens* metamaterials [42, 110]. mmWall is a reconfigurable intelligent surface that uses a novel Huygens metasurface (HMS) metamaterial to reflect/refract and precisely steer incoming mmWave beams towards desired directions, thus enhancing path diversity for mmWave networks. Work has shown that surfaces that can steer incoming mmWave transmissions in this way have the potential to dramatically improve spatial multiplexing [164] and spectral efficiency [158] of networks as a whole. Hence when obstacles like a human body or outdoor foliage blocks the line of sight (LoS) or non-line-of-sight (NLoS) paths, mmWall can often provide an alternative path that is not a simple reflection or a straight-line transmission, and hence would not otherwise exist. In the first scenario, mmWall can refract



(a) 5G/6G outdoor-to-indoor coverage via mmWall. (b) Reflective mode for indoor VR/AR. (c) Beam splitting, for link establishment.

Figure 2.1: mmWall re-focuses outdoor coverage indoors towards the user and potentially around obstacles, provides path diversity indoors by reflection, and splits an incoming beam for fast link establishment (*upper*: without mmWall; *lower*: with mmWall).

mmWave signals from outdoors to steer them directly towards an indoor receiver, making outdoor to indoor communication possible. In the second scenario above, mmWall reflects mmWave beams at non-specular angles (those for which the angle of reflection is not equal to the angle of incidence). And in the third scenario, mmWall can reflect outdoor transmissions at non-specular angles, ameliorating outdoor blockages.

mmWall is electronically reconfigurable to either reflect or refract incoming energy, allowing it to time-multiplex the different roles of each of the three above use cases without human intervention, while installed in a fixed location. Also, its multi-beam functionality (Figure 2.1(c)) enables fast beam search, and support for multiple users at the same time. mmWall has no RF chain, and its electric components draw only a couple-of-hundred microwatts order of power. Consequently, it consumes much lower power compared to a conventional AP that necessitates multiple RF chains for multi-beam operations. To our knowledge, mmWall is the first surface able to achieve near-360° angular coverage (Section 2.5).

This work addresses several hardware and software design challenges that arise in the realization of such a design. Since mmWave transmissions are “pencil-beam” in nature, they work only when the transmitter’s beam is perfectly aligned with the receiver’s beam. To correctly steer the beam towards the receiver, we design a metamaterials-based surface that can precisely control the phases of the incoming signal, focusing signal power in a narrow beam. Secondly, since the size of meta-atom scales with its operating frequency, mmWall’s meta-atoms are much smaller than the conventional antennas and therefore extremely sensitive to coupling. Hence, we not only scale the surface to mmWave frequency but also deliberately design the control lines to avoid undesirable coupling. Lastly, existing systems use their own beam search protocol to find the best alignment. To make mmWall compatible with different mmWave applications, we design an effective beam alignment protocol that leaves the existing systems unchanged [64].

Contributions and Results. mmWall is the first design that can arbitrarily reflect, refract, and split the mmWave beam in a nearly lossless manner. We analyze our meta-atom designs and compare them with simulation results, allowing our designs to scale to different frequencies for potential applications like Terahertz communication. To the best of our knowledge, this is the first study that theoretically analyzes and builds a working prototype of a reconfigurable Huygens metasurface at mmWave frequency. We have designed and implemented mmWall hardware with a novel control network in custom PCB, and in Section 2.5, evaluate its performance through experiments in environments matching the scenarios we outline above. Our empirical results show that when both the AP and the client are in the same room, we can provide an SNR of 25 dB or more for all locations in a $10 \times 8m$ room, using a single mmWall surface. This SNR is sufficient to support 128-QAM in 91% of locations. Moreover, the SNR improves to 30 and 35 dB when we place two surfaces, respectively, on different walls. Finally, we show the effectiveness of mmWall in bringing outdoor mmWave networks indoors. When the AP is 6 meters away from the building, mmWall improves the

Table 2.1: mmWall: comparison to existing surface works.

Related Works	Freq.	Steer-able	Refl.	Trans.	Multi-beams	Loss (dB)	Same pol.	Low power
Passive HMS [42, 160]	<10GHz	No	No	Yes	No	-7 to -2	Yes	Yes
Active HMS [21, 87, 149]	<10GHz	Yes	Yes	Yes	No	-9 to -1	Yes	Yes
Passive non-HMS [114, 152]	mmWave	No	Yes	No	No	-4.5 to -1.7	Yes	Yes
Active non-HMS [30, 136, 134, 145, 79]	mmWave	Yes	Yes	No	No	-10 to -1	Yes	Yes
C-resonators [168, 86, 161]	Sub-THz	No	No	Yes	No	-10.5 to -8	No	Yes
Repeaters [7]	mmWave	Yes	Yes	No	No	Amplifier	Yes	No
Multi-arm repeaters [74, 77]	mmWave	Yes	Yes	No	Yes	Amplifier	Yes	No
mmWall	mmWave	Yes	Yes	Yes	Yes	-1.31	Yes	Yes

SNR by up to 30 dB, providing an SNR of 20 dB or more in all locations in a room using a single surface placed on a window.

2.1 Related Works

Metamaterials are artificially created surfaces with unique electromagnetic properties not found in naturally occurring materials. These engineered materials can directly alter existing signals in the environment, including negative refraction index [122], engineering complex beam patterns [70], and rotating the polarization of wireless signals [70]. Recent research has focused on 2D metamaterials with planar structures, known as *metasurfaces*, with applications in EM cloaking [122], imaging [125], and motion sensing [32]. While these designs have shown promise in controlled experiments at 10 GHz and higher, they have not been integrated into end-to-end networked systems that optimize signal paths in real-time.

In particular, Huygens’ metasurfaces (HMSs) [42, 160, 21, 87, 149] have gained attention as a new paradigm for beam refraction, beamforming, and reflection. They comprise a layer of co-located orthogonal electric and magnetic dipoles facing each other across a dielectric substrate in a 3D structure. This arrangement introduces a discontinuity in the electromagnetic fields, enabling manipulation of all attributes of the incident field, including magnitude and phase, in both reflective and transmissive directions. As summarized in Ta-

ble 2.1, earlier works [42, 160] have introduced passive HMSs, but they lack the capability to reconfigure and to achieve both transmission and reflection. Prior work in actively-controlled HMSs [21, 87, 149] use varactors or PIN diodes to tune each element in a continuous or binary (*i.e.*, on-off) manner, respectively. While these designs have shown great promise at low frequencies, below *ca.* 10 GHz, they do not scale to higher mmWave frequencies in a straightforward way, due to a mismatch between the required meta-atom size and a varactor’s size, and the attenuation that commonly available substrate would induce on an incident mmWave signal. Scaling these designs also requires narrower trace widths that are hard to fabricate and prone to breaking during diode soldering. More importantly, they focus on steering one beam in a one-sided direction, rather than steering one or more beams in a reflective and/or transmissive direction. mmWall is the first mmWave work to do so.

Recent work in mmWave metasurfaces has primarily focused on reflective designs, each with its own approach and limitations. In contrast, mmWall stands out with its unique ability to operate in both reflective and transmissive modes, a capability not present in existing works. Passive non-HMS based mmWave metasurfaces have been proposed that reflect signals at angles different from the incident angle [114, 152]. While innovative, these designs are limited by their inability to tune to specific receiver locations, resulting in wasted incident energy. Moreover, they operate solely in reflection mode, unlike mmWall which can both reflect and refract signals.

Work in actively-controlled mmWave metasurfaces includes a solely reflective, PIN-diode based surface at 2.3 and 28 GHz [30], whose evaluation at 28 GHz states a gain of 19 dBi, but which stops short of further experimental evaluation of steerability or any further end-to-end evaluation at 28 GHz. Tang *et al.* describe similar PIN-diode, reflective surfaces at 27 and 33 GHz, model path losses in such scenarios, and experimentally evaluate [136]. Tan *et al.* consider a similar design at 60 GHz [134], but neither consider HMS-based designs such as mmWall’s, which can shift between reflective (on both sides of the surface) and transmissive modes instantly via electronic control. NTT DoCoMo has described reflective,

outdoor-to-indoor surfaces operating at 28 GHz in press releases ([a], [b], [c]). This work states top-line experimental results but does not disclose design details or details of their experimental evaluation. Some sub-THz designs are similar in concept yet differ in substrate, leveraging switched 65 nm CMOS process controlling ring resonators for a 0.3 THz design [145]. However, evaluation efforts in this group of prior work stop short of realistic end-to-end experiments. Kim *et al.* proposes NR-Surface [79], a varactor-controlled patch antenna reflective surface working at 24 GHz that integrates with 5G New Radio’s beam management, considering realistic end-to-end design. However, NR-surface does not offer a transmissive mode, limiting applicability. Also, existing reflective surfaces only reflect on one side, while mmWall can reflect on both sides, working for both indoor and outdoor non-specular reflections from a fixed location. More importantly, mmWall offers both reflective and transmissive modes, enhancing its applicability and flexibility in various environments. Additionally, our work encompasses a comprehensive control design, details a real hardware implementation, integrates with existing network protocols, and presents significant new evaluation results from realistic and diverse scenarios.

Single-layered non-HMS metasurfaces are typically limited to reflection, but surfaces incorporating C-shaped resonators [168, 86, 161] can manipulate high-frequency signals passing through them. The phase control in these resonators is determined by their geometrical configuration, specifically the gap width, while amplitude control is achieved through angular orientation. However, they lack reconfigurability as the geometric configuration and angular orientation of these resonators cannot be changed once fabricated. Also, they operate in a linear cross-polarization scheme where phase control is effective only for transmitted beams with polarization perpendicular to the incident beam. This necessitates rotating the polarization after passing through the surface to maintain consistent polarization between incoming and outgoing waves. mmWall, on the other hand, maintains the polarization of the incident beam. Lastly, they exhibit higher transmission loss due to their 2D structure. While C-shaped resonator metasurfaces have a maximum transmission amplitude ranging

from 0.3 to 0.4 (-10.46 to -8 dB), our mmWall prototype achieves a maximum transmission amplitude of 0.86 (-1.31 dB) in real-life, demonstrating significantly lower transmission loss.

Recent amplify-and-forward proposals for Wi-Fi [166] use a mesh topology but do not scale to mmWave frequencies. At mmWave [7], these proposals are limited to indoor reflection. Recent complementary approaches leverage multi-beam transmission [77, 74], sensing and leveraging ambient reflectors [150], and use Wi-Fi as a control plane to discover mmWave links [129, 84]. While they align with mmWall’s goals, such approaches cannot create paths whose reflection angles diverge from their incident angles, or refract through a surface. More importantly, they rely on one or more power-hungry RF chains and amplifiers, limiting their use in practice. On the other hand, mmWall operates without an RF chain or amplifier, drawing only a couple-of-hundred microwatts order of power.

2.2 Primer: Huygens Metamaterials

HMSs comprise a layer of co-located *magnetic* and *electric* meta-atom, etched onto the two respective sides of a dielectric substrate (Figure 2.2, *inset*). The magnetic meta-atom is an enclosed metallic ring with one split, while the electric meta-atom has two splits and a metal strip in the center (Figure 2.4(b)). As the incident wave passes through the magnetic meta-atom, the wave’s magnetic field H_i induces a rotating current (green arrows in Figure 2.4(b), *upper*) on the magnetic meta-atom that, in turn, creates a magnetic response (\vec{M}_s along the z-axis in Figure 2.2). Likewise, the electric meta-atom is excited by the wave’s electric field \vec{E}_i , resulting in two symmetric, oscillating current loops (green arrows in Figure 2.4(b), *lower*) that create an electric response (\vec{J}_s along the y-axis in Figure 2.2). These responses interact with the wave’s fields, causing an abrupt phase shift. By varying the applied voltage to a tunable component loaded on each meta-atom, the surface precisely controls the responses, thereby allowing any phase shift from 0° to 360° with near-unity transmission and/or reflection.

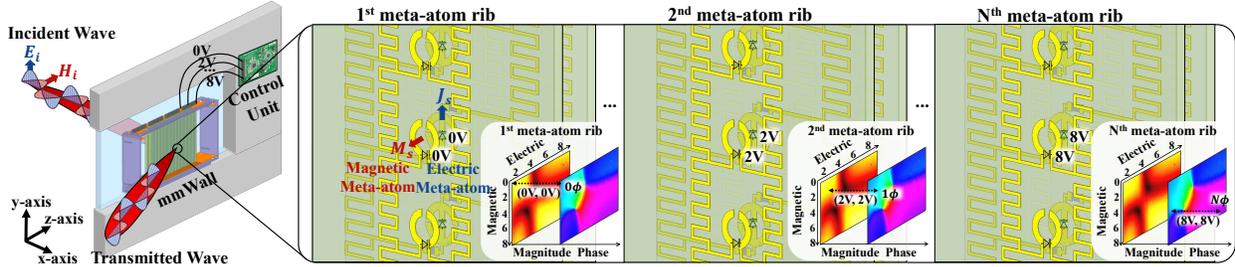


Figure 2.2: mmWall’s design converts an incident mmWave beam to a refracted (or reflected, not shown) beam via field discontinuities created by current in its resonators. *Inset*: magnetic meta-atoms are shown in front of the electric meta-atoms.

2.3 mmWall: Design and Analysis

We describe in turn mmWall’s hardware (§2.3.1), their control mechanism (§2.3.2), and their link layer integration (§2.3.3).

2.3.1 Surface Hardware

mmWall’s unit cells (also known as *meta-atoms*) are stacked vertically with a $\lambda/3$ separation, on each Rogers substrate board (also known as a *meta-atom rib*), as shown in Figure 2.2 (see Section 2.3.2 for a discussion of vertical and horizontal unit cell spacing considerations for beamforming). A control unit connected to mmWall provides a set of voltages to the ribs. In Figure 2.2, 0V is applied to both magnetic and electric meta-atoms on the first rib, causing the meta-atoms to shift the phase by 0ϕ with minimal loss. For the second rib, 2V is applied to shift the phase by 1ϕ . Ultimately, the beam is steered by all N ribs collectively forming an array factor.

Design Goals

The two primary design goals of the unit cell are to simultaneously **1**) achieve transmission T or reflection Γ loss levels as close to zero as possible, and **2**) effect any phase shift in $[0, 2\pi]$ on the incoming signal, both at mmWave frequencies.

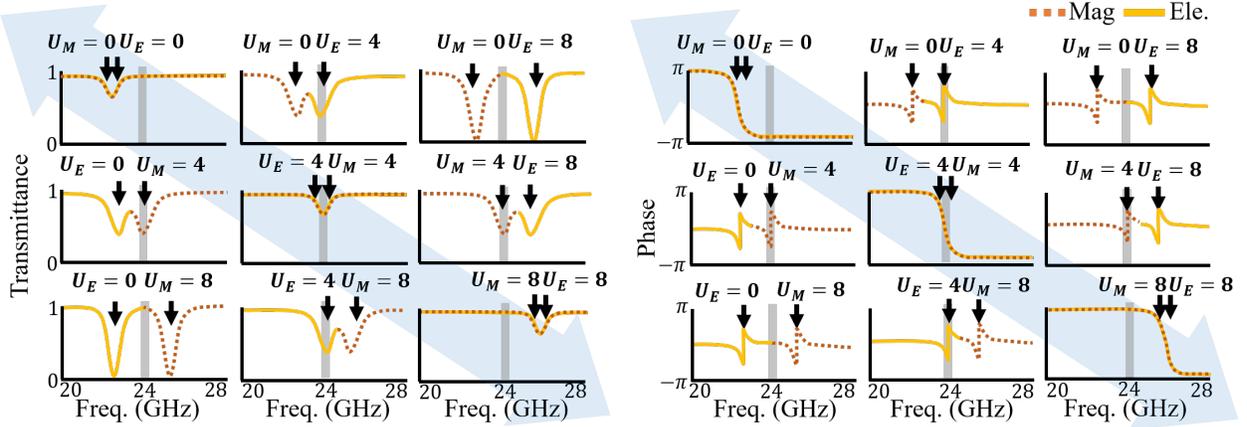
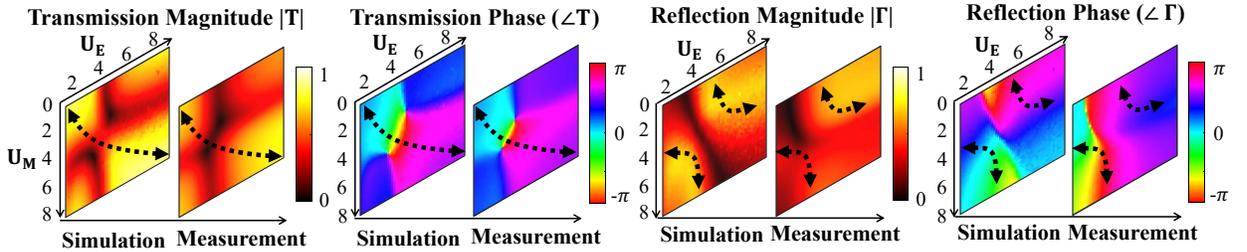
(a) Transmission magnitude $|T|$ (b) Transmission phase shift $\angle T$ (c) T and Γ magnitude, phase at 24 GHz.

Figure 2.3: Unit cell response *v.* electric- and magnetic-side control voltages U_E and U_M — (a): magnitude and (b): phase. (c): HFSS simulation (*left*) and near-field, real world VNA measurement (*right*)— arrows indicate control voltage pairs that yield a 360° phase shift of the incoming signal, with high transmission or reflection magnitude.

The unit cell consists of two meta-atoms, *magnetic* and *electric*. The magnetic (electric) meta-atom induces a magnetic (electric) field response to the incoming signal that can resonate at different, tunable frequencies by varying the applied voltage to the varactor of the magnetic- (electric-) meta-atom.

Without loss of generality, we now describe how transmission works (reflection is fully complementary to transmission, and we refer the reader to Section 2.8.1 for a rigorous mathematical exposition of both). In Figure 2.3(a), we observe that increasing the voltage applied to the magnetic meta-atom U_M from 0 to 8 V (down the three leftmost subplots) shifts its resonance frequency (lowest transmission magnitude point of the red dotted line¹) to the

¹In operation we largely avoid the lowest transmission nulls.

right (we will analyze how this frequency shifting works in Section 2.3.1). Similarly, the electric meta-atom induces an electric response and its resonant frequency can be shifted by its own varactor (reading similarly across the three topmost subplots). Together their effects are superposed and we manipulate the collective magneto-electric response that interferes with the incident plane wave.

The key characteristic that allows near-perfect amplitude with full phase coverage appears when the two responses overlap at the same frequency. Otherwise, the phase response undergoes a sharp change of only π and its magnitude dips to nearly zero at its resonant frequency, as we see in Figure 2.3(a) and Figure 2.3(b) when the voltages applied to the magnetic and electric meta-atom differ by 8 V. However, as the two resonances start to overlap, transmission loss decreases and the phase shift becomes 2π (on-diagonal sub-figures, Figure 2.3(a) and Figure 2.3(b)). As a result, we achieve 2π phase coverage with near-unity magnitude by increasing the voltage applied to both the magnetic and electric meta-atoms together (Figure 2.3(c), at control voltages indicated by the black curves).

While the overlapped resonances can reach a perfect unitary transmission magnitude in theory, the Huygens pattern from our measurement shows a lower transmission magnitude on the area where abrupt phase shifts occur due to various reasons, including the sensitivity at mmWave frequency, fabrication loss, and measurement errors.

Design Process

We now describe challenges we overcame in scaling the resonance of the mmWave unit cell to mmWave frequencies. By definition, the meta-atom behaves as an LC circuit with resonant frequency $1/(2\pi\sqrt{LC})$, determined by the capacitance or inductance of the meta-atoms. Hence, we must markedly *decrease* the inductance and capacitance of prior microwave designs (Section 2.1), if we can hope to achieve a mmWave resonant frequency. As we will see next, the smaller the ring is, the higher the resonant frequency becomes. However, the state-of-the-art approach to scale the frequency of a Huygens resonator (Figure 2.4(b)) requires a

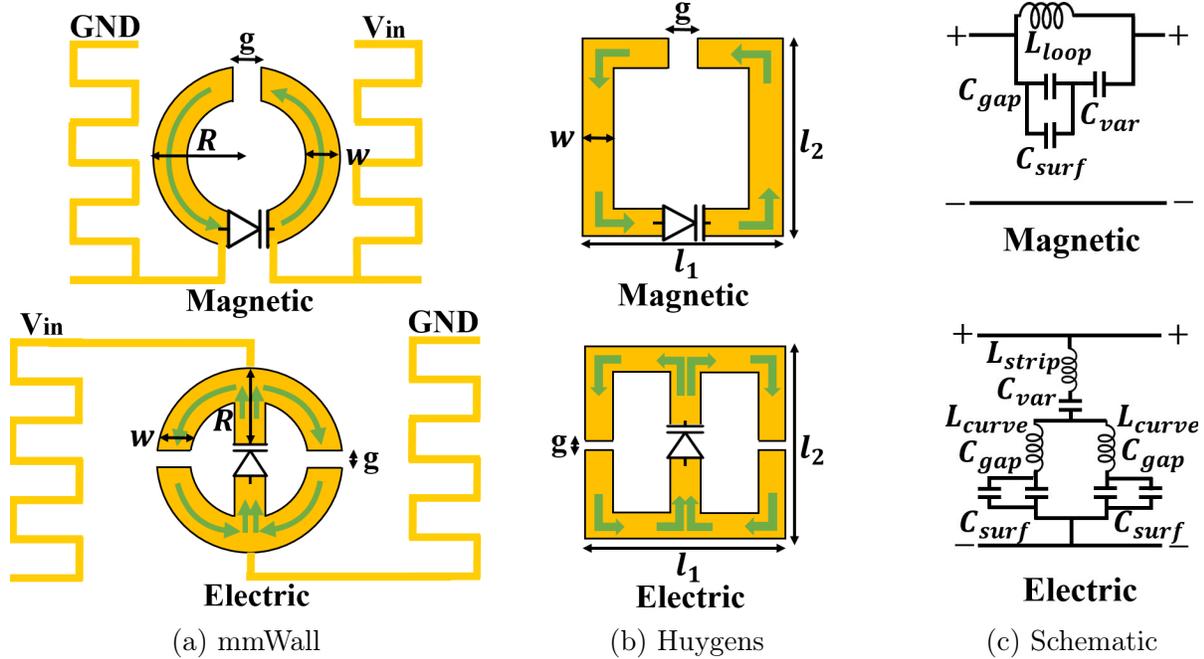


Figure 2.4: mmWall, prior Huygens unit cell designs (*top*: magnetic; *bottom*: electric side), and equivalent circuits. Green arrows indicate the oscillating current loops, and V_{in} indicates where the input voltages connect.

loop width l_1 and loop height l_2 of $\lambda/10$. At mmWave, however, the varactor packaging itself would significantly distort the tailored electromagnetic surface properties when a meta-atom is sized $\lambda/10$, and so the straightforward approach fails. Moreover, the varactor is soldered with heat, causing tighter designs to become more fragile. Changing the rectangular cell shape to a circular one with equal diameter reduces size while preserving varactor placement on the diameter.

We thus instead adopt the design shown in Figure 2.4(a), but this is only tenable with a careful tradeoff of meta-atom design parameters *radius* R , *trace width* w , and *trace gap width* g (*cf.* Figure 2.4) as we next describe.

Magnetic meta-atom. Figure 2.4(a) (*upper*) shows the design parameters that determine inductance L_m and capacitance C_m . L_m ($= L_{loop}$, the inductance of the physical conductor loop), is largely proportional to R (also $L_{loop} \propto t^{-1}$, w^{-1} , and g^{-1}). C_m consists of three

capacitance values, C_{gap} , C_{surf} , and C_{var} :

$$C_m = \left(\frac{1}{C_{\text{gap}} + C_{\text{surf}}} + \frac{1}{C_{\text{var}}} \right)^{-1} \quad (2.1)$$

Here, C_{gap} is the parallel-plate capacitance induced by the gap in the ring ($\propto g^{-1}$), C_{surf} is a capacitance induced by the metallic surface ($\propto R$ [140]), and C_{var} is the capacitance of the varactor, a voltage-dependent capacitor. While L_{loop} , C_{gap} , and C_{surf} are fixed after fabrication, C_{var} varies with control voltage. Increasing U_M decreases C_{var} (see Figure 2.19 in Section 2.8.1 for the precise relationship), and thus C_m (Equation (2.1)), which in turn increases the resonance frequency, as depicted in Figure 2.3.

When tuning the physical loop design parameters, we fix $C_{\text{var}} = 4$ V for both the magnetic and electric meta-atoms since at that voltage, the resonant frequency is at our desired mmWave frequency and an abrupt phase change occurs. Figure 2.5 shows our chosen design parameters (denoted with black circles) and its corresponding magnetic side resonant frequency when $U_M = 0, 10$ V. Calculated (curves) and simulated (markers) data in our sensitivity analysis show that among all feature dimensions, decreasing R , followed by increasing g has the greatest effect on increasing resonant frequency for the magnetic meta-atom. We note that after fixing our meta-atom geometry as shown in the figure, 24 GHz lies in the middle of the resulting resonant frequency range. Also, we observe that PCB manufacturing tolerance ($\pm 5\%$) does not greatly shift the resonant frequency (we refer Section 2.5.4 for meta-atom sensitivity analysis against fabrication tolerance).

Electric meta-atom. Figure 2.4(a) (*lower*) shows the electric meta-atom, in which current oscillates in two different directions, while the current of the magnetic meta-atom oscillates in one direction only (*cf.* green arrows in Figures 2.4(a) and 2.4(b)). Hence, we analyze its inductance L_e as the combination of the inductances of the half-circular loop on the left half (L_{curve}), the inductance of the other half on the right half (L_{curve} , by symmetry), and the inductance from the metallic strip shared by two loops (L_{strip}). Since the two half-loops

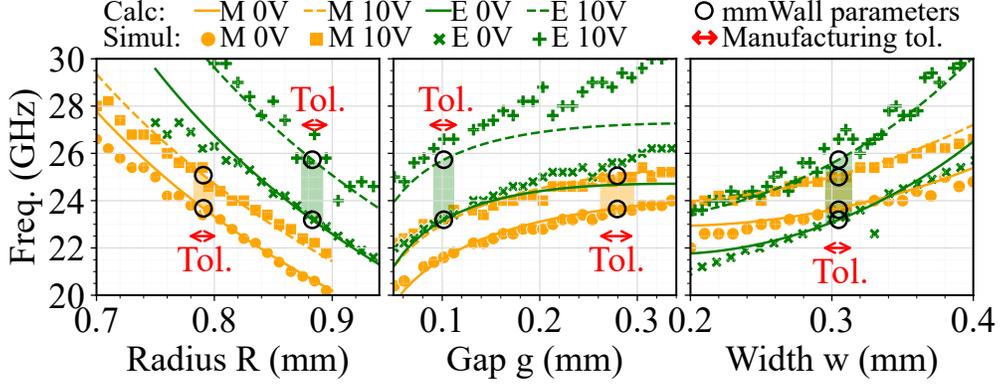


Figure 2.5: mmWall design parameter sensitivity analysis.

are arranged in parallel, with the metallic strip arranged in series, $L_e = (L_{curve}/2) + L_{strip}$ [44]. Since inductance generally depends on the surface area of the copper trace, $L_{curve} \propto R$, and $L_{strip} \propto w^{-1}$, L_e largely depends on both R and w , but not g . We see the impact of w on the resonant frequency in Figure 2.5: compared to magnetic meta-atom, the resonant frequency of the electric meta-atom increases steeply as w increases due to L_{strip} . To minimize the difference in resonant frequencies between the electric and magnetic sides as desired, Figure 2.5 guides us to design an electric meta-atom with equal w as the magnetic meta-atom, greater R and lesser g . The electric meta-atom has two gaps and two surface capacitances, with respective associated capacitances C_{gap} and C_{surf} , all in parallel, and that combination in series with C_{var} :

$$C_e = \left(\frac{1}{2(C_{gap} + C_{surf})} + \frac{1}{C_{var}} \right)^{-1} \quad (2.2)$$

Because there are many capacitances in parallel, changes in C_{var} lead to a wider frequency shift than analogous varactor tuning of the magnetic side. Using more precise equation-based analysis (available in Section 2.8.1) that matches our qualitative analysis, we cross-check and finalize design parameters R , g , and w for the magnetic and electric meta-atoms. We refer Section 2.8.1 for the values of the design parameters and voltage distributions for different steering angles.

In Figure 2.5, we observe that the difference in resonant frequencies between 0 and 10 V for the electric meta-atom are larger than the magnetic meta-atom. Hence, since the effect

Biasing lines—failed attempts (a)–(c):

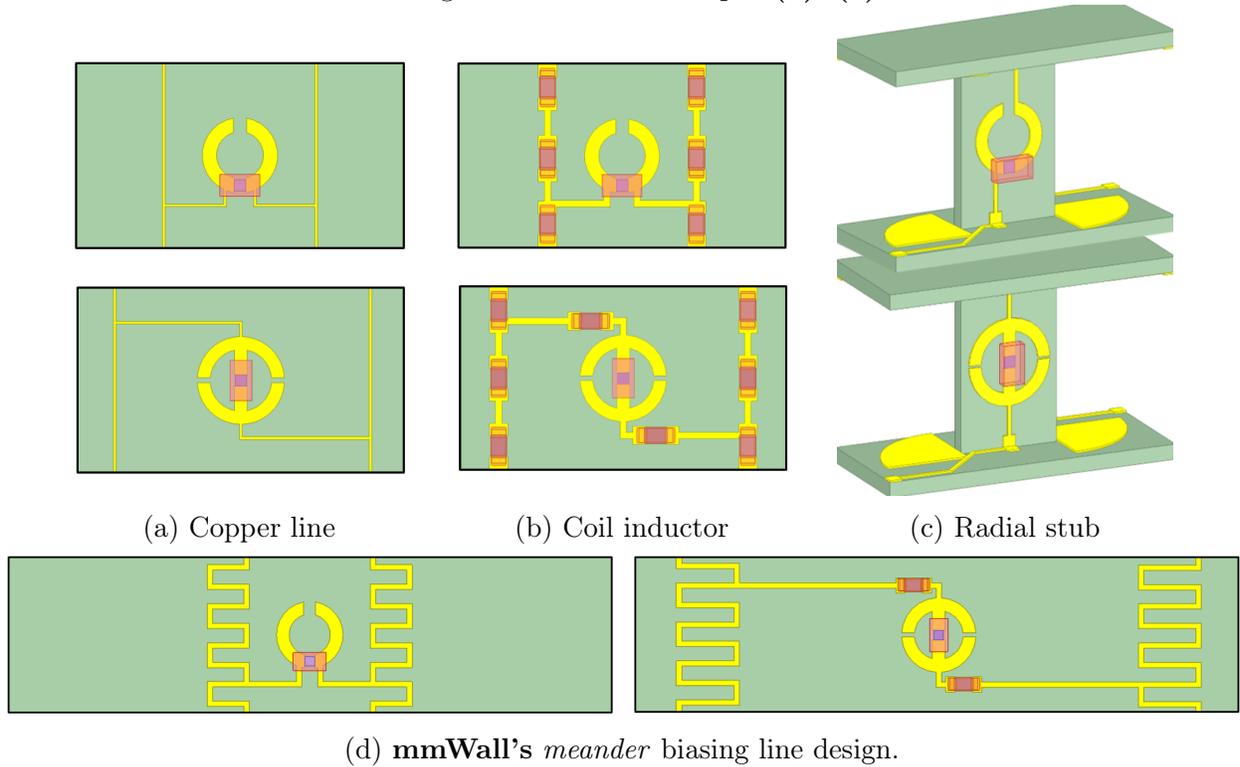


Figure 2.6: Biasing line designs: notable failed attempts include (a) straight microstrip, (b) coil inductor, and (c) radial stub. (d) mmWall uses an inner meander line for magnetic, and an outer meander line for electric meta-atoms.

of C_{var} differs, overlapping of resonance will not always occur when $U_M = U_E$. Rather than of simply finding the area where $U_M = U_E$ as suggested above, we instead need to search for the voltage pair for every desired phase that also maximizes the reflection or transmission magnitude. We do this by running one-time optimization that searches for the voltage pair that maximizes $|T|$ (or $|\Gamma|$) for each phase and generates a static lookup table that will later be used for beam steering.

2.3.2 Surface Control

To control the meta-atoms, we connect an off-surface control unit via ribbon cables with on-surface *biasing lines*, which altogether comprise the entire *control network* (Figure 2.2 on p. 51).

Biasing lines

This design process concerns the problem of designing the on-surface control network to interact with mmWave-frequency meta-atoms. Directly connecting a line to the meta-atoms changes the performance of the meta-atom, which causes mmWave signal loss and invalidates the design process described previously (§2.3.1). To mitigate such adverse effects, we seek to design biasing lines that incorporate radio frequency (RF) chokes, low pass filters that block RF signals within a certain frequency band from propagating on direct current (DC) signal paths. Our primary design goals are to design a biasing network that **1)** minimizes the use of extra components, **2)** avoids a large amount of copper on the panel where the meta-atom is placed, and **3)** is straightforward to fabricate. This is challenging because mmWave meta-atoms are sensitive to the shape and placement of the choke.

Failed attempts. Figure 2.6 shows various biasing line structures we have considered. First, we try a straight copper line design (**a**). We use a narrow width resembling a very high impedance transmission line, to try to attenuate the RF signal while the DC biasing voltage is applied. However, to achieve the desired impedance, a very narrow width transmission line (0.07 mm) is required which is not possible to fabricate by common PCB manufacturing techniques.

Second, we try the use of inductors to create a high-impedance line (**b**). The impedance of an inductor is determined by the RF frequency and is proportional to its inductance. However, inductance of mmWave inductor components are limited. Hence, we would need to apply at least four inductors in series to achieve the desired isolation, introducing significant surface complexity and also internal resistance that adversely affects unit cell efficiency.

Third, a radial stub which is an open ended transmission line is employed. The length of the stub determines the input impedance of the line, and so thus acts as an RF “choke” that blocks mmWave signals, while a DC biasing voltage is applied to the cell from the control network. The required length of the stub is one-quarter wavelength, which is comparable to the cell size. But if the stub is designed on the same panel, the stub itself would reflect most

of the wave, stealing energy to illuminate the cell itself. To avoid this problem, one can put the stubs on a perpendicular panel, as shown in Figure 2.6(c). This could potentially solve the wave reflection issue, but would complicate implementation, since there would be one perpendicular panel for each unit cell.

Proposed meander structure. To achieve our design goals, we have formulated a meander structure (Figure 2.6(d)) that acts as an RF choke, but at the same time connects the vertically adjacent meta-atoms. Longer and thinner traces provide more inductance, so by bending the straight wire vertically and horizontally, we enable the control network itself to be an inductor that outperforms the multiple off-the-shelf inductors. But this increases capacitance between the two meander lines on opposing sides of the unit cell, which also invalidates our meta-atom design process. So mmWall places the meander line of the magnetic meta-atom in a non-overlapping configuration relative to the meander line of the electric meta-atom. To compensate the loss from the microstrip that connects the electric meta-atom and the meander lines, we add two off-the-shelf inductors next to the electric meta-atom.

Beam steering and splitting

A conventional phased array transmitter has a net radiation pattern multiplying the radiation pattern of a single element by the *array factor* (AF), the pattern induced by the array. Unlike prior mmWave receive-transmit relay systems which require two phased antenna arrays (one to receive and another to transmit a new phase-shifted signal), mmWall uses only a single array of meta-atoms to directly shift the phase of an existing mmWave signal. For L omnidirectional antennas with d separation, each with transmit amplitude A , $AF = A \sum_{n=0}^{L-1} e^{2\pi jnd(\cos\theta)/\lambda}$ with radio wavelength λ and steering angle θ .

mmWall applies different phase shifts to each meta-atom rib for beam steering. Specifically, by searching over the space of control voltages to maximize reflection or transmission amplitude subject to achieving the desired phase (Figure 2.3(c)), we construct a look-up

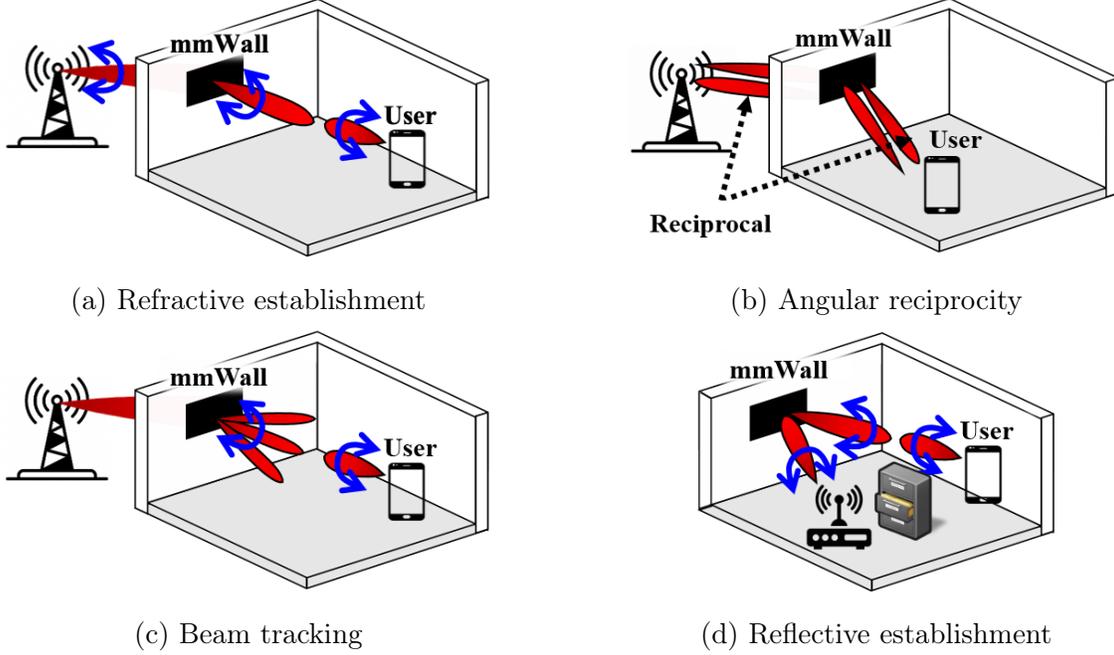


Figure 2.7: mmWall’s refractive link establishment, angular reciprocity property, tracking, and reflective link establishment.

table that maps steering phase φ to the chosen unit cell voltage pair (and without loss of generality) transmission coefficient: $\Phi(\varphi) \rightarrow \langle U_M, U_E, \Gamma \rangle$. The difference with conventional beamforming is that element amplitudes vary, so mmWall’s net radiation pattern becomes $\sum_{n=0}^{L-1} \Phi_{\Gamma}(\phi) e^{j\phi}$ where $\phi = 2\pi nd \cos \theta$.

To transform a single beam into multi-armed beams, we modify the above AF to account for angles θ_1 and θ_2 :

$$\sum_{n=0}^{N-1} (\alpha \Phi_{\Gamma}(\phi_1) e^{j\phi_1} + \beta \Phi_{\Gamma}(\phi_2) e^{j\phi_2}) \quad (2.3)$$

where $\phi_k = 2\pi nd \cos \theta_k$, and α and β are weighting terms that determine the power of each beam.

2.3.3 Link Layer Design

Recall that mmWall operates in two different modes, a lens mode and a mirror mode.²

1) In lens mode, a mmWave signal refracts through mmWall allowing, *e.g.*, a user inside

²Reflective mode and mirror mode are equivalent.

the building to communicate with the base station (*ENodeB*) in a cellular network. This requires two beam alignments: one between the *ENodeB* and mmWall, and another between mmWall and the user. **2)** In mirror mode, mmWall reflects mmWave signals. For example, in wireless LAN settings, it reflects the beam between the AP and user, which requires beam alignment between the AP and mmWall, and again between mmWall and the user.

mmWall electronically switches between the two modes because different users may be located outdoors and indoors. Hence, mmWall sweeps the beam in both lens and mirror mode to align to the user during a beam search.

Our development here follows the outline of the existing 5G New Radio (NR) beam management protocol, but adapts it to mmWall’s unique capabilities. The current 5G NR beam search proceeds in three steps: **1)** the *ENodeB* sweeps its beam, the user equipment (*UE*) selects a best direction, and reports it to the *ENodeB*; **2)** the *ENodeB* refines its beam (*i.e.*, sweeping a narrower beam over a narrower range), the user detects the best direction and reports it to the *ENodeB*; **3)** the *ENodeB* fixes a beam and the *UE* refines its receiver beam.

To establish a link from a cold start, the *ENodeB* sweeps different directions such that the user can detect the best beam for an initial link establishment (Figure 2.7(a)). If the *UE* cannot detect the beam or the beam strength is low, it turns mmWall to a lens mode and signal it to simultaneously sweep the beam received from the *ENodeB*, via sub-6 GHz control. At the same time, the *UE* scans its receiving beam to various directions. After the search, the *UE* knows the combination of the *ENodeB*’s transmit beam angle, mmWall’s beam refraction angle, and its receive beam angle that maximizes the SNR of downlink signals. Given an initial link, *ENodeB* and mmWall refine the beam by simultaneously sweeping narrower beams over narrower ranges, and lastly, the user refines its receiving beam.³ *ENodeB*-mmWall alignment takes $O(n)$ steps (for n directions), and mmWall-*UE* alignment takes $O(n^2)$ steps, so cold-start beam alignment as described above takes $O(n^3)$ steps, but only

³We note that some full-duplex relays [7] require the relay node’s receive direction aligned to the *ENodeB*, which is not necessary with mmWall.

once *ever* when mmWall is installed, because both ENodeB and mmWall are stationary. As long as mmWall remains in the same location, the one-time initial beam alignment is kept constant. Hence, the common case of cold-start beam establishment between mmWall and user in fact requires $O(n^2)$ steps (*cf.* Figure 2.7(c)). Also, the above notably does not require modifications to the existing 5G NR protocols.

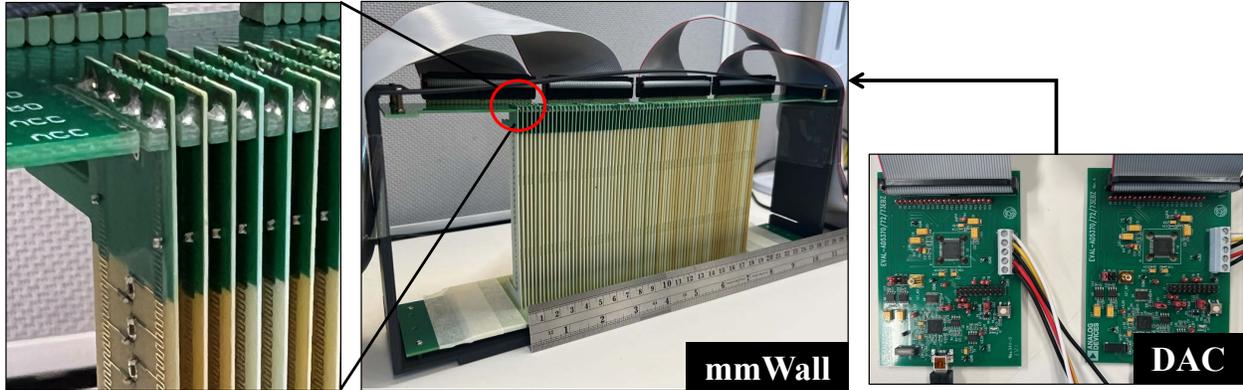
As illustrated in Figure 2.7(b) and demonstrated experimentally in Section 2.5.4, mmWall refracts beams in one direction at the same angle as they arrive at the surface from the other side of the surface (angular reciprocity), which obviates the need for separate downlink and uplink link establishment. If the downlink has already been established, mmWall does not reconfigure for the uplink. Instead, ENodeB simply switches the direction of its receiving beam to match its transmit beam, and the user transmits in the direction of its receiving beam. This facilitates a quick transition between downlink and uplink.

Since the UE controls mmWall, the user can alternate between the ‘lens’ mode for outdoor-to-indoor communication and the ‘mirror’ mode for indoor communication. For example, when the user switches from an outdoor to an indoor ENodeB, it signals mmWall to re-establish the beam estimation process for indoor usage.

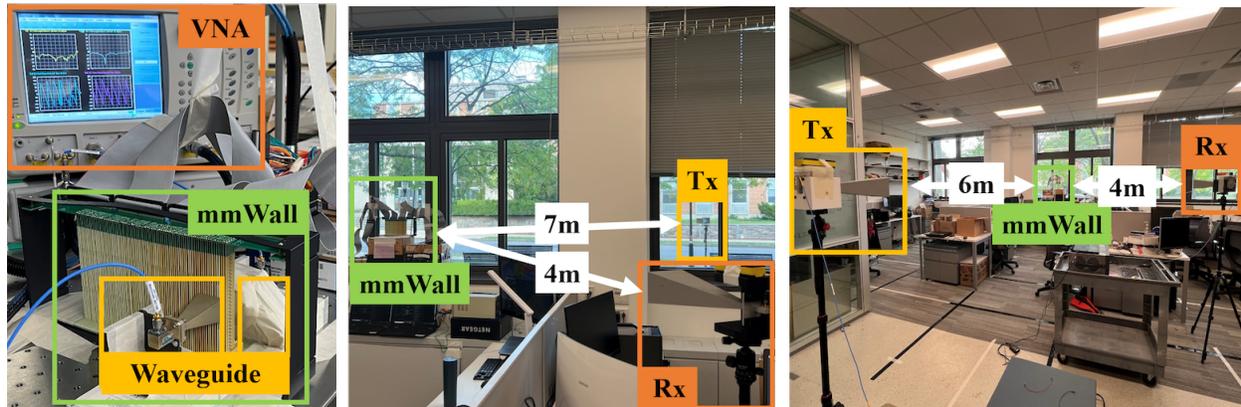
Multi-beam search. mmWall can create irregular beam shapes such as multi-arm beams (Section 2.5.3), which allows it to leverage state-of-the-art beam searching algorithms that exploit the sparsity of the mmWave channel to accelerate beam search [8, 116] by orders of magnitude improvement (essential for agile and mobile applications such as VR), now for the first time at a surface.

2.4 Implementation

We have fabricated and assembled a complete hardware prototype of mmWall, summarized in Figure 2.8. mmWall’s meta-atoms are fabricated on a 16 by 120 mm *rib* made of Rogers



(a) mmWall hardware



(b) Near-field testing

(c) Outdoor-to-indoor

(d) Indoor-to-indoor

Figure 2.8: mmWall’s hardware implementation, transmissive (‘lens’) and indoor reflective (‘mirror’) evaluation scenarios. We placed mmWall at the same location for both scenarios.

4003C printed circuit board (PCB) substrate, as shown in Figure 2.9. We assemble the PCB and constituent Macom MAVR-000120-1411 varactor diodes⁴ and 026011C-1N7 inductors.

In total, we have fabricated 76 ribs, each consisting of 28 vertical meta-atoms. These ribs are mechanically hold together with two perpendicular FR4 panels; one in top and the other in bottom of the structure. The top FR4 also provides control lines as it is shown in Figure 2.10. Each rib’s control pads are then soldered to the upper holder board, which connects the ribs to a DAC through its microstrip traces and pin headers. The lower holder boards are installed to position and the ribs fixed into these boards. For holding the ribs and FR4 panels steady, a 3D printed enclosure is fabricated that provides a standing support,

⁴We have modeled this varactor based on its *Simulation Program with Integrated Circuit Emphasis* (SPICE) model (see Appendix, Figure 2.19).

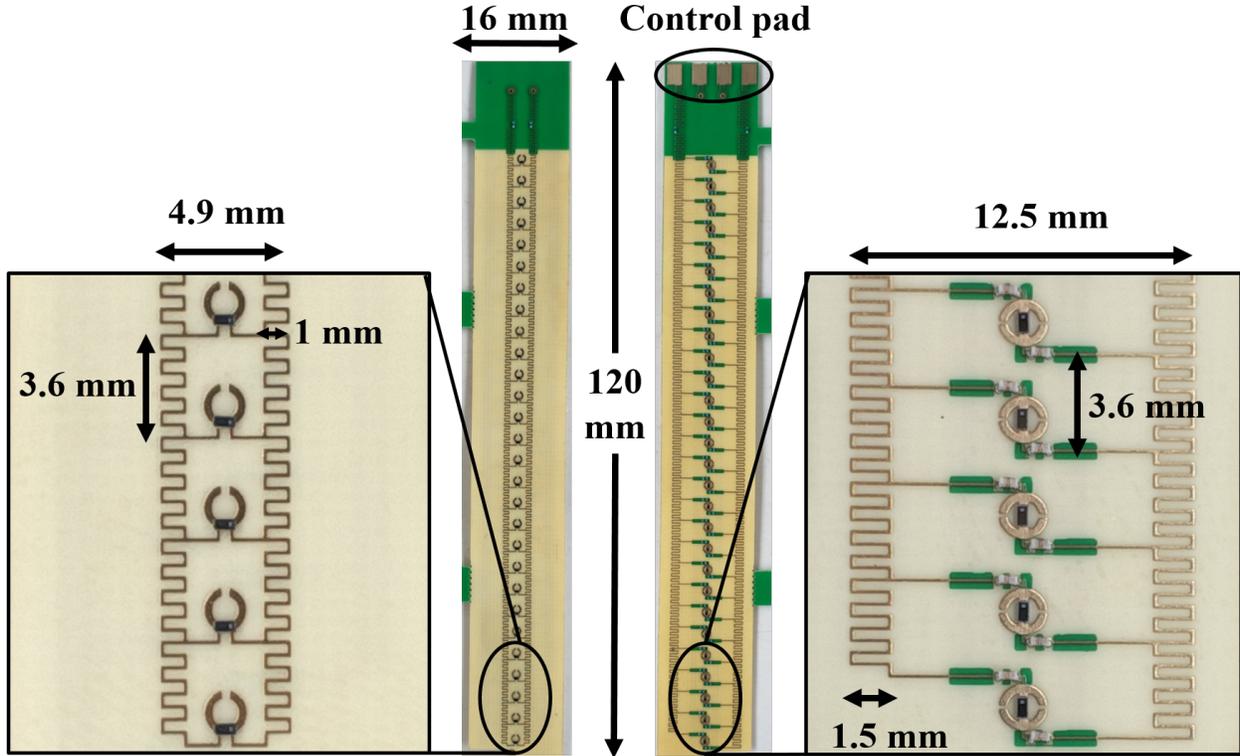


Figure 2.9: mmWall’s ribs, comprised of our proposed meta-atom design fabricated on a Rogers printed circuit board.

as shown in Figure 2.8(a). The spacing between the adjacent ribs are 2.6 mm, making the dimension of our mmWall prototype 120×197.6 mm. We note that scaling up our prototype with identical ribs and expanded FR4 holder boards is straightforward.

Four 40-channel AD5370 16-bit DACs from Analog Devices allow independent control of both electric and magnetic cells of every mmWall rib. Each DAC supplies a variable 0 to 10 V control voltage for each of 40 channels (*i.e.*, one DAC per 20 boards with one channel for U_E and U_M apiece). A laptop is connected to four DACs and listens for control signals from the UE. Once a signal is received, it sends a command to the DACs, which then apply the voltage levels, corresponding to a particular steering angle. Different voltage levels are found from a pre-stored voltage-to-phase look-up table. This control program is written in Microsoft Visual C++, and it can be executed from EnodeB, instead of UE. mmWall hardware, including the DACs, takes $20 \mu\text{s}$ to reconfigure the beam. The speed of DACs is

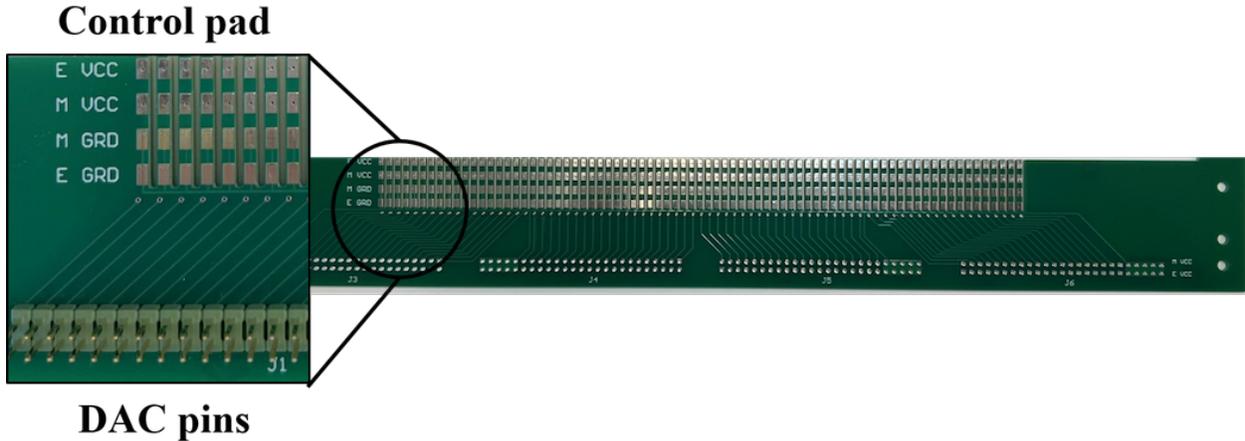


Figure 2.10: mmWall’s FR4 holder/control board.

the key determinant of the total latency, and deploying faster DAC hardware will lower the latency.

2.5 Evaluation

We begin with field studies that quantify mmWall’s SNR gain compared to the best NLoS environment path for both indoor-to-indoor and outdoor-to-indoor links (§3.4.2). Moreover, we explore the SNR gain and link failure rate under dynamic link conditions. We then evaluate multi-armed beams created by mmWall at various receiver locations (§2.5.3). We conclude with microbenchmarks to characterize mmWall’s steering performance, its support for wide steering angle, angular reciprocity, operation across wide bandwidths, and the impact of the surface size (§2.5.4).

2.5.1 Methodology

We conduct evaluations of various indoor and outdoor scenarios. For indoor-to-indoor settings, we place both the receiver (circles in Fig. 2.11) and transmitter (triangles in Fig. 2.11) in an office measuring 10×8 m, which includes interior walls, windows, and a server room. Between the three windows, there are two brick walls (black rectangles in Fig. 2.11). For

the outdoor-to-indoor testbed, we locate the transmitter outside the office, approximately 6 – 7 m away from the window, while the receiver is inside the office. During the experiments, we place mmWall in front of the window inside the room, and the loss of window is approximately -4 to -5 dB. For each outdoor-to-indoor and indoor-to-indoor experiment, we conduct two sets of experiments, each with a fixed transmitter location and 23 receiver locations. In the first set, the transmitter is perpendicularly facing mmWall and is 6.3 m away (upper subfigures of Fig. 2.11(b) and left two subfigures of Fig. 2.11(a)). The second set has the transmitter 6.8 m away from mmWall, and its beam hits the surface at approximately 30° to 40° angle (lower subfigures of Fig. 2.11(b) and right subfigures of Fig. 2.11(a)). During the beam search, mmWall steers the angle by the step of 0.5° . For end-to-end performance, we report SNR with a noise floor of 80 dBm.

Near-field experiments. Given that the measured Huygens pattern is likely to deviate from simulated results due to manufacturing tolerances, it is crucial to conduct accurate measurement through near-field experiments and compile a voltage-to-phase look-up table. Specifically, we collect near-field reflection and transmission coefficients of mmWall using two-port Anritsu MS4647B VNA, operating from 70 kHz to 70 GHz, as shown in Fig. 2.8(b). The Huygens pattern measured from the VNA is shown in Section 2.5.4. To minimize measurement error, we perform a two port calibration before acquiring the data. For data collection, we program the VNA using LabVIEW, which communicates with four DACs through the socket. During the measurement, mmWall is placed in between two waveguide horn antennas that are connected to the VNA. We place two horn antennas closely to mmWall to resemble the near-field simulation. Since the area of mmWall is larger than the aperture of waveguide horns, we collect the pattern on multiple locations of mmWall. In Section 2.5.4, we present a measured Huygens pattern at different locations of mmWall and demonstrate the robustness of mmWall against fabrication variations.

Far-field experiments. A standard mmWave base station is equipped with highly directional phased array antennas and supports an average EIRP range of 55-60 dBm [47, 46] or

more. With a 25 dBi transmit horn antenna, the maximum EIRP we achieved is 31 dBm, which is in accordance with FCC rules [48]. We use the same antenna at the receiver but apply a -10 dB correction to reflect typical UE antenna gain. Specifically, to generate mmWave signals, we use off-the-shelf phase-locked loop (PLL) frequency synthesizers ADF4371 with integrated VCO and frequency quadrupler, which quadruples 6.125 GHz VCO signals to 24.5 GHz. At the transmitter, the PLL output power is < -13 dBm, and we use the PLL in conjunction with a variable gain amplifier (VGA) HMC997LC4, which amplifies signals by 18 dBm.

2.5.2 In-situ Performance

In this section, we evaluate the end-to-end performance of mmWall for indoor and outdoor scenarios.

SNR improvement over the best environment path. To evaluate the effectiveness of mmWall in improving SNR in scenarios with blocked LoS paths, we conduct SNR measurements at multiple transmitter and receiver locations (two locations for the transmitter and 23 locations for the receiver). For each link, the transmitter and receiver (and mmWall if deployed) search for an NLoS path that maximizes the SNR.

Fig. 2.11 illustrates the measurements taken prior to and following the deployment of mmWall. Specifically, Fig. 2.11(b) presents the SNRs obtained when the transmitter was positioned towards the window at 0° (upper subfigure) and 30° (lower subfigure) in an indoor testbed. The results of both subfigures show that our indoor testbed has a rich scattering environment, with some receiver locations achieving SNR levels exceeding 25 dB in the absence of mmWall. However, receivers located at either end of the room experience SNR levels below 20 dB. With mmWall, all receivers, including the ones in the corner, achieve SNRs of at least 24 dB. Also, the nodes located within mmWall's steering angle of -45° to 45° has SNRs greater than 30 dB. This improvement in SNR is particularly evident in Fig. 2.12. In Fig. 2.12 (*left*), we plot a CDF of the best environment SNRs (black curves) alongside the

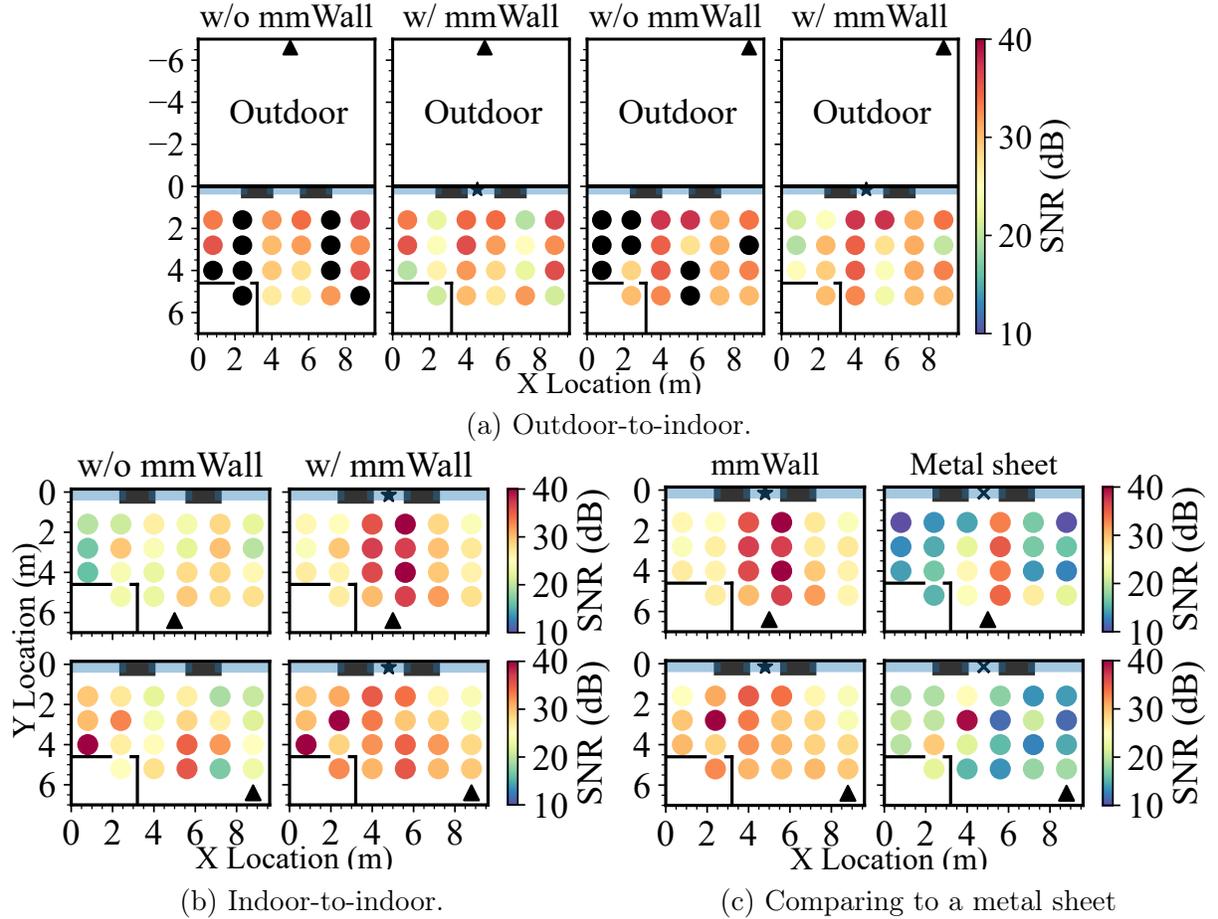


Figure 2.11: mmWall’s SNR improvement over the best NLoS environment path in two scenarios: (a) when both transmitter and receiver are located indoors (*upper*: transmitter facing mmWall perpendicularly; *lower*: transmitter facing 30° away from mmWall) and (b) when the transmitter is located outdoors (*left*: transmitting perpendicularly; *right*: transmitting at a 30° off-angle). We use the following notations: mmWall \star , transmitter \blacktriangle , receiver \circ . \bullet indicates no signal.

SNRs of mmWall links at the corresponding receiver location (rectangles). Fig. 2.12 (*center*) shows the CDFs of maximum SNRs between the environment and mmWall links, while Fig. 2.12 (*right*) shows the CDFs of the SNR gains over the environment path per receiver location. As shown in Fig. 2.12 (*upper*), mmWall ensures outage-free communication for 91% of receiver locations at 128 QAM [105] mmWave data rates, while only 40-50% of receivers achieve the same rate in the absence of mmWall. Moreover, among 80% of receivers that experience the gain from mmWall, some receive more than a 15 dB SNR boost.

In Fig. 2.11(a), we present the SNR improvement in outdoor-to-indoor scenarios. Without mmWall, receivers unable to establish an NLoS link through the window experience complete link failure. With mmWall, on the other hand, all receivers achieve SNRs of at least 19 – 20 dB. Fig. 2.12 (*lower*) shows the CDFs of outdoor-to-indoor SNR improvement. A single mmWall guarantees 64-QAM for almost all receiver locations and a 30 dB SNR boost for 40% of the links. Our results demonstrate that mmWall is highly beneficial for improving mmWave signals quality in the cases of wall blockage.

Deploying multiple mmWalls. To evaluate more than one mmWall, we place another mmWall (downward triangles in Fig. 2.12) in front of the window on the right side of the room. Fig. 2.12 (*upper*) demonstrates the SNR gain from deploying two mmWalls for indoor-to-indoor links. Compared to the gain from a single mmWall, an additional mmWall provides ≤ 5 dB SNR gain for some links. As shown in Fig. 2.12 (*lower*), there is almost no gain from adding an extra mmWall for outdoor-to-indoor links. The results indicate that a single mmWall is sufficient to provide good coverage (at least 128-QAM for reflective and 64-QAM for transmissive links) in a 10×8 m office room. In a static environment another mmWall will not help *if* a mmWall path is already available.

Improving reliability for dynamic links. While a single mmWall delivers good SNRs throughout all receiver locations, it is still possible for blockages to occur on mmWall links. Likewise, even if there is a robust NLoS path present, it can still be blocked. At mmWave frequencies, the indoor environment typically provides three to four strong paths, including

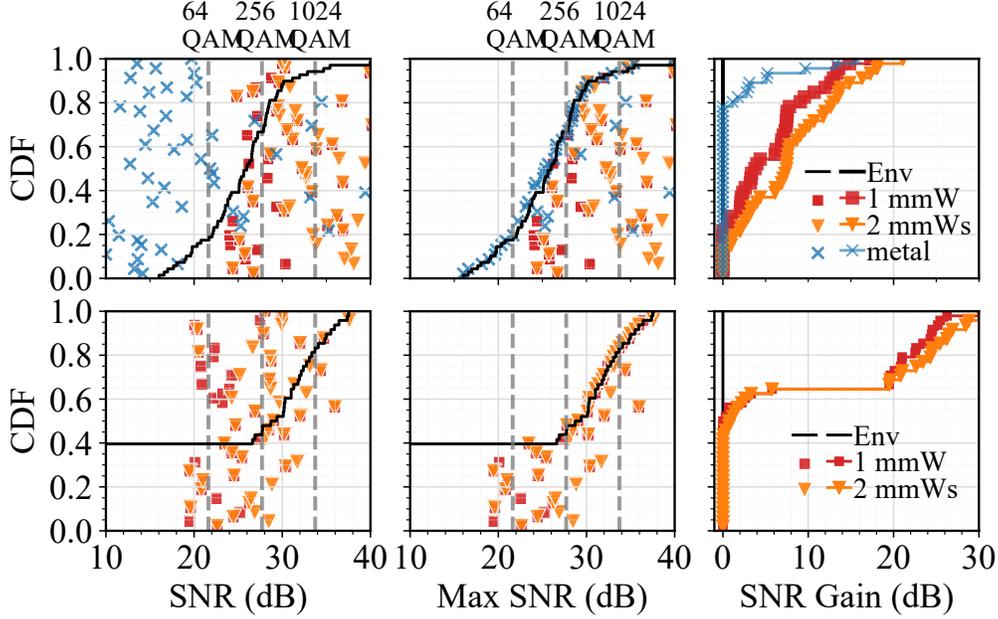


Figure 2.12: The SNR improvement from the use of one or more mmWalls at various receiver locations in indoor-to-indoor (*upper*) and outdoor-to-indoor (*lower*) scenarios. SNRs collected from a given receiver location are plotted on the same y-axis value (*left*: CDFs of the best environment SNRs in black curves alongside the SNRs of mmWall links at the corresponding receiver location in rectangles. The maximum SNRs between two mmWalls placed in different locations are denoted with downward triangles; *center*: the best available SNRs with or without one or more mmWalls; *right*: the SNR gains attained with one or more mmWalls compared to the best environment path in various Rx locations).

the LoS path [103]. Due to the limited number of available paths, an increase in blockages can easily result in link failure, which exacerbates when these obstructions are in motion. One of the primary benefits of using one or multiple mmWalls is the enhancement of link reliability. By providing a diverse, strong alternative path, mmWall reduces the probability of link scarcity. In Fig. 2.13, we demonstrate the SNR gain across various Rx locations as a function of the *blockage probability* β^5 for both environment and mmWall links. In indoor-to-indoor scenarios, a single mmWall and two mmWalls reduce the probability of link failure by a ratio of up to 10% and 20% under 80% path blockage, respectively. For the outdoor testbed, the probability of link failure decreases by 40% for a single mmWall and 45% for

⁵A blockage probability is equivalent to a probability of complete link failure for each path. Under various available paths, the blockage probability of one path is independent from the other.

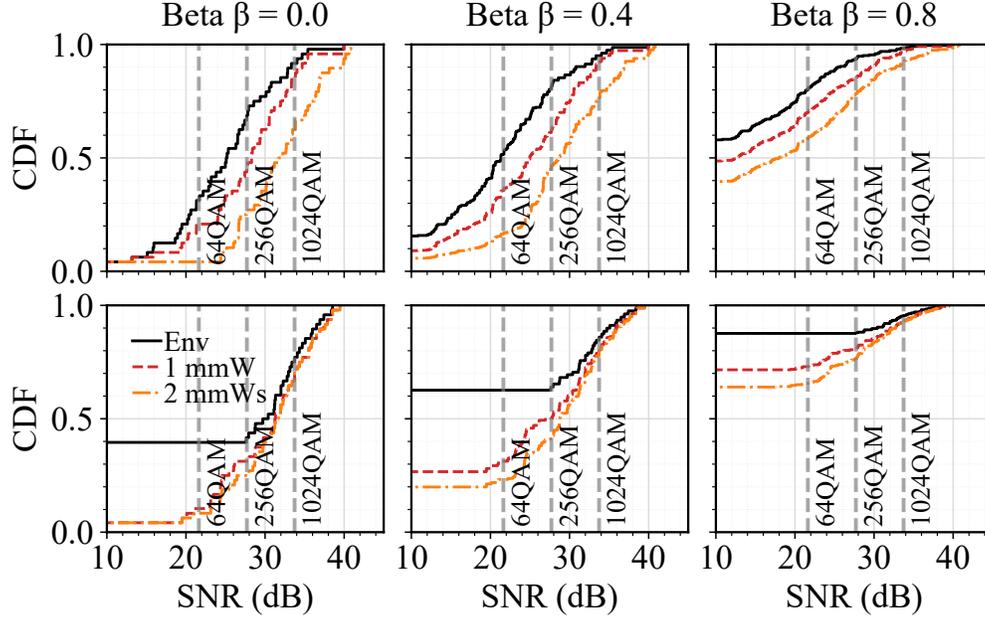
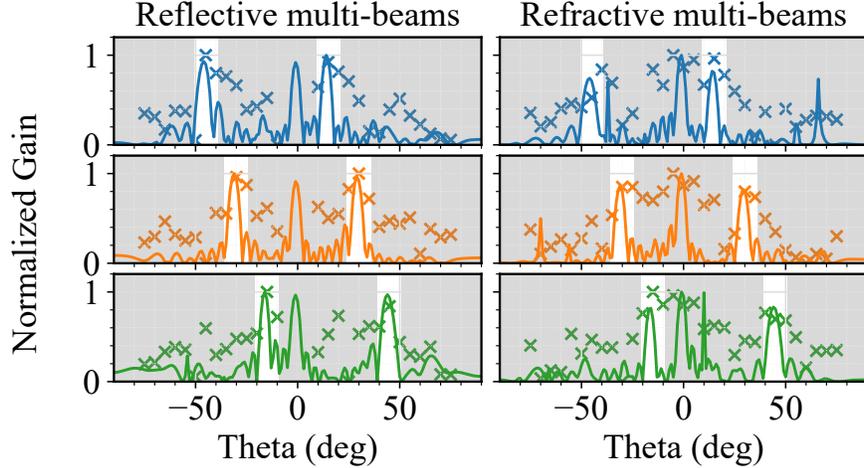


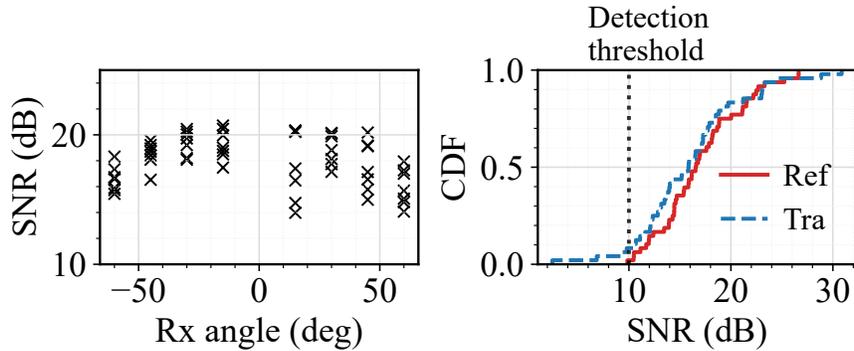
Figure 2.13: The SNR improvement (multiple mmWalls) for dynamic links (*upper*: indoor-to-indoor; *lower*: outdoor-to-indoor scenarios). β is a blockage probability.

two mmWalls under 40% blockage probability. Hence, we conclude that multiple mmWalls are beneficial when channel environments are highly dynamic.

One may argue that deploying a simple reflective metal sheet could help, but mmWall's ability to steer the beam has a significant impact on the extent of coverage. We evaluate the SNRs of links reflected by a 60×60 cm metal sheet, along with the SNRs of links steered by a 10×20 cm mmWall. As shown in Fig. 2.11(c) (*right*), only 10% of the receivers achieves SNRs above 30 dB, and the remaining 90% have SNRs below 15 dB. It is also worth noting that for a metal sheet, the SNRs depend largely on the location of the receiver and transmitter. In Fig. 2.11(c) (*right*), only the receivers that are placed and perfectly aligned with the angle of specular reflection achieve a high SNR. In Fig. 2.12, only 8% of the receivers achieve more than 5 dB SNR gain from the metal sheet. On the other hand, mmWall guarantees at least 25 dB SNRs across all areas. We conclude that, compared to fixed-angle reflection, mmWall links are less sensitive to the location of the transmitter, receiver, and surface, making them much more robust.



(a) mmWall’s multi-armed beam pattern (*upper*: $-45/15^\circ$ split; *middle*: $-30/30^\circ$ split; *lower*: $-15/-45^\circ$ split). Empirical points are denoted \times , with simulation curves.



(b) The SNRs of aligned multi-beams (*left*: a fixed distance between the transmitter and mmWall and between the receiver and mmWall; *right*: various Tx and Rx locations in the office setting.)

Figure 2.14: Evaluation of mmWall’s multi-armed beams.

2.5.3 Multi-armed Beams

We next evaluate mmWall’s capability to generate multi-armed beams. Fig. 2.14(a) presents our measurements on the multi-armed beams, along with simulation results from HFSS. Here, mmWall splits an incident beam into two beams at $-45^\circ/15^\circ$ and steers these multi-beams to $-30^\circ/30^\circ$ and $-15^\circ/45^\circ$. To measure the beam pattern, we position the transmitter and receiver three meters away from mmWall and record the gain of mmWall as we move the receiver from a -90° to 90° angle with respect to mmWall. Since we did not measure the beam pattern in an anechoic chamber, the received beam interfered with signals reflected off the indoor environment. Despite the interference, we observe that the gain peaks at

the angles where mmWall splits the beam. Furthermore, as mmWall steers its multi-armed beams, the measured peaks change accordingly. Our results show a peak at 0° due to leakage that was directly fed from the transmitter to the receiver. Reducing the distance between the transmitter and receiver and/or increasing the size of mmWall will reduce the peak at 0° .

We then measure SNRs as mmWall generates and steers various multi-beams, the beams that are 15° to 120° apart from each other. The distance between the transmitter and mmWall and between the receiver and mmWall are fixed to 2 m. Fig. 2.14(b) (*left*) reveals that as the beam is split into a wider angle, SNR drops.

To demonstrate the feasibility of a beam search using multi-armed beams, mmWall again splits the beam into two beams that are 15° to 120° apart from each other. Then it aligns the beam with the receivers at 23 different locations in the room. Fig. 2.14(b) (*right*) reports the SNRs of mmWall’s multi-beam links aligned with various receivers. The results show that more than 90% of multi-beam links achieve SNRs above 10 dB. Considering that no signal is detected in many locations under outdoor-to-indoor settings, 10 dB SNR is enough for the receiver to detect the beam and start the alignment. We conclude that mmWall can generate multi-armed beams that are sufficiently strong to accelerate beam search.

2.5.4 Microbenchmarks

We now evaluate mmWall’s steering performance, its support for wide steering angle, angular reciprocity, operation across wide bandwidths, and the impact of the surface size. The microbenchmark testbed consists of the receiver and transmitter modules that are three meters away from mmWall. Fig. 2.15 presents both the actual experimental measurements (markers) and simulated results (curves) acquired from HFSS.

Since mmWall does not have an amplifier, the effective aperture A_e is the primary factor that determines its gain. A well-defined relation for the effective aperture in terms of the aperture gain G is $A_e 4\pi/\lambda^2$. We define the aperture gain as our capacity and compare it

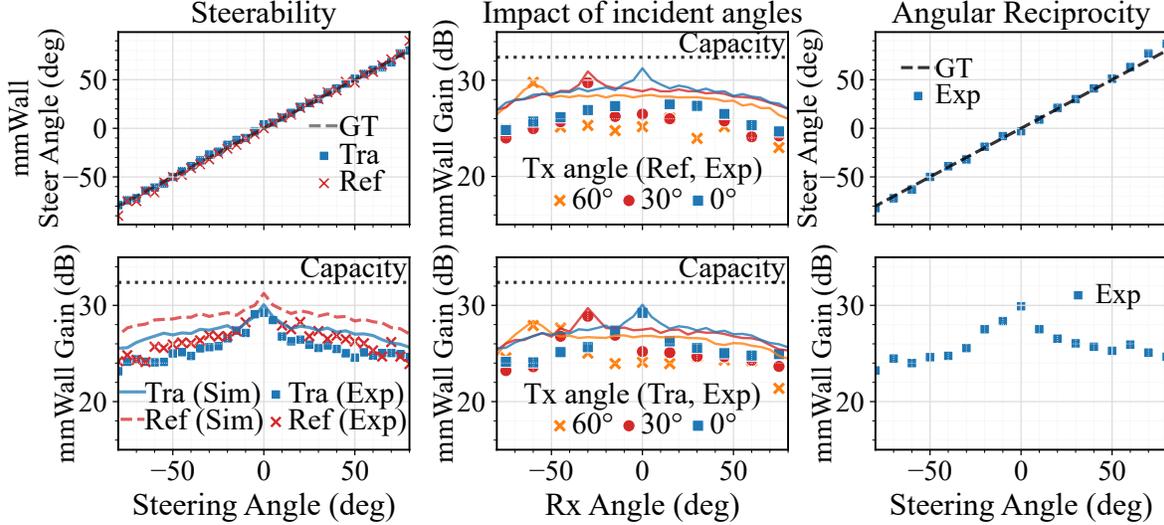


Figure 2.15: Microbenchmarks evaluating (*left to right*;) surface steerability, performance sensitivity of the incident wave angle, and angular reciprocity. Empirical points are denoted with markers, with simulation curves.

against our measured mmWall gains in our microbenchmarks. A rigorous analysis on mmWall gain is available in Section 2.8.2.

mmWall controllability. Fig. 2.15 (*upper first*) presents mmWall’s beam alignment accuracy. We place the receiver at 37 locations in our testbed and find the angle that provides the maximum SNR as mmWall sweeps the beam from -80° to 80° angle. During the experiment, the transmitter is facing mmWall at 0° angle. For both reflection and transmission, mmWall accurately steers the beam with at most 3° difference from the groundtruth (GT). Second, we evaluate the effect of a steering angle on the mmWall gain in Fig. 2.15 (*lower first*). As mmWall increases the steering angle, the gain slowly decreases. Furthermore, reflection provides a slightly higher gain than transmission.

Support for wide steering angle. In this microbenchmark, we evaluate the effect of incident beam angles on the mmWall gain jointly with the steering angle. Here, we move the transmitter to three different locations and the receiver to 37 locations. Fig. 2.15 (*upper second*) and Fig. 2.15 (*lower second*) show the impact of incident angles for reflection and transmission, respectively. For both scenarios, increasing the incident beam angle does not

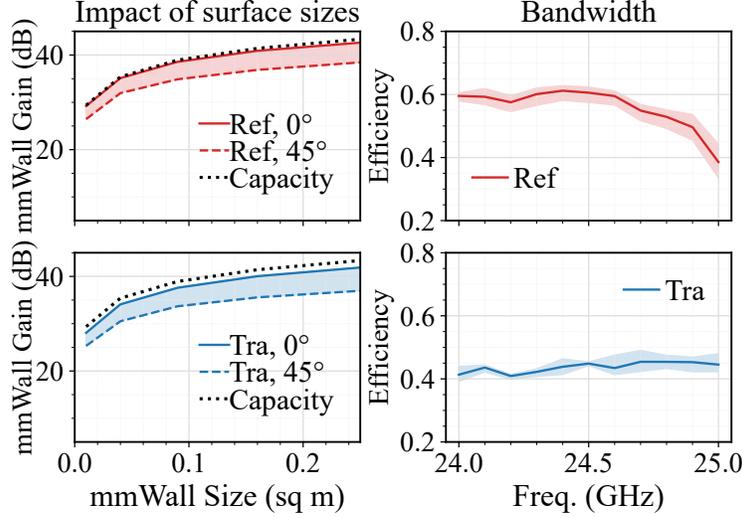


Figure 2.16: Microbenchmarks evaluating (*left to right*;) surface size and frequency bandwidth.

greatly reduce the mmWall gain. An important observation is that even with 135° steering angle (*e.g.*, Tx angle at 60° and Rx angle at -75°), mmWall achieves more than 22 dB gain, indicating that mmWall is capable of refracting the beam in a very wide angle.

Angular reciprocity. Once mmWall achieves alignment for the downlink channel, the uplink channel also becomes aligned due to its angular reciprocity. To demonstrate this property, we evaluate the accuracy of uplink beam alignment and the corresponding mmWall gain when downlink alignment is already established. In Fig. 2.15 (*upper third*), uplink alignment using reciprocity is very accurate and is within an error of 3° . Also, Fig. 2.15 (*lower third*) shows that all corresponding mmWall gains are above 23 dB using reciprocity.

Increasing mmWall size. Fig. 2.16 (*fourth*) shows an increase of simulated gains as mmWall size increases from 10×10 cm to 50×50 cm with 0° (for reflection, it is a specular reflection) and 45° steering angle. Also, we compare our simulated results against the effective aperture-based capacity. mmWall gains at both 0° and 45° steering angle increase with increasing surface size, following the capacity trend.

Operation across wide bandwidths. To demonstrate mmWall’s phase coverage across a wide bandwidth, we present our VNA measurements from 20 to 30 GHz. In Fig. 2.17, each

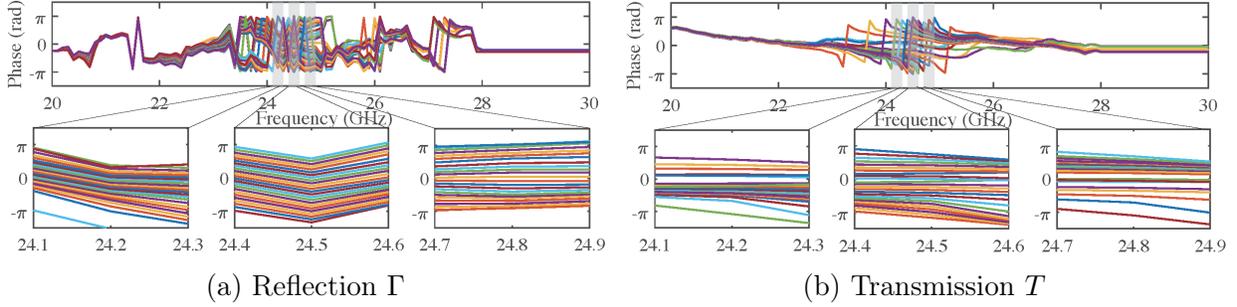


Figure 2.17: mmWall’s phase coverage and consistency (VNA measurement) across different frequencies. The curves indicate the voltage pairs (U_M, U_E) that provide -180° to 180° phase shift with the step of 15° at 24.5 GHz. The phases are unwrapped across the mmWall’s operating bandwidth.

curve indicates the phase response of voltage levels in our lookup table that we compiled at our center frequency, 24.5 GHz. Here, we emphasize three points. First, mmWall provides a full phase coverage from $-\pi$ to π over the 200 MHz 5G mmWave link bandwidth. Second, within 200 MHz bands (highlighted in gray), the phase distributions are mostly constant, allowing improvements over the entirety of these bandwidths. Third, mmWall can operate in the entire 23.5 to 25.5 GHz band, as it provides a wide range of phases there. Hence, mmWall operates over the mmWave 5G bandwidth. More importantly, our meta-atom design goal is to reduce transmission or reflection *loss level* with a full phase coverage. To quantify both magnitude and phase coverage at the same time, we define *efficiency* as $\sum_{\phi=-180}^{180} (T e^{-1j\phi})/360$ where T is a set of points obtained from the near-field transmissive (reflective) Huygens pattern that provides the maximum magnitude for -180° to 180° phases. Fig. 2.16 (*right-most*) demonstrates that for both reflection and transmission, the efficiency is consistent and declines after 24.9 GHz. Since targeted operational bandwidth for 5G mmWave is 200 MHz, we conclude that mmWall operates within the 5G bandwidth.

Meta-atom controllability and sensitivity. We present the Huygens pattern measured from the VNA in Fig. 2.18(a). We measure the near-field Huygens pattern in three different areas of mmWall to evaluate its sensitivity against fabrication variation. For all three areas, we observe a 360-degree phase variation with high magnitude for both transmission

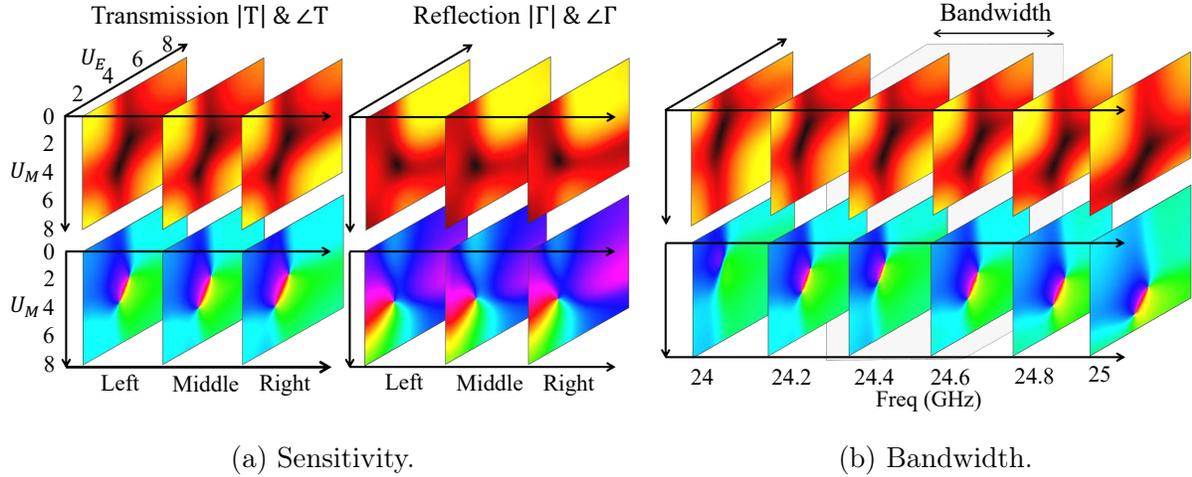


Figure 2.18: Microbenchmark on meta-atom controllability and sensitivity with near-field, real world VNA measurements.

and reflection. Moreover, the patterns do not vary across the different areas of the surface, signifying that manufacturing tolerance do not greatly affect mmWall’s near-field performance. We also demonstrate the Huygens pattern across mmWall’s operating bandwidth in Fig. 2.18(b). Within the 200 MHz bandwidth, the pattern is consistent.

2.6 Discussion

In this section, we discuss several limitations of mmWall and our potential solutions.

3D beamforming. 3D beamforming technique is beneficial in massive MIMO communications as it sophisticatedly controls the beam in different directions in spatial domain. With mmWall, meta-atoms on the same rib share the voltage level, and therefore it is structured as a 2D linear array. To achieve 3D beamforming, we can simply separate mmWall control lines in the vertical direction. In the future, we will vertically partition mmWall and separately control them.

Tinted window. While mmWave waves propagate through a glass wall with virtually no loss, the penetration loss increases when the glass is metal-coated. Our window has approximately -4 to -5 dB loss with a light tint. If our window is tinted more, SNRs will

drop, and this decrease will be equivalent to the increase of penetration loss from a different level of a tint. There is an existing work [169] that measures reflection coefficients and penetration loss for common building materials at mmWave. According to the paper, the penetration loss may increase by more than 20 dB when the window is heavily tinted. With such windows, we can remove the tint of the small area of the window (approx. 0.02 sq m) for mmWall.

Indoor AP as a 5G mmWave relay. An indoor mmWave AP can serve as a relay when the outdoor cell coverage fails to reach indoors. To accomplish this, cellular operators require indoor infrastructure to install an AP capable of receiving 5G signals. This AP then communicates with an internal modem through an Ethernet cable, and the modem wirelessly transmits the signal to the user through Wi-Fi. This deployment is not only costly and time-consuming but also hard to implement. On the other hand, a single mmWall at a fixed location can achieve all three use cases, including 5G outdoor-to-indoor and outdoor-to-outdoor coverage, and indoor WiFi.

Temperature and humidity tolerance. Thinner meander transmission lines provide more inductance. Initially, we printed these lines with a trace width of 3 mil. However, we found that at this width, the transmission lines were prone to breaking due to high soldering temperatures. To prevent line breakage during the soldering process, we recommend using a transmission trace width of 5 mil or greater and applying a soldering mask on the transmission line around the SMT component. Our components demonstrate robust performance across a wide temperature range: -40°C to $+125^{\circ}\text{C}$ for inductors and -65°C to $+150^{\circ}\text{C}$ for varactor diodes. The inductors also have a Moisture Sensitivity Level 1 rating, allowing for unlimited floor life at $<30^{\circ}\text{C}$ and 85% relative humidity. The Roger 4003C PCB substrate's dielectric constant increases linearly from approximately 0.997 to 1.005 as temperature rises from -50°C to $+150^{\circ}\text{C}$. While lower dielectric constant values are preferred at high frequencies to minimize power loss, this change is negligible. Surface performance remains stable within

this temperature range. Both electrical components can withstand a maximum soldering temperature of +260°C.

2.7 Conclusion

This chapter presents mmWall, the first Huygens metasurface that can reconfigures itself to relay an incoming mmWave beam as either a non-specular “lens” or “mirror.” Our prototype steers single- or multi-armed beams at non-specular directions, arbitrarily in real-time. We conduct an extensive evaluation in various indoor and outdoor settings, demonstrating significant SNR improvement, and describe how scaling to even larger sizes is eminently possible.

2.8 Appendices

In this section, we derive our mathematical analysis of mmWall’s magnetic and electric unit cell and the path loss model of the entire surface.

2.8.1 Appendix – Unit Cell Electromagnetic Analysis

We now present a full mathematical analysis of mmWall’s unit cells. Since electromagnetic fields are naturally continuous and will not change the propagation characteristics by itself, we artificially introduce electric and magnetic surface currents (\vec{J}_s, \vec{M}_s) from the electric and magnetic meta-atoms, enforcing a field discontinuity:

$$\vec{J}_s = \hat{n} \times [H_t - H_i], \quad \vec{M}_s = -\hat{n} \times [E_t - E_i] \quad (2.4)$$

where \hat{n} is a unit normal. The average tangential field applied on the meta-atom pair induces (\vec{J}_s, \vec{M}_s). To induce suitable surface currents, we need a proper surface impedance for each

meta-atom:

$$\begin{aligned}\hat{n} \times [E_{avg}] &= Z_e \vec{J}_s = Z_e \hat{n} \times [H_2 - H_1] \\ \hat{n} \times [H_{avg}] &= Y_m \vec{M}_s = -Y_m \hat{n} \times [E_2 - E_1]\end{aligned}\tag{2.5}$$

where Z_e is the electric surface impedance and Y_m is the magnetic surface admittance equivalent to $1/Z_m$. In fact, the electric and magnetic meta-atoms are each described by a surface impedance of LC oscillating circuit containing inductance L and capacitance C . Mathematically, we can formulate the surface impedance of the electric and magnetic meta-atom as

$$Z_e = \left(\frac{2\pi f C_e - 1}{(2\pi f)^2 L_e C_e} \right) j, \quad Y_m = \left(\frac{1 - (2\pi f)^2 L_m C_m}{2\pi f C_m} \right) j\tag{2.6}$$

where f indicates the resonant frequency. Each meta-atom behaves as an LC circuit when its resonant frequency f matches the frequency of the incident wave. Mathematically, the resonant frequency is equivalent to $f = (2\pi\sqrt{LC})^{-1}$.

Given Z_e and Y_m , we can formulate the transmission coefficient T and reflection coefficient Γ of a meta-atom pair:

$$T = \frac{4 - Y_m \cdot Z_e}{(2 + Y_m \cdot \eta)(2 + Z_e/\eta)}, \quad \Gamma = \frac{2(Z_e/\eta - Y_m \cdot \eta)}{(2 + Y_m \cdot \eta)(2 + Z_e/\eta)}\tag{2.7}$$

where η is the wave impedance in free space. Hence, by changing the surface impedance (Z_e, Y_m), we precisely control the phase of the coefficients, creating an arbitrary phase shift on the incident wave [43].

The excitation of the electric and magnetic surface currents, or, equivalently, the values of Z_e and Y_m is tuned by changing the capacitive or inductive loading of the meta-atoms as shown in Equation (2.6). Hence, to make HMS reconfigurable, we load a voltage-controlled capacitor, varactor diode, on each meta-atom. By applying voltage across each varactor, we can arbitrary change the surface impedance, or equivalently, the phase of the transmission or reflective coefficient.

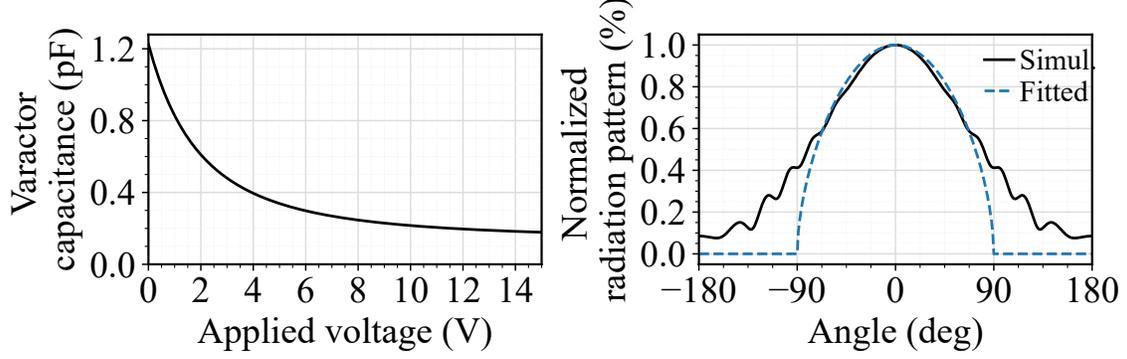


Figure 2.19: *Left*: C_{var} as the voltage applied to varactor changes, modeled with SPICE simulation; *Right*: mmWall element normalized beam pattern $F(\theta)$ simulated with HFSS and fitted function.

Since the electric and magnetic meta-atoms are superimposed on the surface, we dissect the equivalent circuit model for the electric and magnetic meta-atom individually.

Magnetic Meta-atom

In this section, we provide the formulas for the magnetic meta-atom's capacitance and inductance discussed in Section 2.3.1. First, we define the inductance of a circular metallic loop L_{loop} as

$$L_{loop} = \mu_0 R \left(\log \left(\frac{8R_m}{t+w} - \frac{1}{2} \right) \right), \quad (2.8)$$

where R is a mean radius, and μ_0 is free-space permeability. Since there is a gap on the top of a metallic loop, the inductance of our magnetic meta-atom can be calculated as

$$L_m = p_m L_{loop} = \left(1 - \frac{g}{2\pi R} \right) L_{loop}, \quad (2.9)$$

where g is a length of the gap. Now, we present the calculation of C_m . First, the gap in the metallic loop creates a parallel-plate capacitance as follow:

$$C_{gap} = \epsilon \frac{wt}{g} + \epsilon(t+w+g), \quad (2.10)$$

where w is the width of the loop, and t is the thickness of the copper. Here, $\epsilon = \epsilon_0 \epsilon_{eff}$ where ϵ_0 is free-space permittivity, and ϵ_{eff} is effective permittivity, which can be calculated as

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \left(\frac{\epsilon_r - 1}{2} \right) \left(\frac{1}{\sqrt{(1 + 12t/e)}} \right) \quad (2.11)$$

where ϵ_r is the permittivity of the substrate. Second, there is a capacitance induced by the metallic ring itself:

$$C_{surf} = \frac{2\epsilon(t + w)}{\pi} \ln \left(\frac{4R}{g} \right) \quad (2.12)$$

Lastly, the varactor diode adds the capacitance as discussed in Section 2.3.1. We have modeled our varactor, of Macom MAVR-000120-1411, based on its *Simulation Program with Integrated Circuit Emphasis* (SPICE) model and demonstrate our simulated C_{var} values in the left subfigure of Fig. 2.19. Then, we formulate C_m according to Equation (2.1). Finally, we model the circuit diagram as a series impedance where the series impedance itself corresponds to the surface impedance $Z_m = 1/Y_m$.

Electric Meta-atom

Now, we provide the capacitance and inductance calculation for the electric meta-atom. First, we formulate the inductance of a half-circle ring L_{curve} as follow:

$$L_{curve} = (p_e L_{circle}) / 2 = \frac{1}{2} \left(\left(1 - \frac{g}{2\pi R_m} \right) L_{circle} \right). \quad (2.13)$$

Based on [44], we compute the the inductance of the strip as

$$L_{strip} = \mu_0 l / 4\pi \left[2 \sinh^{-1} \left(\frac{l}{w} \right) + 2 \left(\frac{1}{w} \right) \sinh^{-1} \left(\frac{w}{l} \right) - \frac{2(w^2 + l^2)^{1.5}}{3lw^2} + \frac{2}{3} \left(\frac{l}{w} \right)^2 + \frac{2}{3} \left(\frac{w}{l} \right) \right] \quad (2.14)$$

	Radius R (mm)	Gap g (mm)	Width w (mm)
<i>Ele.</i>	0.8831	0.1016	0.3048
<i>Mag.</i>	0.7907	0.2794	0.3048

Table 2.2: mmWall design parameters.

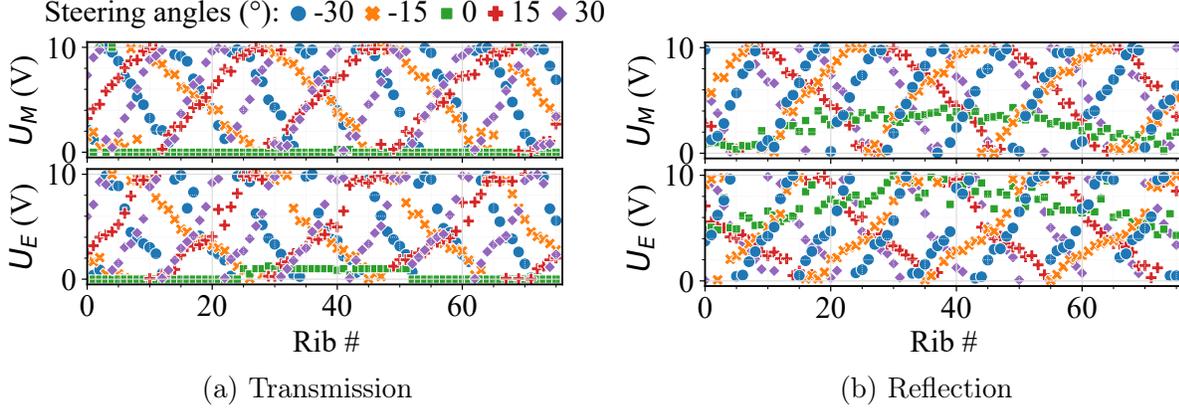


Figure 2.20: *Upper*: a set of voltage levels applied to the magnetic and electric meta-atoms U_M and U_E for transmissive steering; *Lower*: voltage levels applied for reflective steering.

where l is the length of strip, which is equivalent to $2R_m$, and w is the width of the trace.

We then combine all inductance values into L_e as

$$L_e = (L_{curve}/2) + L_{strip} \quad (2.15)$$

The formulas for the gap capacitance and surface capacitance for the electric meta-atom are the same as the magnetic meta-atom, and we define C_e according to Equation (2.2). Finally, the surface impedance of the electric meta-atom corresponds to a shunt impedance.

Design Parameters

We present the exact values for our design parameters, including radius R , gap g , and width w of the magnetic and electric meta-atom, in Table. 2.2. Also, the voltage levels applied to the magnetic and electric meta-atoms for different phase shifts are shown in Fig. 2.20. The y-axis indicates the voltage level, and the x-axis is different ribs. Specifically, Fig. 2.20(a)

demonstrates a set of U_M and U_E required for -30° , -15° , 0° , 15° , and 30° transmissive steering. Similarly, Fig. 2.20(b) shows the voltages values required for reflective steering.

2.8.2 Appendix – Path Loss Model

This section presents a standard path loss model calculation largely following the development in prior similar efforts targeting lower frequencies [135], useful for our purposes to establish the basic feasibility of our design prior to hardware fabrication and full-scale evaluation.

First let us assume that a transmitter directly communicates with a receiver. According to the Friis formula [50], the power intercepted by the receiving antenna with effective aperture Ae_R and distance between transmitter and receiver d is:

$$P_i = S_R Ae_R = \left(\frac{P_T}{4\pi d^2} G_T \right) Ae_R \quad (2.16)$$

where S_R is the received power density, and G_T is the peak gain of the transmitting antenna. Since the effective aperture $Ae_R = \frac{\lambda^2}{4\pi} G_R$ where G_R denotes the gain of the receiving antenna, we rewrite Equation (2.16) as

$$P_i = \left(\frac{P_T}{4\pi d^2} G_T \right) \left(\frac{\lambda^2}{4\pi} G_R \right) = P_T G_T G_R \left(\frac{\lambda}{4\pi d} \right)^2. \quad (2.17)$$

Now we consider a transmitter communicating with the receiver via mmWall. Given Equation (2.17), we formulate the power the nm^{th} meta-atom captures from the transmitter as

$$P_{nm}^i = P_T G_T G_w \left(\frac{\lambda}{4\pi d_{i,nm}} \right)^2, \quad (2.18)$$

where G_w denotes the gain of the meta-atom in the direction of the transmitter, and $d_{i,nm}$ is the distance between the transmitter and nm^{th} meta-atom. Similarly, we can calculate the

power received by the receiving antenna from the nm^{th} meta-atom as:

$$P_{R,nm} = P_{nm}^s G_R G_w \left(\frac{\lambda}{4\pi d_{s,nm}} \right)^2, \quad (2.19)$$

where G_w is the meta-atom gain scattered in the direction of the receiver, $d_{s,nm}$ is the distance between nm^{th} meta-atom to the receiver, P_{nm}^s is the power applied by each meta-atom, and $P_{nm}^s = P_{nm}^i \epsilon$. Here, ϵ accounts for the limited efficiency of meta-atom and insertion losses associated with components. To simplify the formula, we assume $\epsilon = 1$. To calculate the power from the transmitter to the receiver, we then combine Equations (2.18) and (2.19):

$$P_{R,nm} = P_T G_T G_R \frac{G_w G_w}{d_{i,nm}^2 d_{s,nm}^2} \left(\frac{\lambda}{4\pi} \right)^4 \quad (2.20)$$

Here, we emphasize that in the link budget, we must calculate the gain of mmWall twice, one for receiving and another for transmitting. Hence, Eq. (2.20) has two G_w . Since mmWall consists of a large array of meta-atoms, we can formulate the total received power as a sum of the received powers from all meta-atoms as

$$P_R = \left| \sum_{n=1}^N \sum_{m=1}^M C_{nm} \sqrt{P_{R,nm}} e^{j\phi_{nm}} \right|^2, \quad (2.21)$$

where $C_{n,m}$ denotes the transmission or reflection coefficient of the nm^{th} meta-atom, and the phase $\phi_{nm} = 2\pi(d_{i,nm} + d_{s,nm})/\lambda$. In a lens mode $C_{n,m} = T_{n,m}$, and in a mirror mode $C_{n,m} = \Gamma_{n,m}$. We already defined $T_{n,m}$ and $\Gamma_{n,m}$ in eq. Eq. (2.7). Finally, we write the total received power as:

$$P_R = P_T G_T G_R \left(\frac{\lambda}{4\pi} \right)^4 \left| \sum_{n=1}^N \sum_{m=1}^M C_{nm} \frac{\sqrt{G_w G_w}}{d_{i,nm} d_{s,nm}} e^{j\phi_{nm}} \right|^2. \quad (2.22)$$

However, the meta-atom gain G_w is unknown. Thus, we re-define G_w as a power radiation pattern from each meta-atom, which is equivalent to $GF(\theta_{nm})$. G is a gain that depends on

the physical area (*i.e.* the effective aperture) of the meta-atom, and $F(\theta_{nm})$ is the normalized power radiation pattern. Based on the effective aperture formula, $G = (4\pi/\lambda^2)Ae_{nm} = (4\pi/\lambda^2)(xy)$ where x and y are vertical and horizontal meta-atom spacing, respectively. Unlike traditional antennas with $x = y = \lambda/2$, our meta-atom has $x = \lambda/4.8$ and $y = \lambda/3.4$. Moreover, $F(\theta_{nm})$ defines the variation of the power radiated or received by a meta-atom:

$$F(\theta) = \begin{cases} \cos^q(\theta) & \theta \in [0, \pi/2] \\ 0 & \theta \in [\pi/2, \pi] \end{cases} \quad (2.23)$$

where θ are the angle from the meta-atom to a certain transmitting or receiving direction. In the right subfigure of Fig. 2.19, we present a simulated mmWall element beam pattern $F(\theta_{nm})$ as well as the curve fitted with Eq. (2.23). Based on our curve fit, $q = 0.5611$.

Far-field beamforming. In the far-field, we can approximate $d_{s,nm} = d_s$ and $d_{i,nm} = d_i$ since d_i and d_s are much greater than the distance between different meta-atoms. However, we do not approximate $d_{s,nm} = d_s$ and $d_{i,nm} = d_i$ for the phase ϕ_{nm} . Then, we can simplify Equation (2.22) as:

$$P_R = P_T G_T G_R \left(\frac{Ae_{nm}}{4\pi d_i d_s} \right)^2 F(\theta_i) F(\theta_s) \left| \sum_{n=1}^N \sum_{m=1}^M C_{nm} e^{j\phi_{nm}} \right|^2 \quad (2.24)$$

This indicates that we can maximize the received power by configuring each meta-atom's $\angle C_{nm}$ to $-\phi_{nm}$. Finally, the path loss of a correctly reconfigured mmWall as:

$$L_{mmWall}^{-1} = \left(\frac{xy}{4\pi d_i d_s} \right)^2 F(\theta_i) F(\theta_s) \left| \sum_{n=1}^N \sum_{m=1}^M |C_{nm}| \right|^2 \quad (2.25)$$

Since $0 < |C_{nm}| < 1$ for both transmissive and reflective mode, increasing the number of meta-atoms N and/or M reduces the path loss. Assuming $|C_{nm}|$ is close to 1, the path loss of mmWall is proportional to $1/(NM)^2$. While increasing the element spacing x and y seems

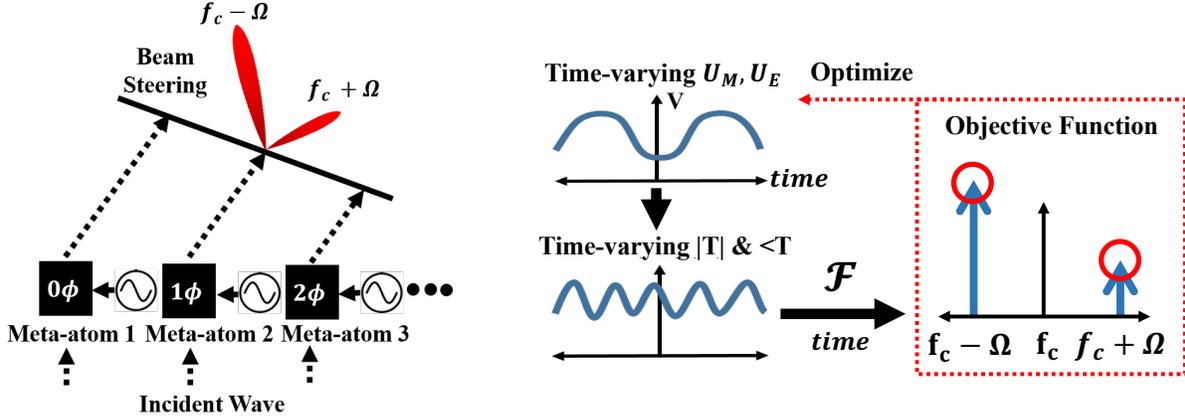


Figure 2.21: Schematic illustration of multi-armed beam generation at $f_c - \Omega$ and $f_c + \Omega$ and multi-armed beam-steering. f_c is the carrier frequency, and Ω is the voltage modulation frequency.

to reduce the loss, it is not always true because $|C_{nm}|$ decreases when x and y increase due to increasing coupling between adjacent meta-atoms.

2.8.3 Appendix – Multi-armed Beam via Time Modulation

Time-modulated beam steering and splitting

To transform an incoming beam into multiple beams at different frequencies, we can introduce *time* as an additional degree of freedom in the control line. This is achieved by convolving a high-frequency mmWave signal with a low-frequency signal to produce a desired Fourier series. Specifically, we apply time-modulation to the voltage signal, resulting in a time-varying transmission amplitude $|T|$ and phase $\angle T$.

When this time-varying transmission signal is periodic, its Fourier transformation produces harmonics, creating multiple beams at different frequencies, known as *sidebands*. By applying an appropriate time-varying voltage signal, mmWave can generate a time-varying signal response, whose Fourier transformation creates a desired number of beams at specific frequencies.

We define the time-varying voltage signals \tilde{U}_E and \tilde{U}_M for the electric and magnetic meta-atoms as $U_{amp}f(t) + U_{offset}$ where U_{amp} is the voltage amplitude, U_{offset} is the voltage

offset, and $F(t) = \sum_n^N a^{(n)} \cos[n(\Omega t - \varphi)] + b^{(n)} \sin[n(\Omega t - \varphi)]$ is a normalized Fourier series with a modulation frequency Ω , time t , and phase φ . Our goal is to find the solution Θ^* to the following optimization problem:

$$\Theta^* = \arg \max_{\Theta} OBJ(F(T(\Theta))) \quad (2.26)$$

where Θ^* is an optimal set of the voltage waveform coefficients $(U_{amp}, U_{off}, \{a\}, \{b\}, \varphi)$. $T(\Theta)$ is a mapping function from the voltage waveform to the Huygens' pattern, and $F(T(\Theta))$ is a Fourier transformation of the time-varying transmission signal. Finally, $OBJ(F(T(\Theta)))$ is the objective function characterizing the scattered power of the desired beams at a desired frequency bin.

Example. Consider the multi-armed beam scenario in Fig. 2.21. The goal is to split the beam into two: one at $f_c + \Omega$ and another at $f_c - \Omega$, where Ω is the voltage modulation frequency. We want to concentrate more energy towards the beam at $f_c - \Omega$ than the beam at $f_c + \Omega$. We redefine our goal to find a proper Θ such that $F(T(\Theta))$ has a large peak at $f_c - \Omega$ and a relatively low peak at $f_c + \Omega$. Our objective function maximizes the sum of the weighted power at $f_c - \Omega$ and $f_c + \Omega$, as denoted with the red circles in Fig. 2.21. After optimizing with a genetic algorithm, mmWall concurrently steers the multi-armed beams by applying different phase shifts to the optimized voltage waveform of each meta-atom pair.

Simulation Results

To assess the feasibility of time-modulation, which is not implemented in our hardware prototype, we conducted simulations. We used 24 GHz mmWave signals with a voltage modulation frequency (Ω) of 30 MHz. Our evaluation covered three scenarios. The first scenario involved a single transmissive beam, where we transformed the incident wave at f_c to a signal at $f_c - \Omega$. The second and third scenarios explored multi-armed beams. In these cases, we split the incoming beam into two separate beams: one at $f_c - \Omega$ and another at

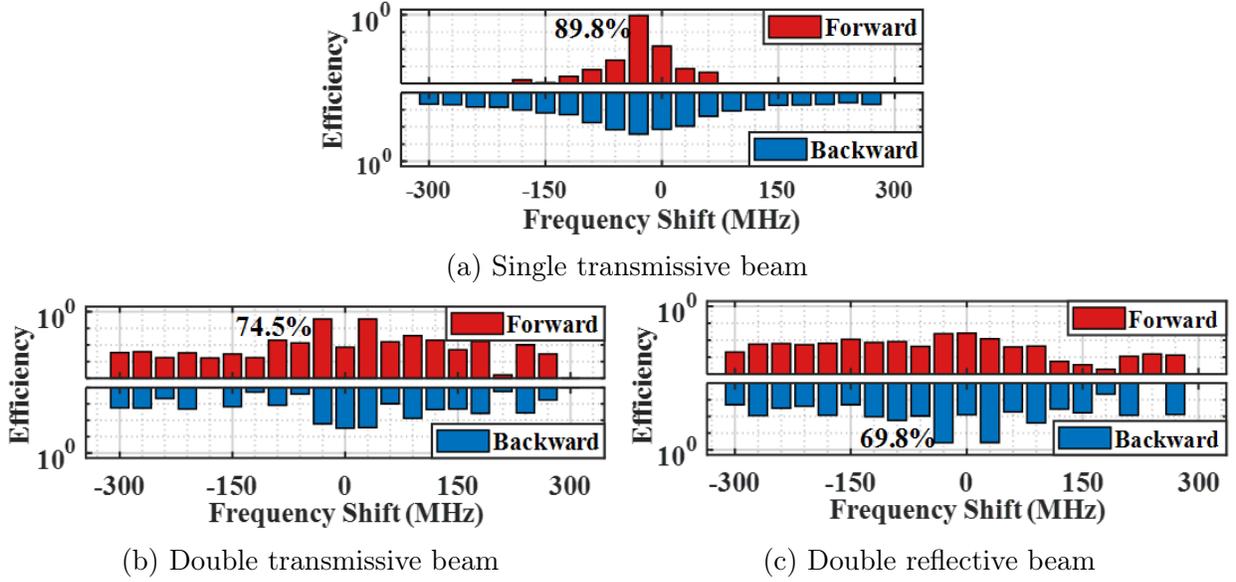


Figure 2.22: Time-modulated multi-beam. Simulated beam efficiency versus frequency shift in log scale.

$f_c + \Omega$. Specifically, the second scenario focused on two transmissive beams, while the third scenario examined two reflective beams.

Beam Efficiency. Figure 2.22 demonstrates the efficiency of mmWall across three scenarios. For the single transmissive beam (Fig. 2.22(a)), we observe a single peak at the -30 MHz frequency shift in the forward spectrum. Here, the frequency shift of 0 MHz indicates the carrier frequency at 24 GHz, and the -30 MHz frequency shift denotes 24 GHz $-$ 30 MHz, which is equivalent to $f_c - \Omega$. This peak has 89.8% efficiency, indicating that nearly 90% of the incident signal is relayed with only -0.46 dB loss. For the double transmissive beams and double reflective beams, the sum of beam efficiency at $f_c - \Omega$ and $f_c + \Omega$ is 74.5% (approximately 1.28 dB loss) and 69.8% (approximately 1.56 dB loss), respectively. As we steer the beam away from 0 degrees, where 0 degrees means no phase shift applied across the meta-atoms, the peak power of the beam weakens slightly. For all three scenarios, there is about -2 dB loss as we steer the beam by -30 or 30 degrees. At -90 or 90 degrees of the beam-steering angle, there is approximately -3 dB signal loss. We further note that for the

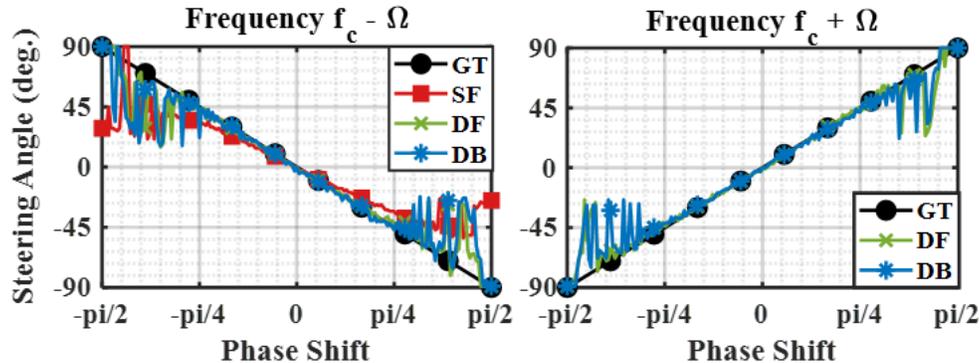


Figure 2.23: Simulated steering angle accuracy with time modulation. GT, SF, DF, and DB stand for groundtruth, single forward, double forward, and double backward, respectively.

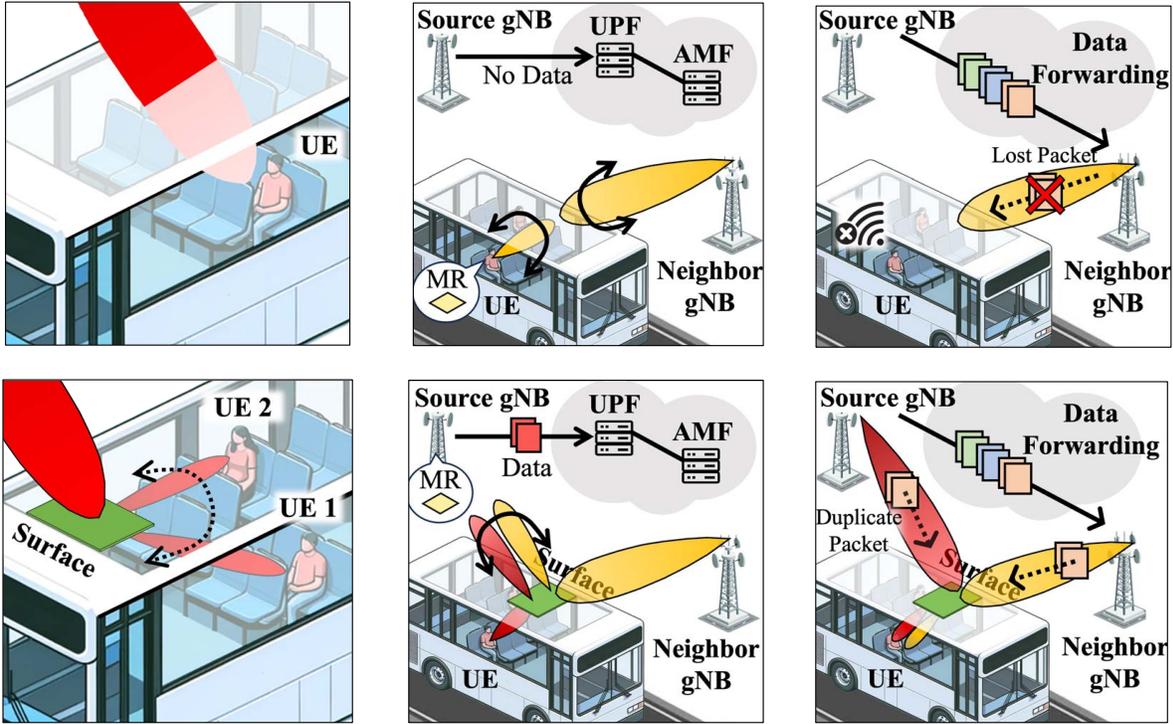
double-beam scenarios, mmWall correctly steers the beam at $f_c - \Omega$ and another beam at $f_c + \Omega$ in different directions, simultaneously.

Beam Steering Accuracy. Figure 2.23 evaluates the beam steering performance across three scenarios as we vary the relative phase value ϕ across different meta-atoms. We specifically measure the angle of the beam's peak. The graph is divided into two parts: the left graph shows the steering performance of the beam at $f_c - \Omega$ for all three scenarios, while the right graph illustrates the steering accuracy of the beam at $f_c + \Omega$ for the double beam scenarios only. Ground-truth values are represented by black markers with circles. For all scenarios, the beam steering angle demonstrates high accuracy, with some exceptions. Accuracy decreases for angles corresponding to phase shifts below $-\pi/4$ and above $\pi/4$. This decrease in accuracy occurs because, as the beam steers to a larger angle, it may become asymmetric. Consequently, the peak may not be at the center of the beam.

Chapter 3

Metasurface-Enabled NextG mmWave for Roadside Networking

This chapter presents the design and implementation of Wall-Street, a smart surface that transforms 5G mmWave connectivity, specifically for vehicular use scenarios. In such high-mobility scenarios, we install **Wall-Street** on vehicles to improve connectivity for users inside. Wall-Street leverages the concept of a programmable mmWave smart surface [27] to manipulate and steer mmWave signals, enabling seamless coverage and reliable communication for vehicle users. By strategically deploying Wall-Street on the exterior of vehicles, we aim to mitigate the impact of signal blockage and attenuation caused by the vehicle body, ensuring that mmWave signals can effectively penetrate and reach users inside the vehicle. Moreover, the handover (HO) process in mmWave networks consumes at least twice the energy per unit distance, compared to low-band HOs [63]. Wall-Street reduces UE HO energy consumption by letting the surface perform neighboring cell signal strength measurements. Three key innovations of Wall-Street (Fig. 3.1) enable seamless vehicular-speed mmWave networking:



(a) Enhance in-vehicle mmWave penetration. (b) Transflexive handover, on-going communication. (c) Enhance handover reliability.

Figure 3.1: Wall-Street’s design innovations electronically split, shape, and steer mmWave transmissions in real time to enable seamless roadside 5G New Radio networks (upper : without Wall-Street; lower : with Wall-Street).

1. In-vehicle mmWave coverage. Wall-Street efficiently steers outdoor mmWave signals into a vehicle, providing consistent coverage for all users within (Fig. 3.1(a)) and eliminates the need for individual users to each undergo separate handover processes.

2. Seamless handover with concurrent communication. Innovating beyond previous programmable mmWave surfaces [27], Wall-Street is the first design that can simultaneously reflect, refract, and split the mmWave beam at arbitrary angles, with low signal strength loss. Leveraging this *transflective* property, the Wall-Street system enables simultaneous data exchange and neighboring cell measurement by reflecting the synchronization signal bursts from neighboring cells directly to the serving BS while refracting the data communication link between the serving BS and UE, as shown in Fig. 3.1(b) (*right*). This innovative dual

functionality (simultaneous transmissive and reflective beam steering) allows for seamless handovers without service interruption, as the UE can continue to communicate with the serving base station (gNB) while the gNB performs the necessary measurements.

3. Make-before-break handover. At the moment of handover, when signal strength is typically weakest, Wall-Street facilitates the transmission of duplicate data packets from both the source and the new serving cell by combining two beams and steering them directly to the UE. This dual-beam combination significantly boosts the reliability during handover execution. By adjusting the power of the respective links, Wall-Street ensures an uninterrupted flow of data, minimizing packet loss and retransmissions.

We have designed and implemented Wall-Street hardware and integrated it into the COSMOS testbed [117], incorporating new PHY features on its USRP-based mmWave gNBs and UE. This integration enables real-time handover experiments in environments that match the scenarios depicted in Fig. 3.1. Section 3.3 provides a detailed description of our implementation of the Wall-Street-integrated COSMOS testbed. We mount Wall-Street on the rear door of an SUV vehicle and deploy multiple mobile UEs on the rear seat and back seat, with three gNBs placed on the first floor of a lab facing the road. We also implemented the 5G Standalone handover protocol and used it as a performance benchmark for our evaluation. To evaluate TCP end-to-end performance, we use a state-of-the-art mmWave ns-3 simulator [99] with our own modifications, in SA mode. The simulation is trace-driven, fed with PHY traces collected from real-time experiments. In multi-user scenarios, our results demonstrate a $2\times$ improvement in TCP throughput, a $0.7\times$ reduction in delay, and a $0.5\times$ reduction in unnecessary handover events compared to the baseline 5G Standalone handover protocol. Also, we observe more than a $0.5\times$ reduction in unnecessary handover events. We further evaluate the impact of vehicle speed and the performance improvement at different handover stages. Microbenchmarks investigate Wall-Street’s in-vehicle coverage improvement, beam tracking, and multi-link capabilities, and beam patterns, providing insights into new capabilities offered by Wall-Street.

3.1 Primer: 5G Mobility Management

We introduce the state of the art in 5G New Radio (NR) Standalone (SA) beam acquisition and tracking (§3.1.1), followed by 5G SA handover (§3.1.2), standardized in 3GPP 38.802 [1]. In contrast to Non-Standalone (NSA) handover, SA handover obviates the need to involve legacy 4G infrastructure in the handover [63], making it the cutting edge in 5G.

3.1.1 Initial Attachment and Tracking

Initial attachment of the UE to the gNB begins via a coarse spatial *beam sweep* of beam directions at both gNB and UE, which 3GPP calls *Procedure 1* (P-1). During the beam sweep in the first 5 ms of each period, the gNB transmits a *synchronization signal burst* comprised of 64 *Synchronization Signal Blocks* (SSBs). The gNB groups M SSBs together, transmitting all of them in the same direction before moving on to the next direction, as shown in Fig. 3.2(a). This transmission happens periodically, every 20 ms in the mmWave FR2 band by default. The UE scans all its receive directions in the corresponding time it takes the gNB to send M SSBs, so rendezvous is possible. The UE sends its measurement of the SSB's signal strength in a *measurement report* (MR) back the gNB, which confirms the transmit beam, thus initially aligning the transmit and receive beams. Finally, the UE and gNB exchange data through the selected beam patterns.

After initial attachment, the beam directions between the gNB and UE need to be refined in the face of UE mobility, a process 3GPP calls *Procedure 2* (P-2). The gNB uses transmit beams narrower than those used in P-1, and centered on the beam direction chosen during P-1. During P-2, the gNB scans different fine-grained directions on different channel state information reference signal (CSI-RS) resource blocks. The UE sends a CSI-RS to the gNB, which chooses the optimal fine-grained transmit beam direction. To track and refine UE beam directions, *Procedure 3* (P-3) uses CSI-RS combined with a static transmit beam and different UE receive beams.

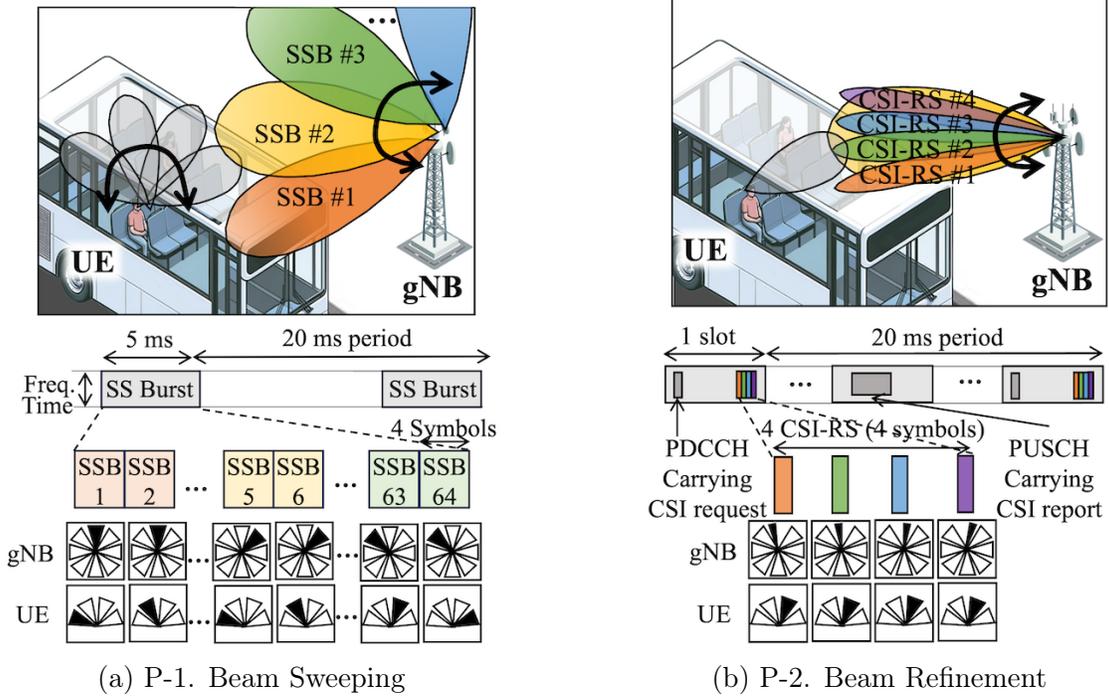


Figure 3.2: 5G SA beam acquisition and beam tracking.

3.1.2 5G SA Handover

Today's state of the art 5G SA handover is called *Xn handover*, named after the *Xn interface* that allows gNBs to exchange information related to the handover process. The handover procedure begins with the UE periodically generating and sending a MR containing the neighboring gNB's identifier, and the corresponding signal strength to the *source* gNB (with which the UE is currently communicating), which then takes the decision to start the handover procedure to the best *target* gNB. Finally, the target gNB completes the procedure. SA Xn handover is divided into three phases: *Preparation*, *Execution*, and *Completion*, as outlined in Fig. 3.3.

(1) Preparation phase. 5G SA relies on client-side feedback to trigger handovers. The source cell requests the UE to stop the ongoing transmission and perform P-1 measurements to measure the received power of the reference signal (*Reference Signal Received Power*, or RSRP) for each neighboring cell. A common trigger for 5G handover is *event A3* [34, 159], which specifies that the signal strength of the neighboring cell is higher than that of the

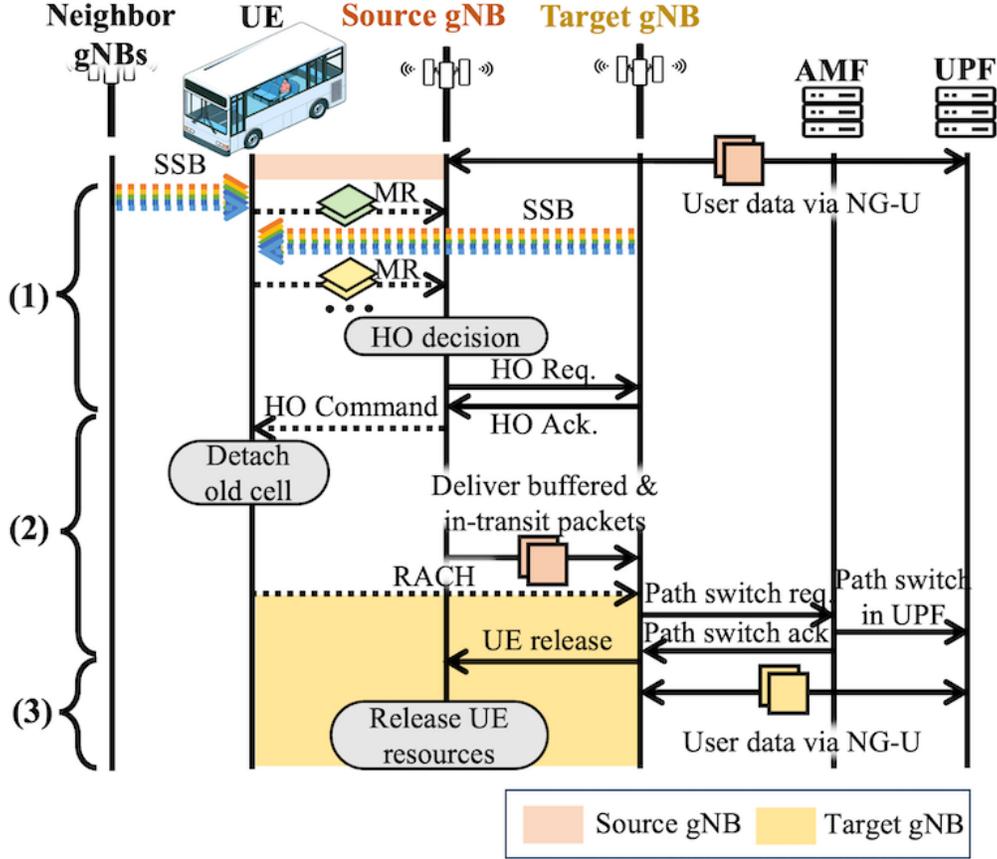


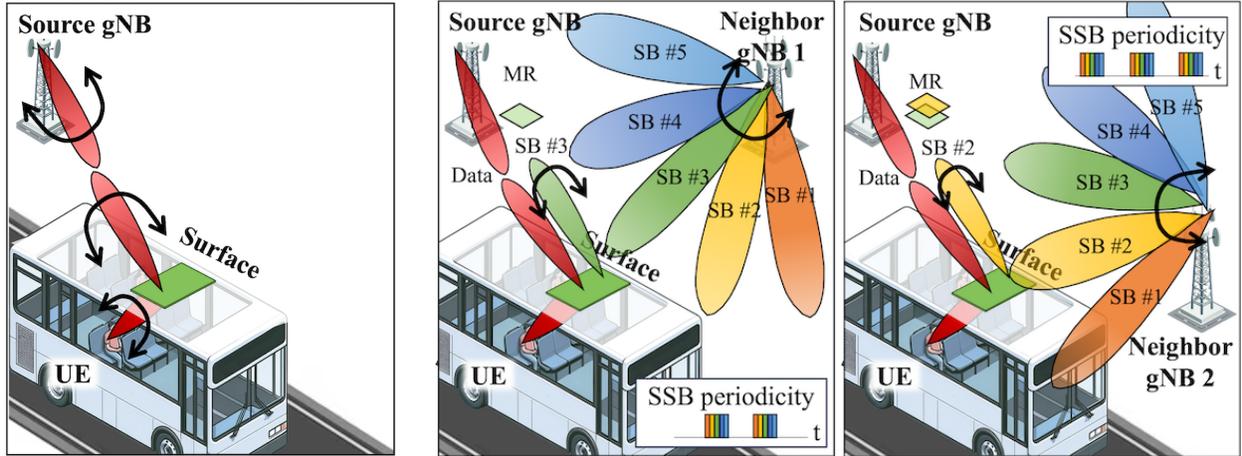
Figure 3.3: 5G SA Xn handover: (1) Preparation phase; (2) Execution phase; (3) Completion phase. Data communication is denoted with orange and yellow rectangles.

current serving cell by a certain hysteresis threshold, *i.e.*:

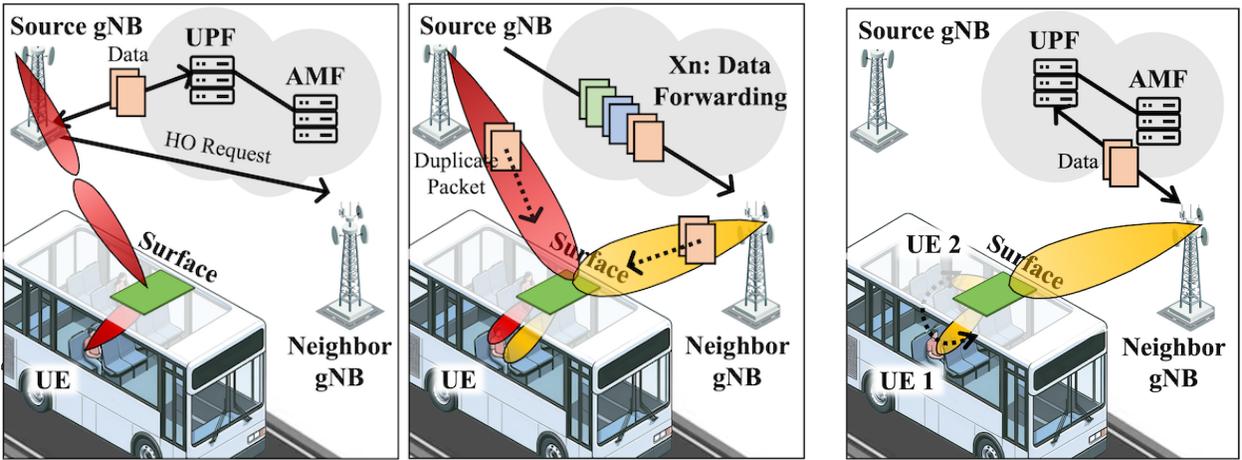
$$M_n - H > M_s \quad (3.1)$$

where M_n and M_s are the respective RSRPs of the neighboring and source cells, and H is a hysteresis parameter. To avoid ping-pong handover caused by signal fluctuations, a time-to-trigger (TTT) is used as the time hysteresis: the HO is triggered when Equation (3.1) holds true for the TTT duration.

(2) Execution phase. The source gNB decides to initiate a handover and sends a *handover request* (HO Req) —Fig. 3.3, Step (2)—to the chosen target gNB over the Xn interface. The target BS responds with an ack, providing the necessary resources for the UE. Upon receiving



(a) Initial attachment between UE and RAN. (b) Handover preparation: Measurement of candidate gNB 1 (left) and gNB 2 (right).



(c) Make-before-break handover execution in the core (left) and physical layer (right). (d) Wall-Street handover completion.

Figure 3.4: Wall-Street design overview: (a) Initial UE-RAN attachment (Section 3.2.2); (b) preparation using transfective surface power measurement to multiple cells (Section 3.2.3); (c) beam combining to enable make-before-break dual-cell connectivity (Section 3.2.3); (d) handover completion (Section 3.2.3).

the ack, the source BS sends a *Handover Command* (HO Command) to the UE, instructing it to detach from the source gNB and synchronize with the target gNB using the *Random Access Channel* (RACH) procedure. The source gNB forwards buffered and in-transit data to the target gNB.

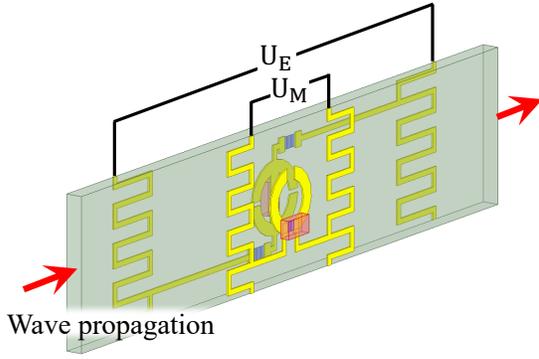
(3) Completion phase. The target gNB sends a Path Switch request message to the the 5G Core (5GC), informing it about the handover and requesting a path update. The 5GC updates the path and sends an ack to the target gNB. The target gNB sends a UE Context Release message to the source gNB over the Xn interface, indicating that the handover is complete. The source gNB then releases the UE’s context.

3.2 Wall-Street: Design

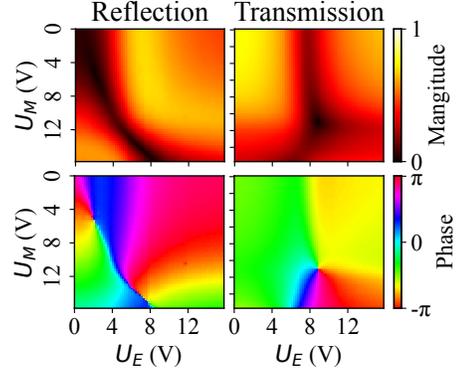
This section presents the design of Wall-Street as shown in Fig. 3.4. Section 3.2.1 describes the Wall-Street surface. Section 3.2.2 explains the normal operations of UE attachment and tracking. Section 3.2.3 walks through Wall-Street’s handover.

3.2.1 Surface Hardware

The Huygens metasurface (HMS) is a metamaterial smart surface that reflects and refracts signals, and that prior work has demonstrated at 24 GHz [27]. Wall-Street for the first time demonstrates an HMS at 26 GHz. HMS comprises a layer of co-located orthogonal electric and magnetic dipoles, facing each other across dielectric substrate in a 3D structure, that introduces a discontinuity in the electromagnetic fields. It hence provides the means for manipulation of all attributes of the incident field, including its magnitude and phase, in both reflective and transmissive directions. Figure 3.5(a) shows Wall-Street’s electric unit cell on the back of the substrate and Wall-Street’s magnetic unit cell on the front of the substrate. Figure 3.5(b) presents Vector Network Analyzer (VNA) measurements of the unit cell response at 26 GHz. By adjusting the voltage applied to both the magnetic (U_M) and



(a) Unit cell design.



(b) Unit cell response.

Figure 3.5: VNA measured unit cell response: reflection and transmission through the surface.

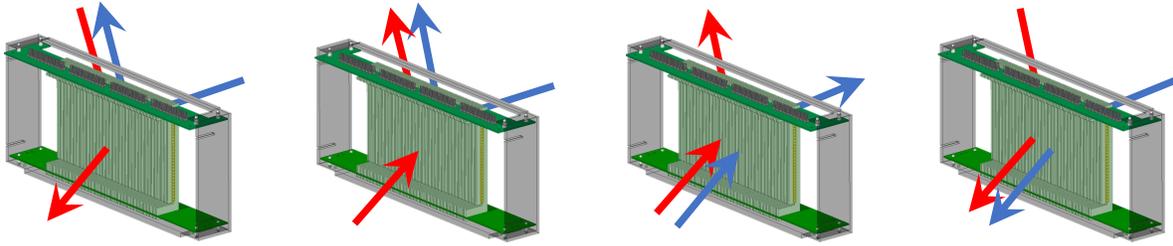


Figure 3.6: Wall-Street’s multi-beam operational modes (*left to right*): bi-directional split and combination for handover preparation and uni-directional split and combination for handover execution.

electric (U_E) unit cells, the surface achieves 2π phase coverage with near-unity magnitude for both reflection and transmission.

The Wall-Street surface is composed of 76 boards, each consisting of 28 co-located magnetic and electric unit cells. A control unit connected to Wall-Street supplies a set of voltages to the boards, and the beam is collectively shaped by all boards forming an array factor. Wall-Street has flexible multi-beam capabilities, including bi-directional splitting and combining and uni-directional splitting and combining, as shown in Fig. 3.6. Due to its angular reciprocity, the voltage configurations of splitting and combining are identical. We refer the reader to Section 3.7 for the details on our design parameters.

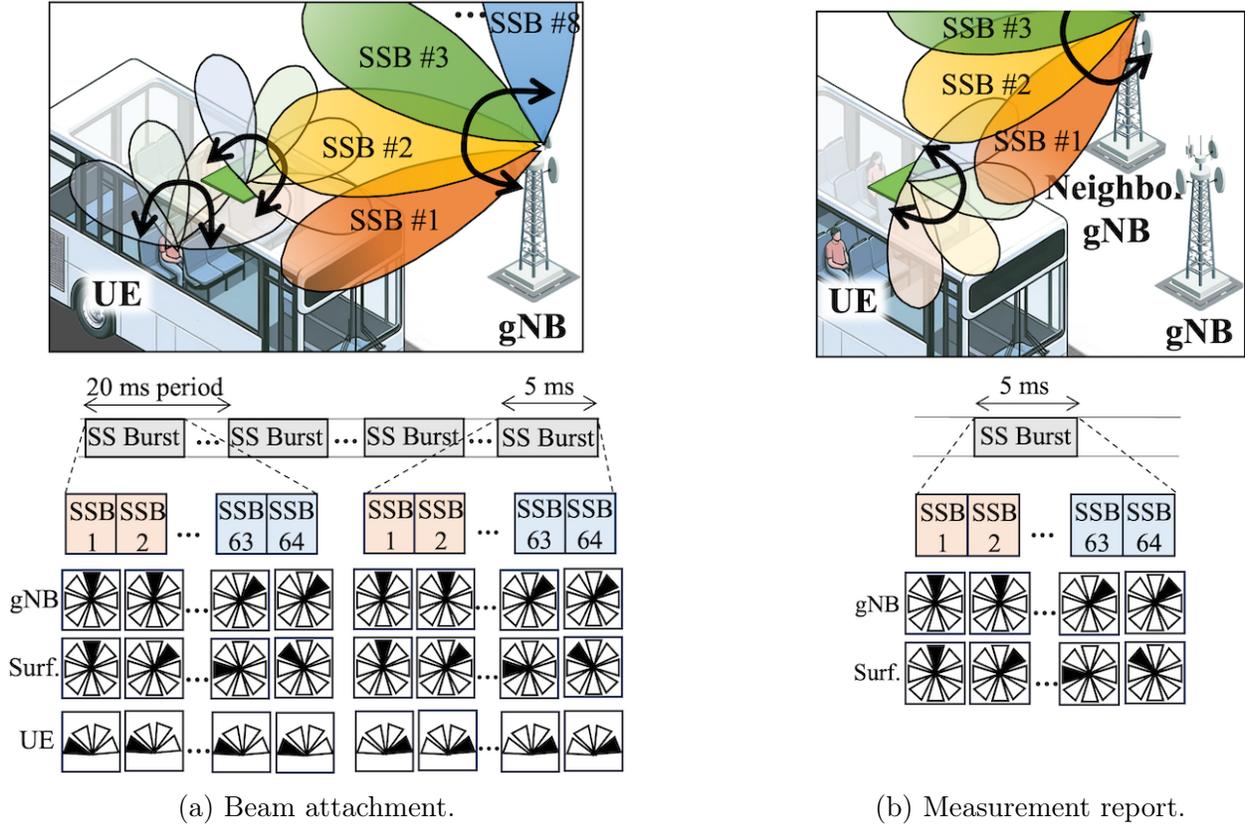


Figure 3.7: Wall-Street process: beam attachment and measurement report process corresponding to Fig. 3.4(a) and Fig. 3.4(b), respectively. Beam attachment and measurement report require four SS bursts and one SS burst of five-millisecond duration.

3.2.2 Attachment and Tracking

Wall-Street's normal operation requires UEs to attach to the RAN, and the RAN to track their movement. We detail how the UE initially attaches to the gNB and determines an optimal set of beam patterns via SSB measurements (Section 3.2.2), and how the gNB tracks the Wall-Street UE in motion (Section 3.2.2)

Initial Attachment

Similar to 5G NR, Wall-Street attachment adapt to function under practical timescales with a surface. First, a UE needs to be able to attach to the gNB of the RAN through Wall-Street. This process, shown at a high level in Fig. 3.4(a), begins with a SSB beam sweep at each of the gNB, Wall-Street, and UE, as detailed in Fig. 3.7(a). During a single SS burst, the

UE fixes its beam angle while both Wall-Street and the gNB spatially sweep their beams. With each UE angle, the UE makes an SS Burst power measurement, each with a different beam pattern—each of which it feeds back to the gNB via a *beam report*. Practically, four SS bursts are needed (corresponding to each of four different UE receive beam patterns) for this initial paging search (80 ms).

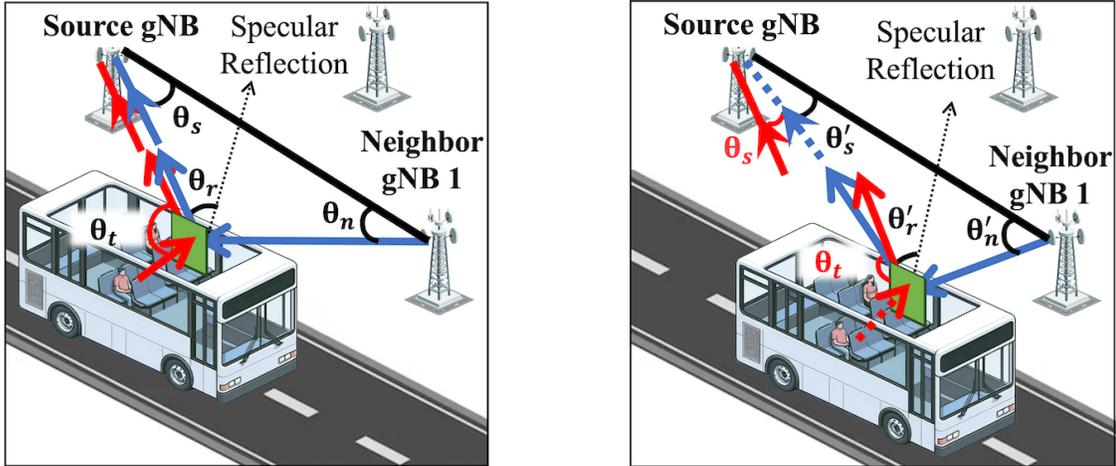
Traditional repeaters require a $O(n^4)$ search for initial attachment, involving the gNB’s transmitting beam, repeater’s beam to gNB, repeater’s beam to UE, and UE’s receiving beam. This process takes about five seconds, which is impractical for rapidly moving vehicles. Our surface, however, directly manipulates the incident beam without separate receiving and transmitting modules, requiring a search in only one degree-of-freedom. This reduces initialization to a $O(n^3)$ search, considering only the gNB’s transmitting beam, surface’s beam to UE, and UE’s receiving beam. Moreover, Wall-Street’s angular reciprocity enables faster switching between uplink and downlink compared to repeaters.

Mobile User Tracking

Wall-Street adopts and leverages 3GPP P-2 and P-3 functionality for gNB- and UE-side beam tracking, respectively (Section 3.1.1). In addition, Wall-Street devises the following surface-driven *path diversity tracking* algorithm that fuses measurement reports acquired from handover preparation, to bolster the reliability of the sensitive mmWave tracking process. Instead of relying on client-side report, the source gNB directly generates measurement reports by triggering Wall-Street to reflect neighboring gNB’s SSBs and receiving reflected SSBs, as shown in Fig. 3.4(b) and Fig. 3.7(b). We will further describe this procedure in Section 3.2.3.

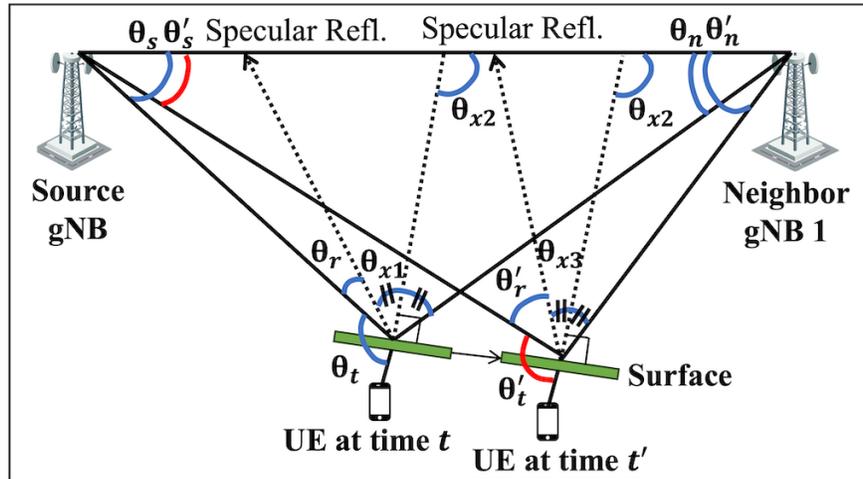
After beam acquisition and measurement report at a given time t , the source gNB has knowledge of **1)** θ_s , its own downlink beam angle, **2)** θ_t , the angle by which Wall-Street deflects the UE’s uplink transmission, **3)** θ_r , the (non-specular) angle¹ by which Wall-Street

¹Specular reflection is a reflection at an equal but opposite angle to the incident light, as on mirrors.



(a) Align (at time t)

(b) Update (at time t')



(c) Wall-Street beam tracking calculation

Figure 3.8: Beam tracking: Wall-Street's surface-enabled beam tracking leverages neighboring gNBs to add path diversity.

reflects a neighboring gNB’s downlink transmission, relative to the hypothetical specular reflection angle, because the source gNB has control over the surface, and **4)** θ_n , the neighboring gNB’s downlink beam angle via mapping a SSB Beam ID from reflected SSBs, as diagrammed in Fig. 3.8(a). Importantly, we observe that every measurement report period of time (160 ms) a neighboring gNB sends an SSB burst, updating the source cell’s knowledge of parameters θ_r and θ_n (which we define θ'_r and θ'_n respectively), as shown in Fig. 3.8(b).

Now our objective is to calculate **1)** $\Delta\theta_s$, a source gNB beam angle update, and **2)** $\Delta\theta_t$, a surface transmissive beam steering angle update, both for the time step $t'-t$. For this, we first calculate the (hypothetical) specular reflection angle of the SS burst from the neighboring gNB $\theta_{x1} = (180^\circ - (\theta_s + \theta_n) - \theta_r) / 2$ as illustrated in Fig. 3.8(c). Then, we calculate the normal direction of the surface with respect to the source and neighbor gNBs $\theta_{x2} = \pi - (\theta_{n1} + \theta_{x1})$. In 160 ms, a vehicle moving at 40 mph travels less than three meters, so the surface orientation remains approximately the same. Hence we calculate the specular angle of SSB signals at t' as $\theta_{x3} = 180^\circ - \theta'_n + \theta_{x2}$. Finally, the source cell angle update is $\Delta\theta_s = \pi - (2\theta_{x3} + \theta_{r2}) - \theta_{n2} - \theta_{s1}$, and the surface transmissive beam deflection angle update is $\Delta\theta_t = (\theta_{x3} + \theta_{r2}) - (\theta_{x1} + \theta_{r1})$.

3.2.3 Cell-to-Cell Handover

This section describes Wall-Street’s cell to cell handover, the first step of which is preparation, which also leverages Wall-Street’s transfective capability to take reflective measurements of neighboring cells while communication through the surface continues uninterrupted (Section 3.2.3). Section 3.2.3 presents our make-before-break handover, made possible by the combining of two beams: one from the source cell and another from the target cell. Section 3.2.3 concludes with a description of handover completion, releasing the source cell. Figure 3.9 summarizes the overall sequence of events, which we refer to throughout this section.

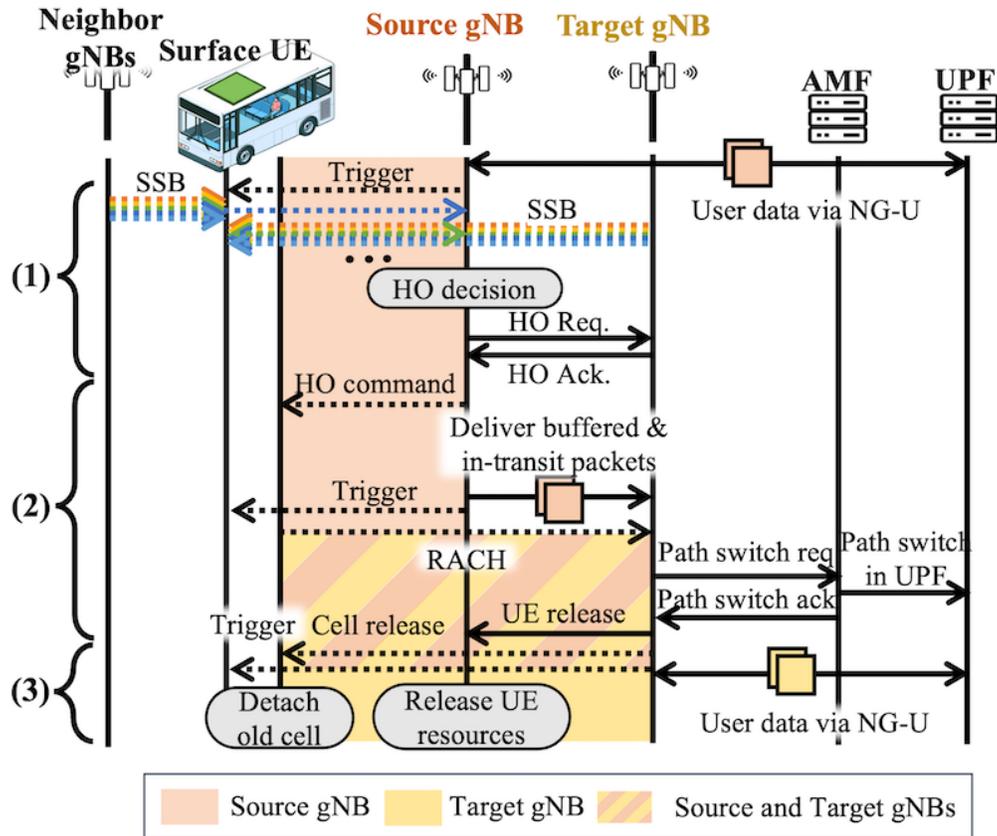


Figure 3.9: Wall-Street’s handover timeline: (1) Preparation, (2) Execution, and (3) Completion phases are indicated.

Handover Preparation

We begin at the network layer, which is implemented inside the 5G core. Before handover, the source and neighbor gNB already have an active Xn (Section 3.1.2) connection, and the UE is in the RRC *connected* mode, sending and receiving data to and from the source gNB. Periodically, the source gNB sends a control signal via the robust, sub-6 GHz FR 1 band to Wall-Street, triggering the preparation at the physical layer that we describe next, a process that measures the RSRP of reflected SSBs. To avoid collisions, the gNBs separate their respective SSBs in frequency and time across different gNBs, allowing the source gNB to listen to other cells’ SSBs while simultaneously exchanging data with the UE.

Handover preparation at the PHY: Upon receiving a beam measurement trigger from the source gNB, Wall-Street alternates between bi-directional beam combination for uplink

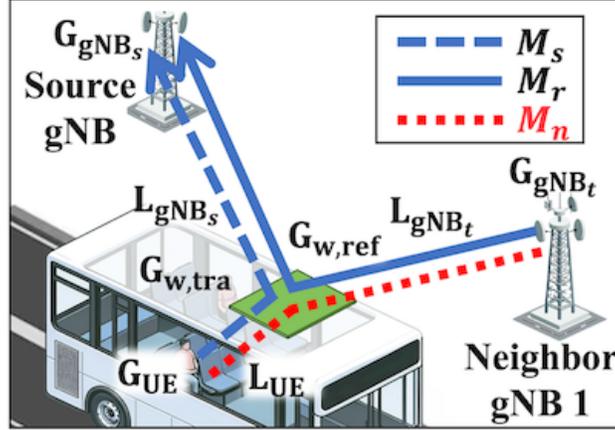


Figure 3.10: Measurement power translation: Inferring 3GPP-standard measurement report power readings based on Wall-Street’s reflective measurement report readings.

data (refracting signals from the UE and reflecting signals from neighboring gNBs) and bi-directional beam splitting for downlink data (refracting signals to the UE and reflecting signals from neighboring gNBs). During this process, the UE’s beam angle, the source gNB’s beam angle, and Wall-Street’s refractive beam angle remain the same, and only Wall-Street’s reflective beam and the neighboring gNB’s beam sweep in one SS Burst, as depicted in Fig. 3.7(b).

As Wall-Street allows flexibility in the proportion of power allocated to its refractive and reflective beams on the uplink, now the issue arises of how to compute a power split. Given that the reflective beam travels a longer distance, it reduces the power allocated to the transmission link to the minimum required for maintaining communication, while directing the remaining power to the reflective link (see Fig. 3.24 in Section 3.4.3).

RSRP inference for HO decision. To correctly trigger the HO event, the source gNB infers 3GPP-standard measurement report power readings based on Wall-Street’s reflective measurement report readings. The RSRP of the neighbor gNB’s measurement report as received by the source gNB, M_r , is distinct from M_n in Eq. (3.1), and so to apply Eq. (3.1), we need to derive a sound measurement for M_n based on the novel reflective measurement M_r that Wall-Street makes. As illustrated in Fig. 3.10, M_n is the signal strength between the neighboring gNB to UE through the surface (red dotted line). To infer M_n , Wall-Street

uses the RSRP of a link between the source gNB and UE (M_s) and between the neighboring cell and the source gNB (M_r). We simplify the link budget of each RSRP as follows:

$$M_s = G_{\text{gNB}_s} + L_{\text{gNB}_s} + G_{\text{w,tra}} + G_{\text{UE}} + L_{\text{UE}} + L_{\text{car}} \quad (3.2)$$

$$M_n = G_{\text{gNB}_n} + L_{\text{gNB}_n} + G_{\text{w,tra}} + G_{\text{UE}} + L_{\text{UE}} + L_{\text{car}} \quad (3.3)$$

where G_{gNB_s} and G_{gNB_n} are the gain of the source and neighbor gNBs respectively, L_{gNB_s} and L_{gNB_n} are the free-space path loss between the source and neighbor gNBs and the surface, respectively, $G_{\text{w,tra}}$ is the transmission gain of the surface, G_{UE} is the gain of the UE, L_{UE} is the free-space path loss between the UE and the surface, and L_{car} is the loss from a car window. Likewise, we can calculate the signal strength measured by the source gNB (M_r) as:

$$M_r = G_{\text{gNB}_n} + L_{\text{gNB}_n} + G_{\text{w,ref}} + G_{\text{gNB}_s} + L_{\text{gNB}_s} + 2 \cdot L_{\text{car}} \quad (3.4)$$

where $G_{\text{w,ref}}$ is the reflection gain of the surface. Summing Eqs. (3.2) and (3.3), we calculate M_n as

$$M_n = M_r - M_s - G_{\text{w,ref}} + 2 \cdot (G_{\text{w,tra}} + G_{\text{UE}} + L_{\text{UE}}). \quad (3.5)$$

We assume the UE-to-surface distance d_{UE} as the maximum possible distance between Wall-Street and the in-vehicle UE. This assumption is based on the fact that signals passing through the surface should reach all corners of the vehicle, particularly since all other UEs in the same vehicle are handed over simultaneously. We then calculate the path loss L_{UE} as $(4\pi d_{\text{UE}}/\lambda)^2$ where λ is a wavelength.

Since the source gNB controls the reflecting surface, it has knowledge of the surface's reflective gain $G_{\text{w,ref}}$ and transmissive gain $G_{\text{w,tra}}$. Also, as the gNB controls the transmit power of the UE in 3GPP,² it also knows the UE's transmit gain G_{UE} . By utilizing this information, the source gNB can infer the RSRP between the target cell and the UE, enabling it to make effective handover decisions.

²The gNB sends Downlink Control Information (DCI) to send power control commands to the UE.

Handover Execution

The source gNB decides whether to initiate a handover using Equations (3.1) and (3.5), and upon proceeding, sends a HO Req message to the target gNB. Upon receiving the corresponding HO ACK, the source gNB then triggers Wall-Street to switch to downlink beam combining and uplink beam splitting, and sends a HO command to the UE. Instead of detaching from the source gNB, however, the UE synchronizes with the target gNB via the RACH, using the 3GPP-standardized Dual Active Protocol Stack (DAPS) [6], maintaining two active protocol stacks. Upon receiving the HO command, the UE suspends the source cell *Signaling Radio Bearers* (SRBs), which are in charge of control-plane signaling (including Wall-Street control), and establishes SRBs for the target cell. The UE receives user data simultaneously from both source and target cells (on different channels), and we reconfigure the Packet Data Convergence Protocol (PDCP) layer to a common PDCP entity for the source and target user plane protocol stacks.

When the UE moves at vehicular speed, packets are prone to loss especially at the moment of handover when signal strength is typically weakest. So to ensure reliable packet delivery, the source and target gNBs send duplicate data packets during handover execution, and the UE's PDCP layer reorders and de-duplicates them. For packet flow in the uplink, the UE sends the same packet to the source and target gNB, which forwards the received ACK through the Xn link and dequeues the packet in the shared buffer. Wall-Street's PHY configuration remains the same for both downlink and uplink as in the preparation phase.

Handover execution at the PHY: Upon receiving a handover execution trigger from the serving gNB, Wall-Street switches between beam combining for the downlink packet flow from two gNBs and beam splitting for the uplink packet flow to two gNBs. Wall-Street simultaneously steers the downlink combination and uplink split at the same angle due to angular reciprocity. Specifically, given a certain biasing voltage configuration applied to the surface that creates a transmissive combination of two different angles for downlink signals, uplink signals impinging on the surface will be redirected toward the same angles in the

uplink. Hence, angular reciprocity facilitates fast beam alignments as the surface configuration optimized for downlink transmissions works under the uplink communication, and vice versa.

During the make-before-break, the goal of the surface is to maintain simultaneous connections with both the source and target cells until the handover completes. Initially, power is equally divided between the two links. If either link fails, we increment its allocated gain to maintain the minimum required signal strength. If this gain increase is ineffective, we revert to the even power distribution.

Handover Completion

When Eq. (3.1) holds true for a threshold *time-to-handover* duration, the target cell completes the handover by sending the source cell a **UE release** indication and triggering a signal to Wall-Street indicate handover is complete. If handover fails, the UE immediately reverts to the source cell configuration and re-activates source cell SRBs for control-plane signaling. This prevents the UE from disconnecting completely.

Handover completion at the PHY: When the source gNB triggers a signal to complete handover, the surface maximizes the power of the transmission link with the new cell by reverting to a single beam mode. If the handover fails, instead maximizes the power of the link with the old cell, maintaining the connection for the UE. This approach eliminates the need to restart communication from the initial attachment phase if handover fails.

Multi-user Support

When a handover is executed for a UE in a vehicle, other UEs sharing the same Wall-Street within that vehicle do not need to undergo a separate handover procedure. Instead, the HO Request encapsulates the contexts of all UEs associated with each Wall-Street, allowing all UEs to be handed over simultaneously using the procedure above.

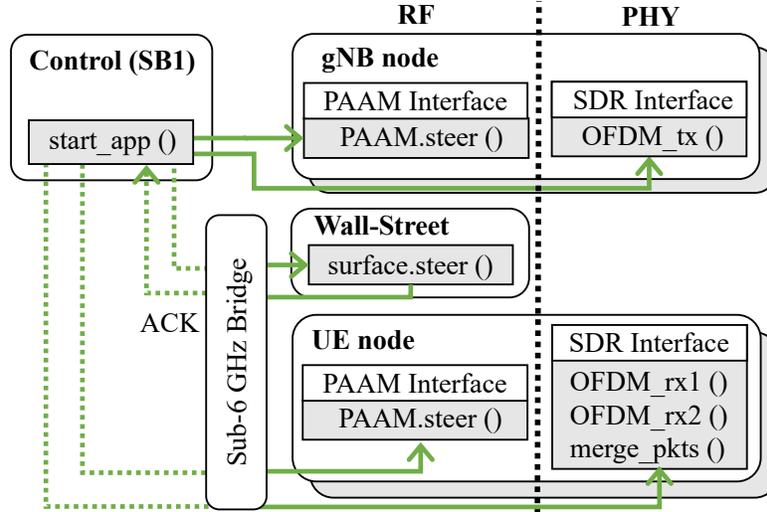
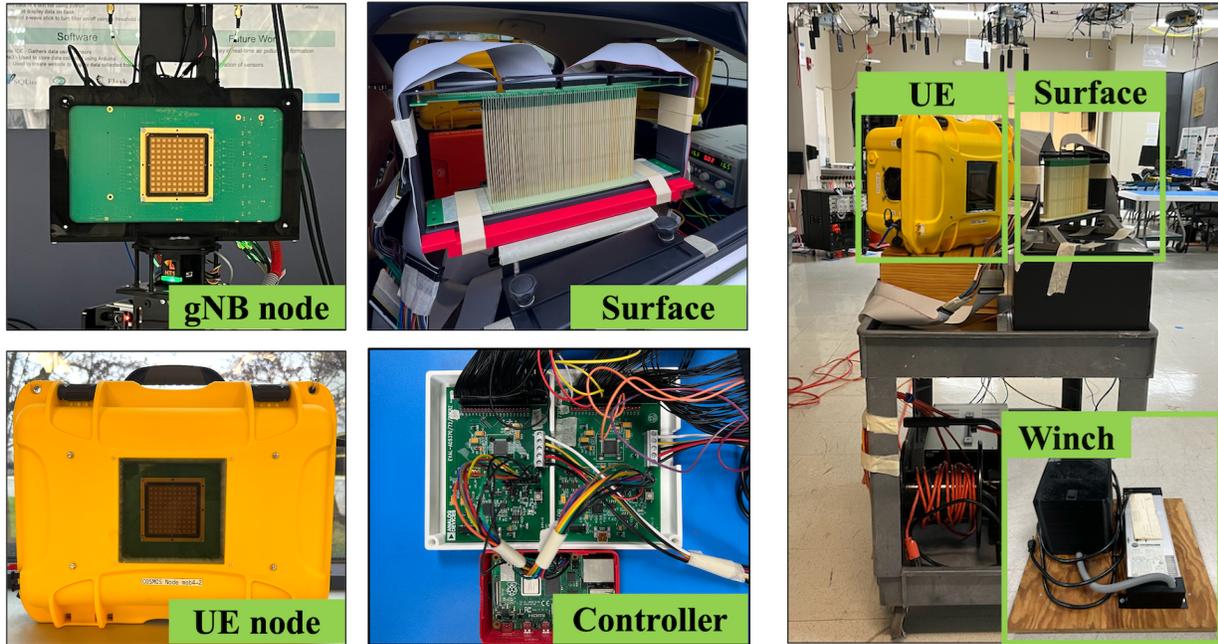


Figure 3.12: We illustrate a simplified control flow of Wall-Street-integrated COSMOS testbed.

logic relative to prior work [27]. The two DACs are each connected to the Serial Peripheral Interface (SPI) of a Raspberry Pi (RPI) through GPIO. The RPI listens for control signals from the COSMOS testbed control network, known as sandbox1 (SB1), via sub-6 GHz. Upon receiving a signal, it sends a command to the DACs, which then apply the appropriate voltage levels. These voltage levels are determined from a pre-stored voltage-to-phase look-up table. The steering speed of the Wall-Street hardware has been optimized to 0.2 ms for real-time implementation.

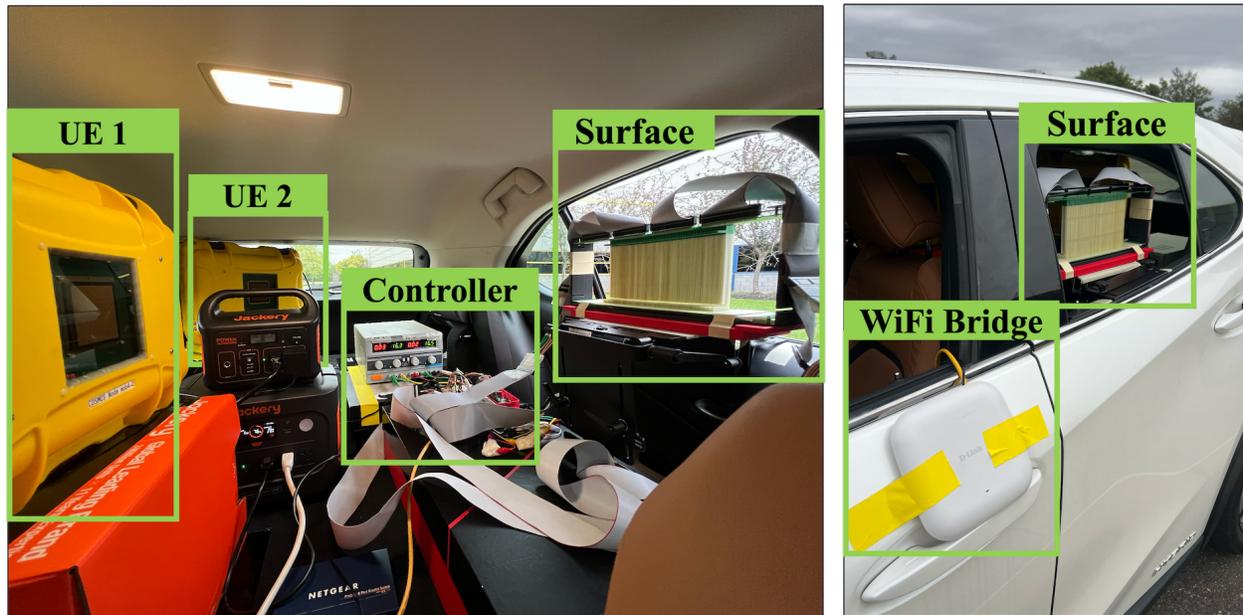
3.3.2 COSMOS Testbed

The COSMOS testbed comprises software-defined radios (USRPs) integrated with the IBM mmWave Phased-Array Antenna Module (PAAM) frontends (Fig. 3.11). The PAAM features 4 tiled RFICs, each with 32 TRx phase-shifting elements, supporting a total of 64 antennas. The PAAM offers full access to beamforming control and latency as low as 10s of nanoseconds for low-latency MAC and hybrid beamforming. As shown in Fig. 3.13(a), our experimental setup employs two types of PAAM devices: two stationary PAAMs integrated with USRP N310s (maximum 65 dB gain) mounted on remotely controllable XY tables,



(a) Wall-Street hardware

(b) Indoor setup



(c) Outdoor in-vehicle multi-UE setup

Figure 3.13: Wall-Street’s hardware implementation and testbed: showing (a) individual hardware components, (b) indoor experimental UE-side setup, and (c) outdoor in-vehicle UE-side configuration.

and three portable PAAMs integrated with USRP 2974s (maximum 31 dB gain). For indoor scenarios, we utilize four gNBs: two stationary PAAMs (gNB 1 and 4) with a 35 dB gain and two portable PAAMs (gNB 2 and 3) with a 31 dB gain. In multi-UE outdoor scenarios, we use three gNBs: two stationary PAAMs (gNB 1 and 2) with a 65 dB gain and one portable PAAM (gNB 3) with a 31 dB gain, with the remaining portable PAAMs serving as UEs. The location of UEs remain the same with and without the surface. Since all nodes must be connected to the same control network (SB1), we use a Wi-Fi bridge operating at sub-6 GHz to send the control commands to outdoor UEs and the surface.

PHY implementation on COSMOS

We developed new features on the PHY layer of the COSMOS testbed, including OFDM packet transmission and reception at one or more channels, extraction of RSRP and Packet Error Rate (PER) from the packet, and IBM PAAM and Wall-Street beam control. Periodic bursts of OFDM packets are sent from two PAAMs in different frequency bands with 3.84 MHz separation, and the receiver simultaneously decodes packets from both transmitters by operating at a larger bandwidth. Specifically, the receiver extracts RSRP from the reference signal in the preamble, the packet sequence number from the header, and the PER from the cyclic redundancy check (CRC). During handover, the source and target gNBs transmit identical packets. Wall-Street combines the downlink channels, enabling the UE to decode packets on both channels simultaneously and compute the combined PER. A C++ script in SB1 sends control commands to the gNB and Wall-Street to follow the procedure, as illustrated in Fig. 3.12.

3.4 Evaluation

We first conduct field studies to evaluate the end-to-end performance of Wall-Street and compare it with a 3GPP standardized 5G SA handover protocol (Section 3.4.2). We then

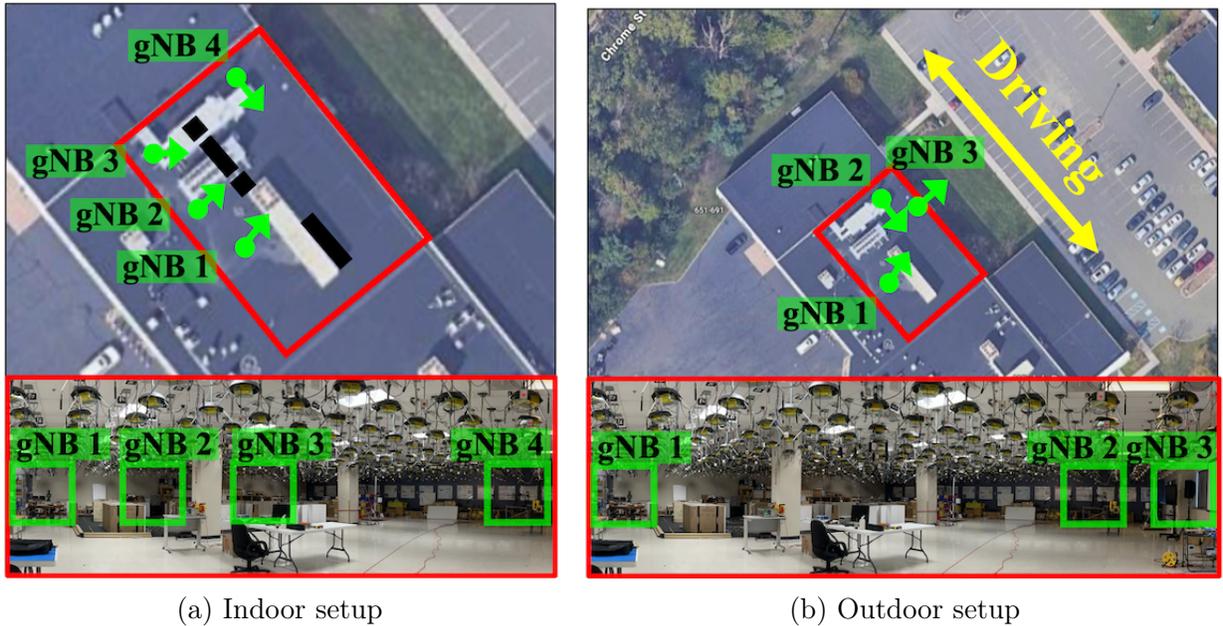


Figure 3.14: Experimental gNB-side testbed setup and map: we deploy gNBs on the first floor of a lab facing an outdoor parking lot. The vehicle is approximately 20 to 30 m from the gNBs. The arrow indicates the direction that each gNB node is facing at 0° , and each node sweeps the beam from -60° to 60° .

present micro-benchmark experiments to provide further insight into which factors impact Wall-Street’s performance (Section 3.4.3).

3.4.1 Methodology

We conduct evaluations of various indoor (single-UE) and outdoor (multi-UE) driving scenarios. For indoor settings, there are four gNB nodes and one UE node in a lab measuring 20×30 m, as shown in Fig. 3.13(b). To test the UE and Wall-Street under mobility, we place them on a cart moving in a constant speed of approximately 1 km/h. For outdoor settings, there are three gNB nodes on the first floor of a lab facing a parking lot. We mount the surface on the rear window of a SUV vehicle and deploy two UE nodes: UE 1 on the opposite side of the rear seat and UE 2 on the trunk, as shown in Fig. 3.13(c). Both UEs face the rear window. The vehicle travels at speeds ranging from 5 to 15 km/h. Both the surface and UEs are battery-powered, with UE positions remaining constant for the baseline

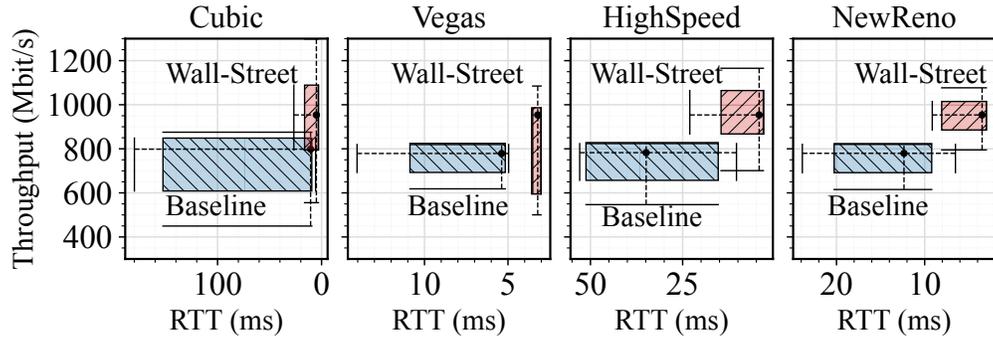
tests. Since stationary PAAMs cannot be moved, our outdoor trajectory is limited to 30 m, as depicted in Fig. 3.14.

We implement the 5G SA HO protocol described in Section 3.1 and deploy it on our testbed for comparison. We term this scheme as the baseline, using it as a performance benchmark in the remainder of our evaluation. For both Wall-Street and baseline, the hysteresis parameter H of A3 event in Eq. (3.1) is set to 10 dB with TTT of 150 ms and the measurement report periodicity of 160 ms. For the baseline, a total duration of every measurement report takes less than 20 ms with three neighbor gNBs. We set our time-to-handover to 300 ms. The outdoor testbed yielded a maximum RSRP of about 10 dB. In the ns-3 simulator, we increased all outdoor RSRP traces by 10 dB for both the baseline and Wall-Street scenarios.

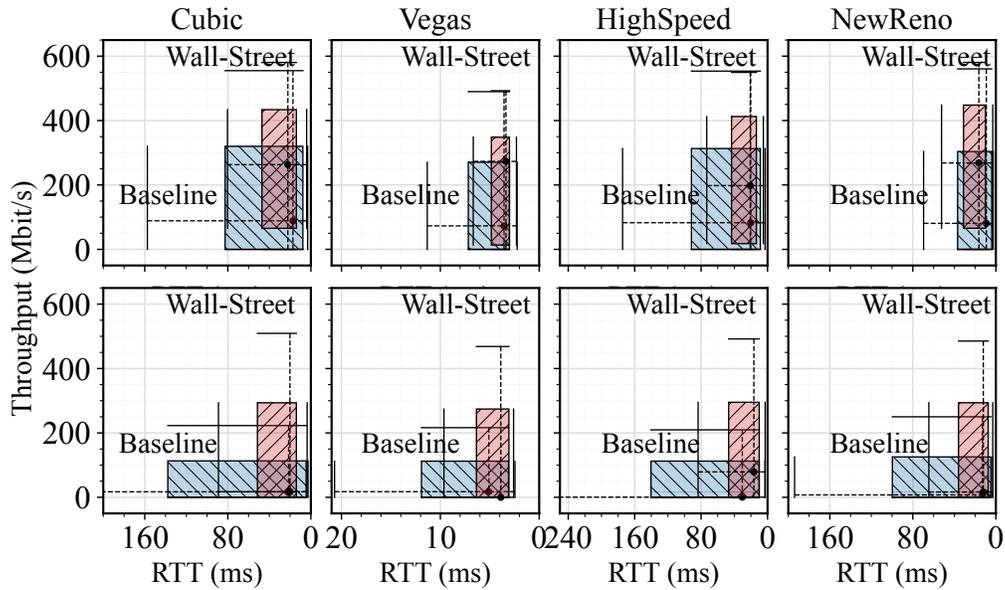
3.4.2 End-to-End Throughput and Delay

We evaluate end-to-end performance of Wall-Street delivering bulk TCP data flows. To quantify performance, we calculate throughput as the total data bits divided by the duration, measured every 100 ms. We average the round-trip-time (RTT) of packets over every 100 ms.

TCP performance under various settings. We evaluate the throughput and RTT of Wall-Street and the baseline under indoor and outdoor settings using four congestion control algorithms: CUBIC, Vegas, HighSpeed, and NewReno. In the indoor setting (Fig. 3.15(a)), Wall-Street enhances average throughput by at least 20% and reduces delay by 20-80% compared to the baseline. For the outdoor multi-user scenario (Fig. 3.15(b)), Wall-Street yields even more significant improvements: for UE 1 (upper), it doubles the average throughput and reduces RTT by 16%, while for UE 2 (lower), it increases throughput by 90% and decreases RTT by 45% on average. Notably, in the baseline setup, UE 2 (located in the trunk) experiences lower throughput and higher RTT than UE 1. Wall-Street improves signal coverage for UE 2, mitigating these issues.



(a) Indoor TCP performance with one user.



(b) Outdoor TCP performance with two users (*upper*: UE 1; *lower*: UE 2).

Figure 3.15: Throughput and round trip time (RTT) achieved by four congestion control algorithms. The right and lower edge of the box represents the 10% percentile, respectively. The left and upper edge give the 90% percentiles. The intersection point of the horizontal and vertical error bar represents the median value.

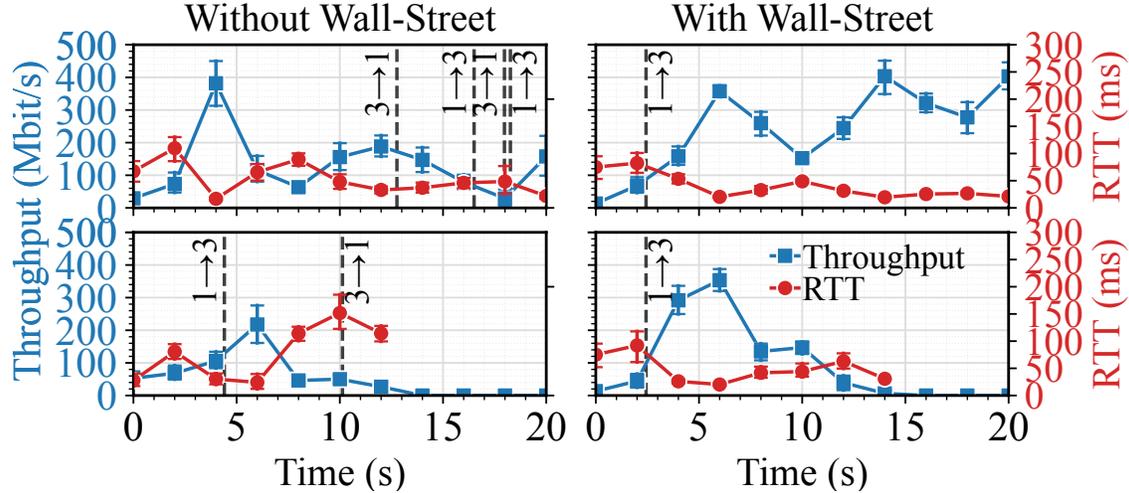


Figure 3.16: Throughput and delay of UE 1 (*upper*) and UE 2 (*lower*) achieved without Wall-Street (*left*) and with Wall-Street (*right*) as the vehicle is moving along the same trajectory. Vertical dashed lines indicate the handover event.

TCP performance over time. To further demonstrate Wall-Street’s performance, we divide the outdoor multi-user experiment into 10 equal intervals of two seconds each. Fig. 3.16 shows the median throughput and delay for UE 1 (*upper*) and UE 2 (*lower*) in each interval, comparing Wall-Street to the baseline. Vertical dashed lines indicate handover events. Wall-Street consistently outperforms the baseline in throughput throughout the experiment while demonstrating lower latency and significantly fewer handover ping-pongs between gNB 1 and 3. Notably, with Wall-Street, both UEs undergo handover simultaneously at the 2s mark, streamlining the process compared to the baseline where UE 1 and UE 2 handed over at different times. Moreover, Wall-Street maintained a stable connection throughout, whereas the baseline experienced an outage for UE 1 at 18s, immediately following a handover. These results demonstrate Wall-Street capacity to enhance throughput, reduce latency, optimize handovers, and improve connection stability. in multi-user scenarios.

Impact of vehicle speed. Figure 3.17 compares Wall-Street’s performance against the baseline for UE 1 (*upper*) and UE 2 (*lower*) at driving speeds from 5 km/h to 15 km/h. Wall-Street consistently outperforms the baseline across most metrics and speed scenarios.

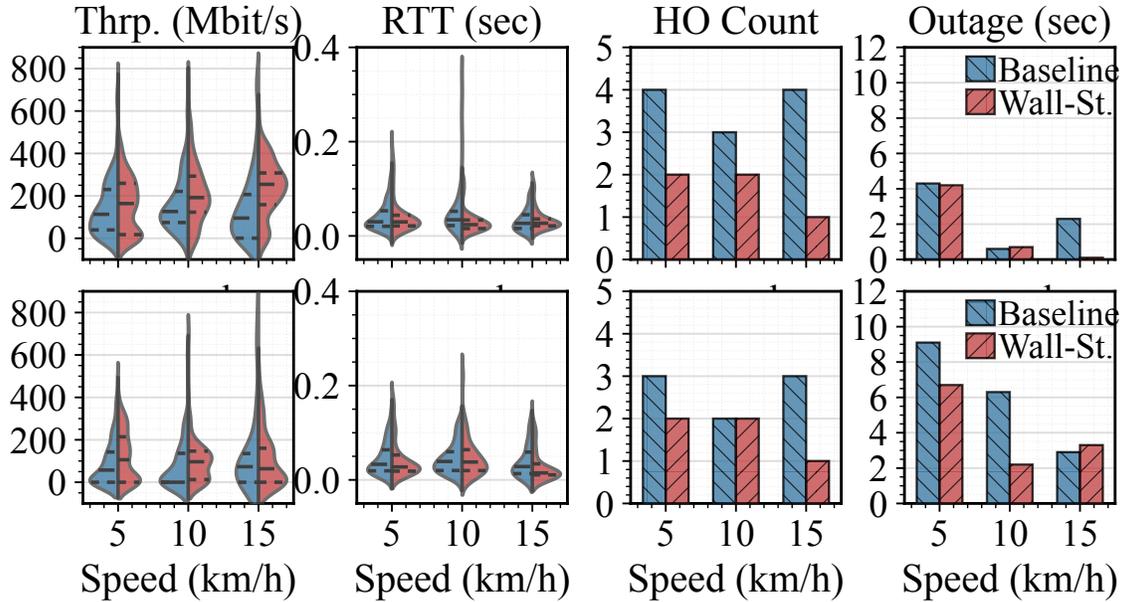


Figure 3.17: Impact of speed: throughput, RTT, HO count (*left to right*) with different speed, ranging from 5 km/h to 15 km/h.

For UE 1, Wall-Street delivers throughput gains of 13%, 30%, and 61% at 5 km/h, 10 km/h, and 15 km/h respectively, while UE 2 sees gains of 34% and 21% for 5 km/h and 10 km/h. Also, Wall-Street reduces handover frequency by an average of 53% for UE 1 and 33% for UE 2, while also reducing outage duration by an average of 27% for UE 1 and 26% for UE 2. This shows Wall-Street’s robustness and ability to maintain stable performance under varying mobility conditions. Notably, UE 2 experiences longer outage durations due to its location in the trunk, where the car body easily obstructs the signal.

Impact of handover preparation. Figure 3.18 illustrates sequence number progression during the outdoor experiment, with the left subfigure showing the entire duration and the right subfigure focusing on the handover preparation phase. To enhance visibility of trends, sequence numbers are subjected to a modulo operation. Wall-Street allows a steeper increase in sequence numbers compared to the baseline, indicating that more packets are successfully delivered. During handover preparation, the baseline’s sequence numbers show irregular progression and a lower slope due to service interruptions as the UE stops communicating with the source gNB to detect SSB bursts from neighboring gNBs. While SSB detection

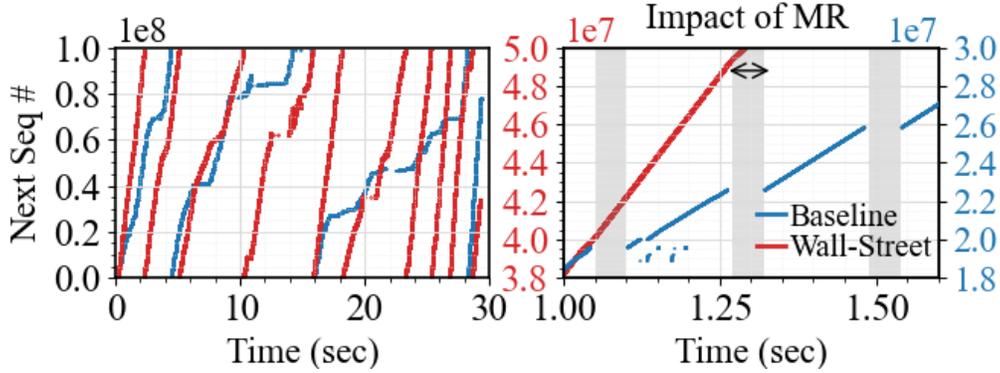


Figure 3.18: Measurement report impact: increase trend of sequence numbers during the entire 30s experiment (*left*) and handover preparation (*right*).

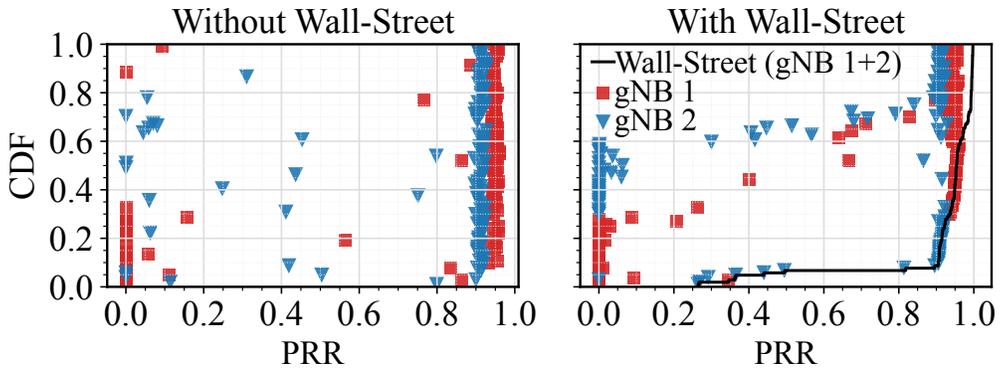


Figure 3.19: Make-before-break impact: packet reception rate of a link from each gNB without Wall-Street (*left*) and that of each gNB and of two links (packet duplication) from two gNB with Wall-Street along the same trajectory (*right*).

takes under 20 ms, its effects persist for 50 ms, significantly disrupting baseline packet delivery. In contrast, Wall-Street maintains steady sequence number progression throughout the handover preparation phase, demonstrating more stable communication.

Impact of make-before-break handover. Figure 3.19 presents the cumulative distribution function (CDF) of the packet reception rate (PRR) during soft handover. PRRs are calculated by decoding CRC for every 2000 packets and plotted on the same y-axis for both subplots. The black line represents Wall-Street’s combined PRR with packet duplication from gNB 1 and 2. In Fig. 3.19 (left), markers show PRRs for each link without Wall-Street, while Fig. 3.19 (right) shows PRRs for links through Wall-Street without packet duplication. Interestingly, individual links using Wall-Street exhibit lower PRRs compared to the base-

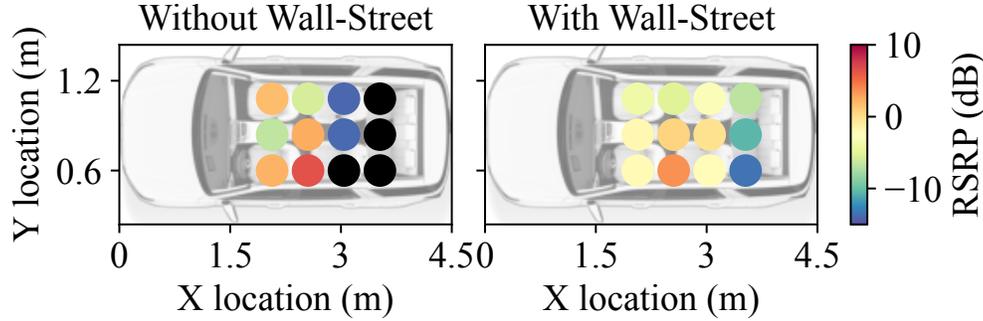


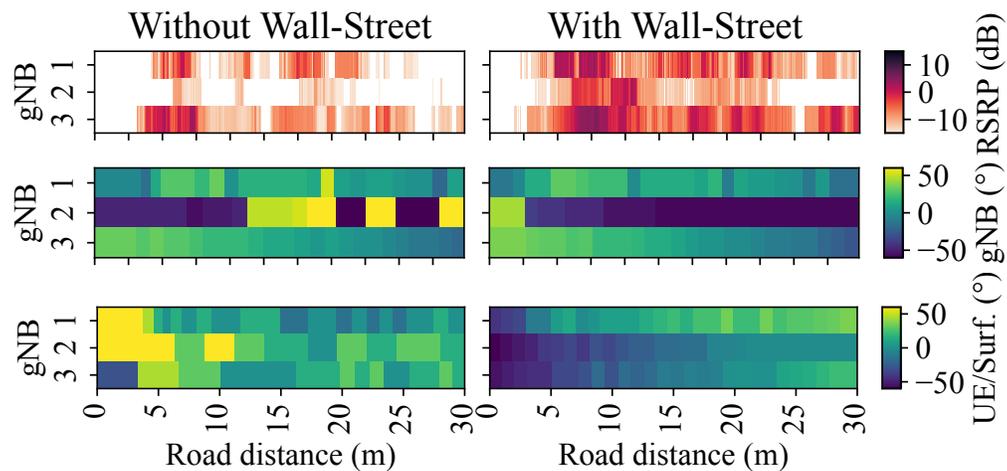
Figure 3.20: In-vehicle mmWave coverage. ●: no signal.

line. This is because Wall-Street’s beam combining approach divides power across two links, resulting in weaker signal strength per link. However, by combining the beams and sending duplicate packets, Wall-Street ensures a PRR above 90% for 92% of handover execution times and eliminates instances of packet error. In contrast, the baseline system suffers from low PRR on each link, potentially leading to packet loss and performance degradation. These results suggest that Wall-Street’s make-before-break approach enhances handover reliability, particularly during moments of weak signal strength.

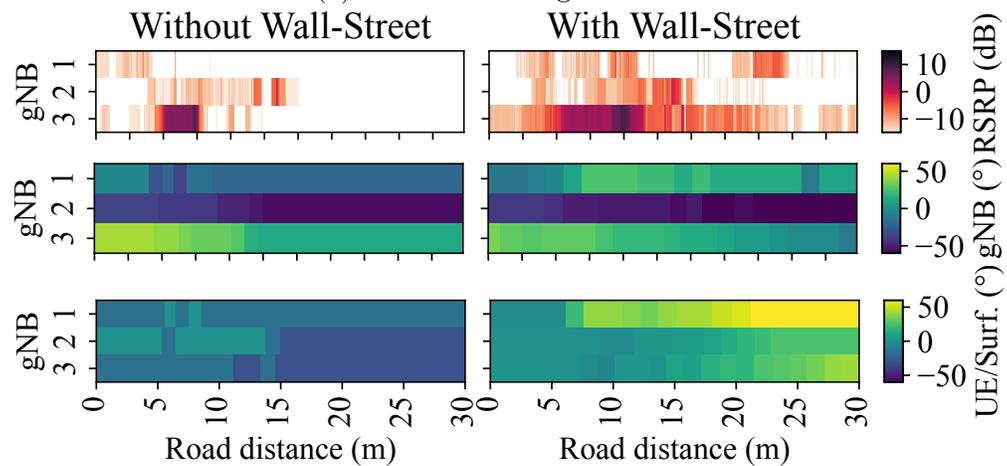
3.4.3 Microbenchmarks

We evaluate Wall-Street’s in-vehicle signal coverage, multi-beam operations at the physical layer, and far-field beam pattern measured with the spectrum analyzer.

In-vehicle mmWave coverage. To evaluate Wall-Street’s signal coverage enhancement, we conduct an experiment with the surface mounted on a vehicle’s rear window and a gNB positioned 30 meters away. We measure the maximum RSRP at various in-vehicle UE locations, exhaustively searching through combinations of gNB and UE beam angles and surface steering angles. For Wall-Street, the UE’s receiving beam is fixed towards the surface. As illustrated in Fig. 3.20, Wall-Street significantly outperforms the baseline. While the baseline experiences signal outage for UEs in the backseat area, Wall-Street ensures outage-free coverage throughout the entire vehicle interior. Furthermore, Wall-Street provides a substantial signal strength improvement with over 12 dB gain at various in-vehicle locations.

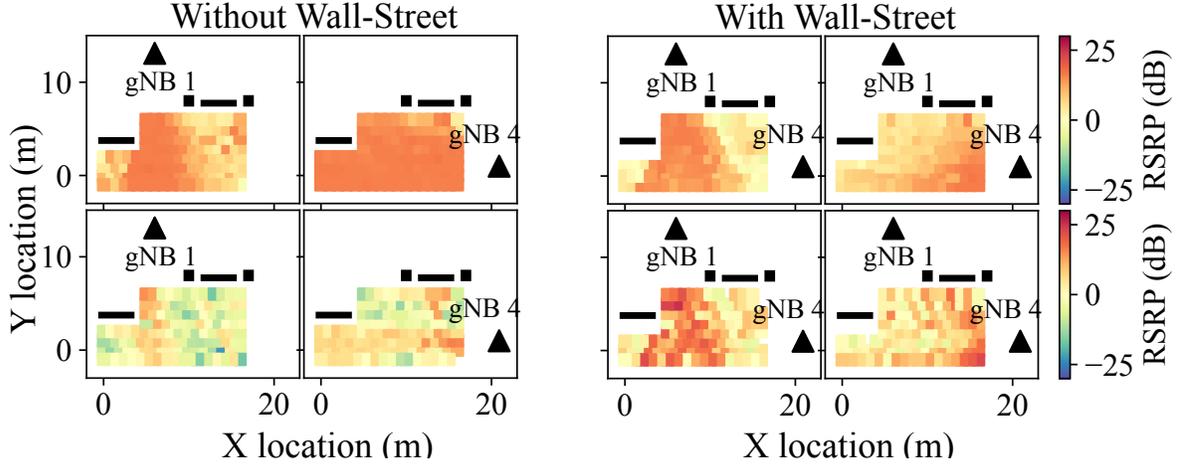


(a) Outdoor coverage for UE 1



(b) Outdoor coverage for UE 2

Figure 3.21: RSRP heatmap (*upper*) and corresponding gNB beam angles (*middle*) and UE/surface angles (*lower*) heatmaps (*left*: without Wall-Street; *right*: with Wall-Street and UE beam fixed to the surface). In-vehicle location of the UE is the same in both scenarios.



(a) *Upper*: UE aligned to the gNB; *lower*: UE not aligned to the gNB. UE fixed to 0° . (b) *Upper*: dual connection from two gNBs; *lower*: inferred MRs. UE is fixed to 0° .

Figure 3.22: Multi-link operations. We use the following notations: gNB \blacktriangle , UE \square , blockage \blacksquare .

Beam tracking under mobility. Figure 3.21 illustrates the RSRPs and corresponding beam angles for gNB and UE/surface as the vehicle traverses a 30-meter outdoor path, with separate plots for UE 1 (Fig. 3.21(a)) and UE 2 (Fig. 3.21(b)). In the Wall-Street setup, the UE angle is fixed towards the surface. The gNB angle distributions remain similar for both Wall-Street and baseline scenarios. However, Wall-Street demonstrates better signal coverage overall. This improvement is particularly evident for UE 2, which loses all gNB signals beyond the 15-meter mark in the baseline setup, whereas Wall-Street maintains signal coverage from gNB 1 and/or 3 throughout the entire path.

Multi-link operations. We evaluate Wall-Street’s capability to manipulate multiple beams at various indoor locations. Fig. 3.22 compares RSRP heatmaps with and without Wall-Street. Fig. 3.22(a) shows single-link baseline scenarios: upper heatmaps with UE beams aligned to gNB 1’s beams (left) or gNB 4’s beams (right), and lower heatmaps with gNB 1’s beams (left) or gNB 2’s beams (right) directed at the UE but UE angle fixed at 0° . Fig. 3.22(b) illustrates Wall-Street’s multi-link scenarios: upper heatmaps during make-before-break handover execution, and lower heatmaps during handover preparation. The

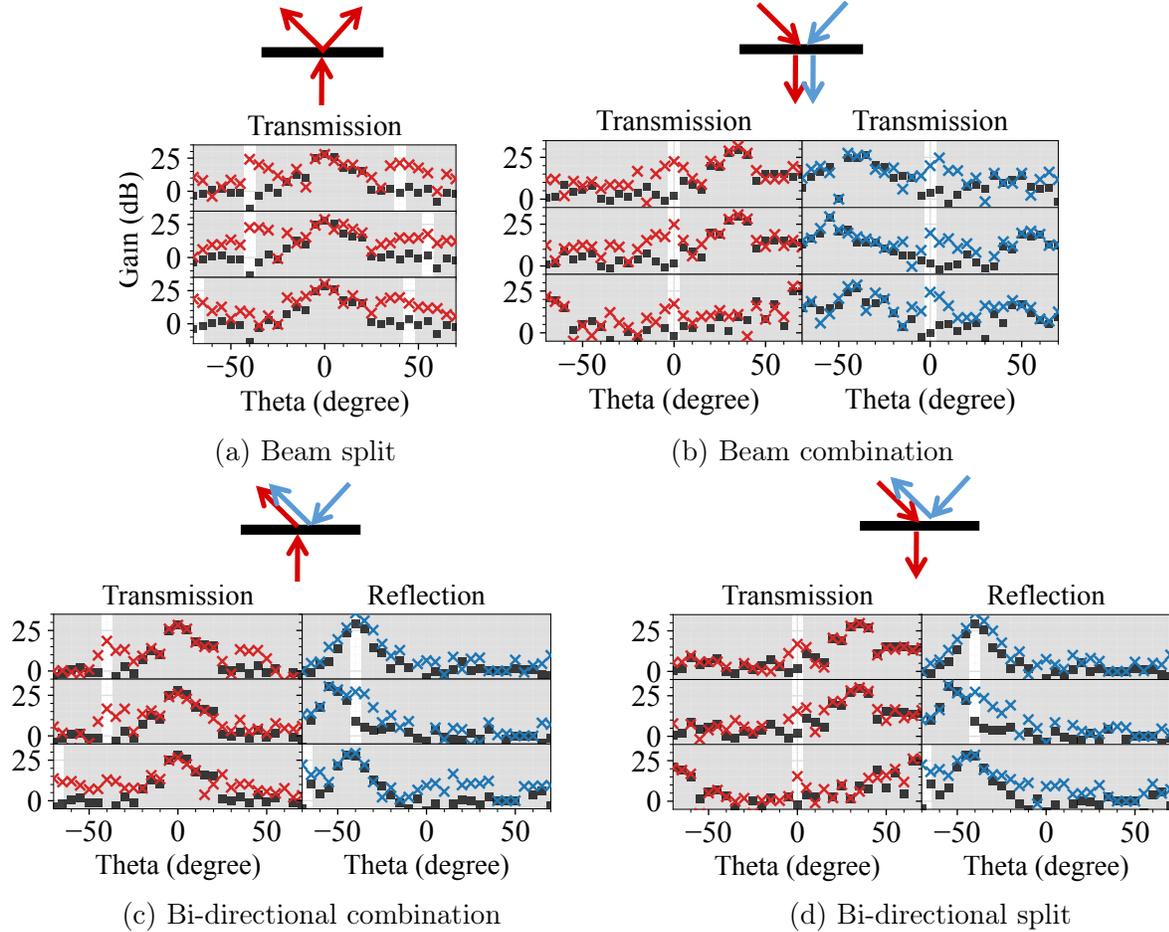


Figure 3.23: Wall-Street’s beam patterns over four operations. \times : empirical spectrum analyzer data; \blacksquare : benchmark data with the surface is set to 0V. The steering angles are highlighted in the background. 0 dB represents the spectrum analyzer noise floor (*upper*: $-40^\circ/40^\circ$; *middle*: $-40^\circ/55^\circ$; *lower*: $-68^\circ/45^\circ$). Red and blue markers are measured at 26 GHz and 26.1 GHz, respectively.

upper heatmaps in Fig. 3.22(b) show received RSRPs from gNB 1 (left) and gNB 4 (right) during the simultaneous packet reception. Due to power splitting, these RSRP values are slightly less than the aligned, single-link scenario. The lower heatmaps in Fig. 3.22(b) depict inferred RSRPs from packets received at gNB 1 (left) and gNB 4 (right) as Wall-Street reflects signals. These results demonstrate that signals reflected by Wall-Street can be accurately inferred, validating the method described in Section 3.2.3.

Far-field beam patterns. Figure 3.23 illustrates beam patterns for Wall-Street’s four operation modes: beam split, single-directional beam combination, bi-directional beam com-

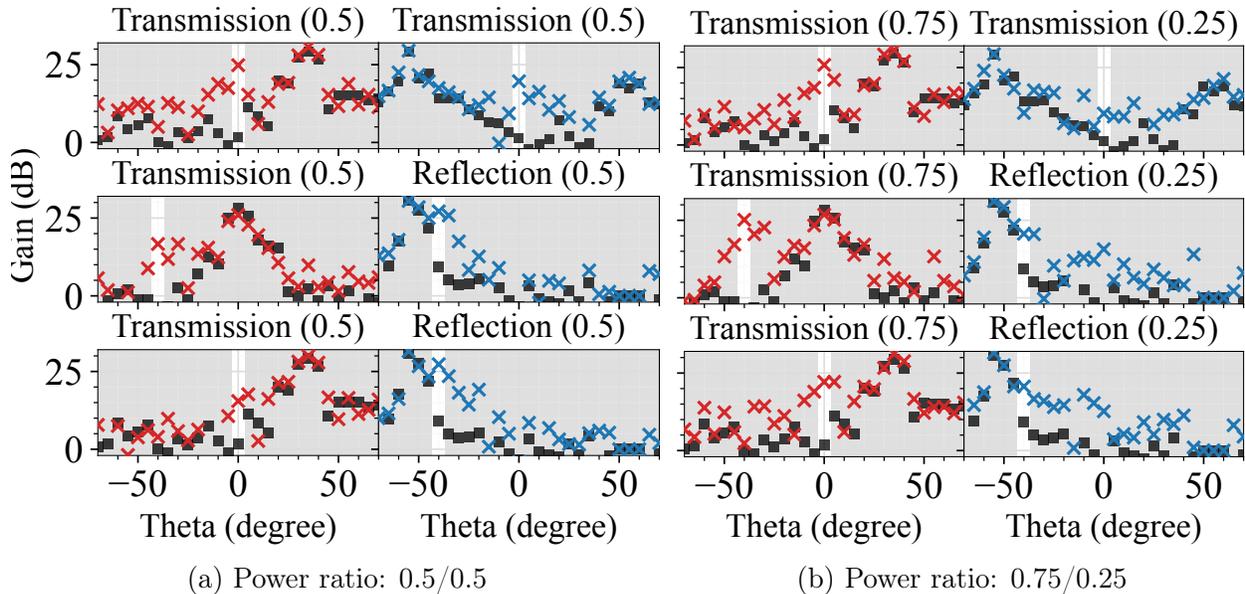


Figure 3.24: Power ratio adjustment with the steering angle of $-40^\circ/55^\circ$ (*upper*: beam combination; *middle*: bi-directional combination; *lower*: bi-directional split).

bination, and bi-directional beam splitting. We select four sets of beam angles ($-40^\circ/40^\circ$, $-40^\circ/55^\circ$, and $-68^\circ/45^\circ$), near the intersection of two gNBs' coverage areas, each with a 0.2 km radius. To measure beam patterns, we position transmit and receive horn antennas two meters from Wall-Street, recording its gain as we move the receiver from -70° to 70° .

Due to the absence of an anechoic chamber, beams passed around the surface. To ensure accurate pattern capture, we measure patterns with Wall-Street set to 0V (not steering the incident beam), using these as benchmarks (denoted by ■). We set 0 dB as the spectrum analyzer's noise floor and scale results accordingly. Wall-Street operates on two channels: left transmitter (red arrows) at 26 GHz and right transmitter (blue arrows) at 26.1 GHz.

Compared to the benchmark, we observe 20-25 dB gain in each direction, except when a steered beam aligns with the transmitter's specular reflection (*e.g.*, $-40^\circ/40^\circ$ reflection in Fig. 3.23(c) and Fig. 3.23(d)). Wider beam steering angles result in decreased gain. Identical voltage configurations in Figs. 3.23(a) and 3.23(b) demonstrate that angular reciprocity holds even under multi-beam operation, enabling rapid uplink/downlink packet flow switching during handover.

Beam power ratio. Figure 3.24 compares beam patterns with different power ratios: 0.5/0.5 and 0.75/0.25 split. Across all three multi-beam operations, shifting the ratio from 0.5/0.5 to 0.75/0.25 results in increased gain for the left beam (red markers) and decreased gain for the right beam (blue markers). This demonstrates Wall-Street’s ability to flexibly adjust power distribution between two beams.

3.5 Related Works

Prior roadside/trackside handover designs Polycorn [104] and Wi-Fi Goes to Town [126] orchestrate packet flow for (low band) LTE and Wi-Fi trackside and roadside networks, respectively, transferring packets from one cell or access point in the radio access network in order to best serve highly-mobile users. Wall-Street, in contrast, leverages the unique capabilities of smart surfaces and (high band) mmWave communications to split and combine beams to and from roadside mmWave cells in a way these prior designs do not anticipate. REM [83] uses Orthogonal Time Frequency Modulation in the LTE low band (sub-6 GHz) to simplify the conventional low-band signal strength based handover sequence, but does not target the larger frequency bandwidths and smaller cell sizes of mmWave 5G NR.

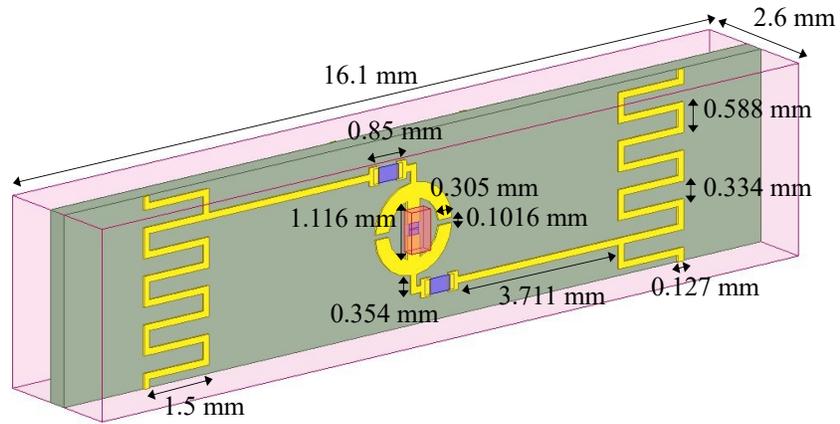
Work in mmWave beam alignment maintains a connection once it has been established via beam search. Representative works include Agile-Link [64], which proposes multi-armed beams and a postprocessing algorithm to accelerate the alignment time to be logarithmic in the number of possible beam directions. LiSteer [57] uses mmWave access point indicator LEDs and an array of mobile client-mounted light sensors to infer client bearing to the access point. BounceNet [77] proposes an architecture and algorithm for aligning many walking speed clients’ 802.11ad links with many nearby 60 GHz access points, in a relatively dense access point deployment scenario, targeting a different problem than Wall-Street.

3.6 Conclusion

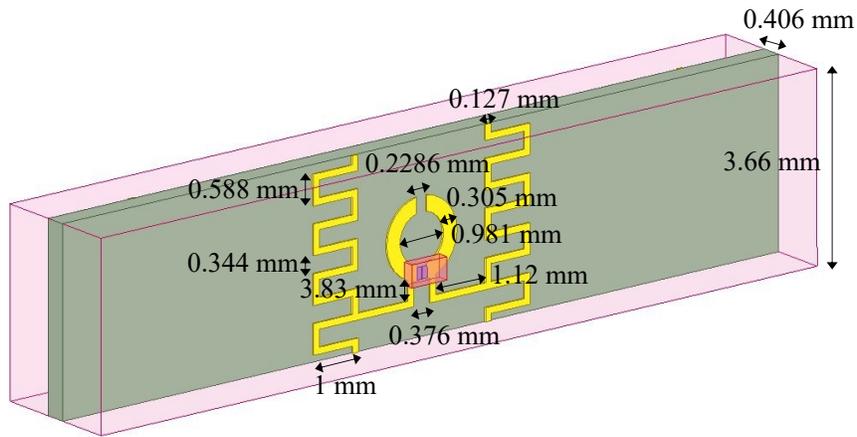
This chapter introduces Wall-Street, the first smart surface-enabled 5G mmWave roadside networking solution designed to efficiently steer outdoor mmWave signals into vehicles, enable simultaneous data exchange and neighboring cell measurements, and facilitate seamless and reliable handovers through a make-before-break process. We conduct an extensive evaluation in various indoor and outdoor settings, demonstrating significant TCP improvement. We believe Wall-Street is the first step in scaling out the capacity of mmWave networks and paves the way for future advancements in mmWave communications, particularly in vehicular networks.

3.7 Appendix

In this section, we provide design parameters of Wall-Street's electric and magnetic meta-atom, as illustrated in Fig. 3.25. The thickness of PCB Roger 4003C substrate is 0.406 mm and the thickness of the copper trace is 1 oz. To simulate the unit cell response, we create an air box that defines the boundary conditions. The top and bottom boundaries of the box are perfect electric boundary, and the front and back of the box are perfect magnetic boundary.



(a) Electric meta-atom.



(b) Magnetic meta-atom.

Figure 3.25: Wall-Street's unit cell design parameters. The pink box is an air box that defines the boundary conditions to simulate the unit cell.

Part II

Reliable and Low-Latency Satellite Networks

Chapter 4

Dual-band Reconfigurable Metasurfaces for Satellite Networking

Recently, there has been much interest in Low Earth Orbit (LEO) satellite data networking, with multiple companies' networks in various deployment phases. These networks consist of constellations of hundreds of satellites that afford advantages in latency and coverage [62]: examples include SpaceX's Starlink with a constellation of 4,425 satellites. Current systems are designed with a dish antenna that the user mounts outside the buildings, which communicates with the satellite in both the uplink and downlink directions. The dish antenna then communicates with the modem through a wire leading from the dish into the building to a modem, which then wirelessly communicates with the user, typically via Wi-Fi. While satellite networks are already deployed and seeing limited use, we believe intelligent reconfigurable surfaces will expand their applicability and improve their performance in at least the following three scenarios:

- 1. Rail/bus/airplane applications:** For best performance, transportation systems (in particular high speed rail and airplanes) will demand adaptive systems to track the satellite currently serving the vehicle as well as handoff between satellites. An electronically reconfig-

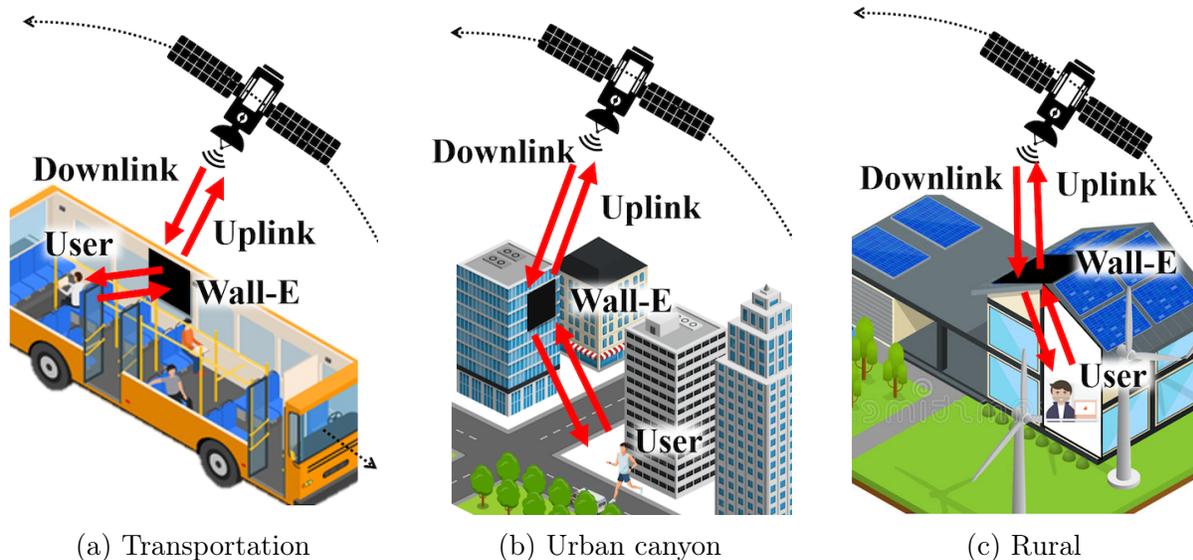


Figure 4.1: Various use cases for a satellite smart surface.

urable surface mounted on the windows and/or skylights of the vehicles can enable dynamic beam alignment to users inside.

2. Service in urban canyons: Tall buildings in a city will reduce satellite lines of sight and preclude areas of coverage at or near street level for satellite networks. While 5G/NextG wireless coverage is maximized in cities, high-frequency financial trading gains an advantage by using such networks [62] and so urban deployment remains relevant. An electronically reconfigurable surface mounted externally mid-way up a skyscraper can enable service at street level via reflection off the building, while also allowing satellite signals to transmit into the building through the surface.

3. Rural service: While current LEO satellite systems require a dish and use a gateway to forward traffic between the satellite link and client, an electronically reconfigurable surface mounted on a window/skylight can refract the satellite link into the home directly, getting rid of the outside dish.

This chapter explores innovations in the design space of LEO satellite networking with a Reconfigurable Intelligent Surface (RIS). In the process, we describe our prototype surface design, **Wall-E**, a dual-band, metamaterial-based RIS design. We first explore fundamental

Ku-band RIS design. In order to bring RIS-enhanced LEO networking to our scenarios, the surface should support both transmission (through the surface) and reflection (off the surface) modes. Huygens’ metasurfaces (HMSs) have shown to be promising in creating such *transmissive* and reflective functionalities in practice [26, 35, 87, 21, 167], thus achieving a full 360 degrees of control over radiated energy. While the basic principles of Huygens unit cells are known, designs that simultaneously make parsimonious use of electronic components (varactors and inductors), resonate at two or more different frequencies (*bi-resonant*), and achieve high efficiency are still open.

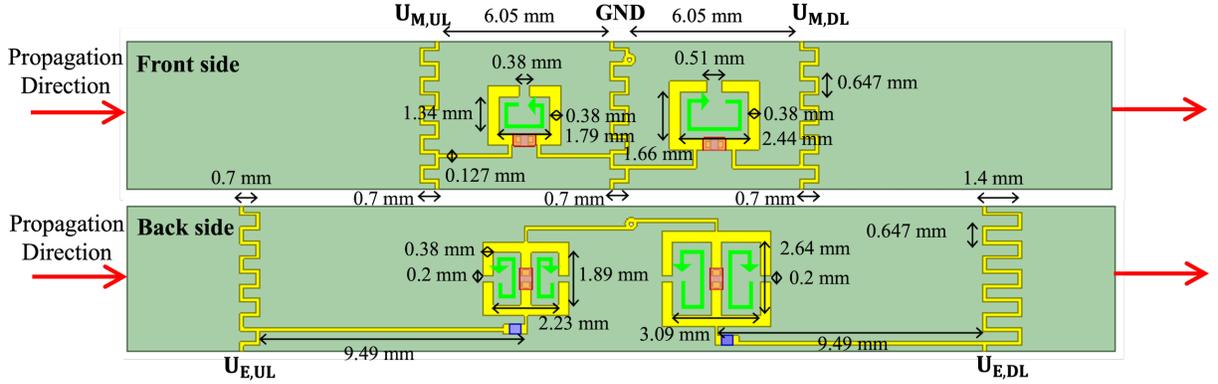
In LEO satellite networking, the process of aligning the physical wireless beam directions among user, surface, and satellite be very complex as both the satellite and user moves. Narrowing our design space to bi-resonant Huygens RIS designs, we next explore how to steer the uplink and downlink beams while preserving angular reciprocity, thus speeding the process of the beam alignment for the uplink via downlink transmissions, and vice-versa. This is of particular importance when both communication endpoints are moving rapidly, which is the case in a transportation scenario, satellite communicating with an airplane or train. In such cases, the LEO satellite network’s use of frequency duplexing division (FDD) allows for real-time, continuous feedback in both the uplink and downlink directions facilitating the constant tracking of the endpoints with respect to the RIS, and associated continuous updating of the RIS’ steering angles.

Finally, we consider starting directions for RISs to enable full end-to-end LEO network designs. We consider the handover process as the constellation of LEO satellites collectively moves over the earth, necessitating a handoff from one satellite to another, serving each user. The ability of the RIS to split uplink radio energy to two satellites and simultaneously combine downlink beam from two satellites makes a *soft* handover a possibility, which we explore further herein.

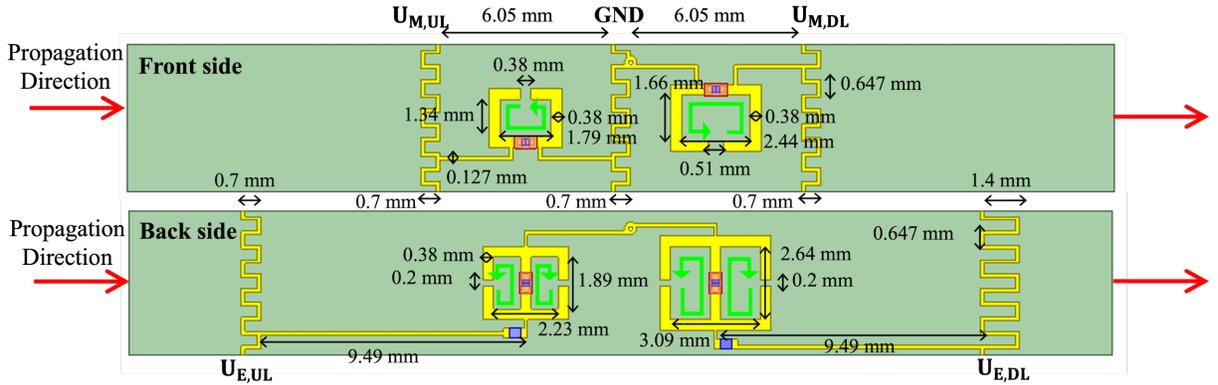
4.1 Design Challenges

By design, HMS-based surfaces consist of a layer of co-located orthogonal electric and magnetic meta-atom, facing each other across dielectric substrate, as shown in Fig. 4.7(a). The key principle is that the pair of two meta-atoms introduces a discontinuity in the impinging electromagnetic (EM) field whereby the meta-atoms manipulate field attributes, including magnitude and phase. To achieve on-demand control of the reflective/transmissive pattern, we mount a tunable, voltage-controlled electric component, known as varactor, on each meta-atom. Since varactors draw only a couple-of-hundred microwatts order of power, Wall-E consumes extremely low power. Unfortunately, HMS unit cells resonate at only one frequency (*mono-resonance*) [110] and thereby cannot act on the FDD links LEO satellite networks require.

Leveraging existing mono resonant structures for satellite networking, two alternative solutions are possible. **Strawman (i)**—building and deploying two single-band RISs (one for uplink and one for downlink). This approach would allow for FDD communication, but demands separate beam training for directional uplink and downlink, thereby doubling the overall delay of beam training. This is an important process because the satellite trajectory is not fully deterministic—it is subject to turbulence and uneven gravitational forces [100, 108]—and the terrestrial user is often mobile. Hence, the required three-party (LEO, RIS, user) beam training needs to be continuously performed for link maintenance. **Strawman (ii)**—Partitioning the surface into two subsets, each resonating at a different frequency. This approach has the advantage of link reciprocity, *i.e.*, since the downlink and uplink resonant elements are co-located, the optimum surface configuration for a downlink transmission is very close (if not exactly the same) as that of the uplink transmission. However, with such partitioning, the number of surface elements is reduced by a factor of two (given a fixed form factor) in each band. Hence, the reduced directivity gain might not be sufficient to close the long-range air-to-ground links.



(a) Design 1 (*Upper*: magnetic meta-atoms; *lower*: electric meta-atoms)



(b) Design 2 (*Upper*: magnetic meta-atoms; *lower*: electric meta-atoms)

Figure 4.2: Strawman (iii) features magnetic meta-atoms on the front side and electric meta-atoms on the back side (meta-atoms on left for uplink and meta-atoms on right for downlink). Green arrows indicate the direction of current oscillation. The design uses 1 oz copper and a Rogers 4003C dielectric substrate with a thickness of 0.813 mm. All meta-atoms are connected to a single ground meander line via through-hole vias, linking the front and back sides.

4.2 Failed Attempts: Design and Analysis

Strawman (iii) offers an alternative approach to designing dual-frequency HMS for satellite communication. Instead of partitioning the surface into two subsets of ribs (*e.g.*, strawman (ii)), each resonating at a single frequency, this design incorporates two types of unit cells within each rib. As shown in Fig. 4.2, each cell resonates at a different frequency, allowing simultaneous operation at both uplink and downlink frequencies. Unlike strawman (ii), this approach maintains the full number of surface elements for each frequency band given a

fixed form factor, thereby preserving directivity gain. The primary trade-off is that each rib becomes wider than in conventional HMS-based surfaces.

4.2.1 Meta-Atom Designs

Design 1, illustrated in Fig. 4.2(a), illustrates the detailed structure of this approach. On the front side of each rib, a smaller magnetic meta-atom operating at the uplink frequency is positioned on the left, while a larger magnetic meta-atom operating at the downlink frequency is on the right. Between these meta-atoms, a meander line on the middle is dedicated for grounding. The back side mirrors this arrangement with smaller and larger electric meta-atoms, each co-located and facing their magnetic meta-atoms across the dielectric substrate.

Voltage control in Design 1 is achieved through a series of meander lines. On the front side, separate meander lines connected to the smaller and larger magnetic meta-atoms apply voltages $U_{M,UL}$ and $U_{M,DL}$ respectively. The back side follows a similar pattern, with outer meander lines applying voltages $U_{E,UL}$ to smaller electric meta-atoms and $U_{E,DL}$ to larger ones. To minimize the number of meander lines, through-holes via connects the smaller and larger electric meta-atoms to the ground meander line on the front side across the dielectric substrate.

We denote current oscillation by green arrows in the figure. In Design 1, the current of smaller magnetic meta-atoms oscillates in the opposite direction to that of the larger magnetic meta-atoms. This is reflected in the placement of varactor diodes installed on these meta-atoms. On the other hand, the current of both smaller and larger electric meta-atoms oscillates in the same direction.

Design 2 has a small modification to this arrangement. While most elements remain the same as in Design 1, the larger magnetic meta-atoms for the downlink frequency are rotated 180 degrees. This rotation aligns the current oscillation direction for both smaller and larger magnetic meta-atoms. The electric meta-atoms, however, maintain their original configuration from Design 1.

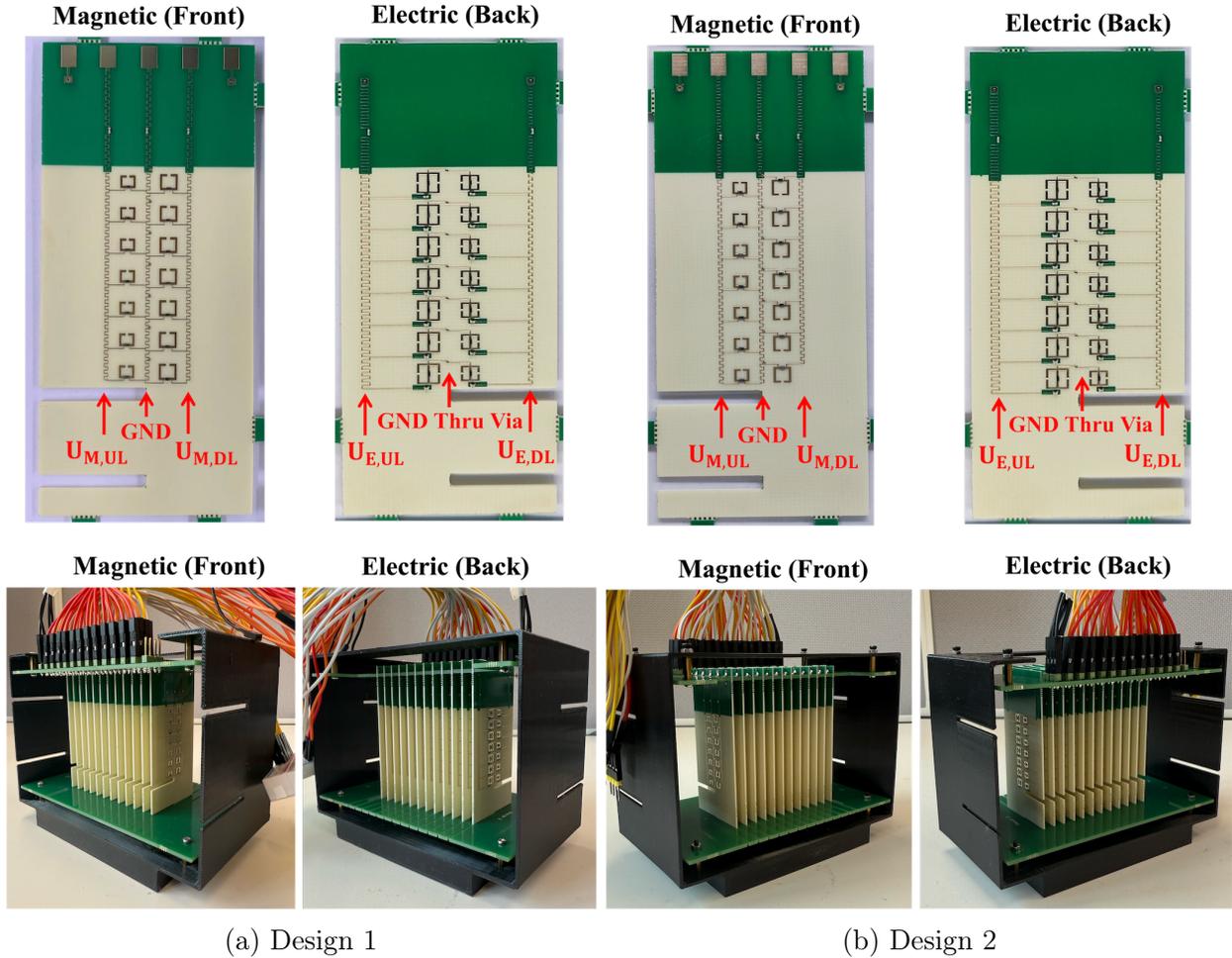


Figure 4.3: Hardware prototype of strawman (iii) design 1 and 2 (*upper*: ribs, comprised of magnetic and electric meta-atoms fabricated on a Rogers printed circuit board; *lower*: complete hardware prototypes).

4.2.2 Implementation

We have fabricated and assembled small hardware prototypes of design 1 and design 2, as illustrated in Fig. 4.3. Each rib, sized 35 by 78 mm, is made of Rogers 4003C printed circuit board (PCB) substrate. The meta-atoms have Macom MAVR-000120-1411 varactor diodes and 026011C-1N7 inductors installed. Our prototype consists of 12 ribs, each containing 14 magnetic and 14 electric meta-atoms. The ribs are held together by two perpendicular FR4 panels at the top and bottom of the structure, with the top panel also providing control lines. Each rib's control pads are soldered to the upper holder board, which connects to a DAC via

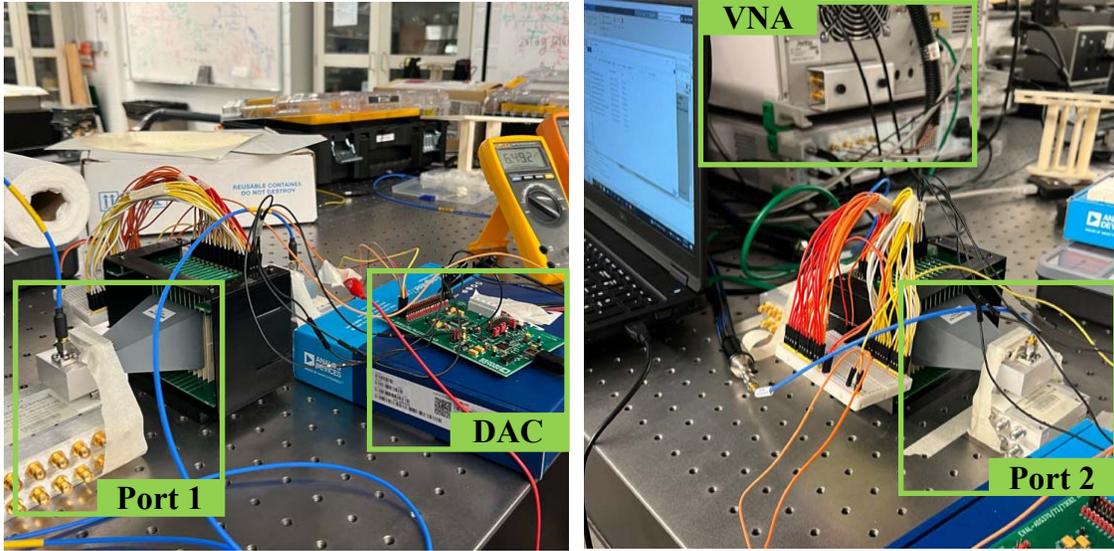


Figure 4.4: Near-field testing with two-port Anritsu MS4647B Vector Network Analyzer (VNA). We supply different voltage levels to our prototypes using an EVAL-AD5370 DAC.

microstrip traces and pin headers. The lower holder board secures the ribs in position. A 3D-printed enclosure provides standing support for the entire assembly. The ribs are spaced 5.5 mm apart, resulting in overall prototype dimensions of $84 \times 70 \times 130$ mm.

We measure the Huygens pattern through near-field experiments, collecting near-field reflection and transmission coefficients of our prototypes using a two-port Anritsu MS4647B Vector Network Analyzer (VNA), which operates from 70 kHz to 70 GHz, as shown in Fig. 4.4. To ensure accuracy, we perform a two-port calibration before data acquisition. The measurement process is automated using LabVIEW, which controls the VNA and communicates with a DAC that supplies voltage to our prototypes. During measurements, each prototype is located between two waveguide horn antennas (WR-75 standard gain horn with 15 dBi gain) connected to the VNA. These antennas are placed in close proximity to the prototype for near-field conditions.

4.2.3 Near-Field VNA Measurements

We conduct voltage sweeps, $U_{E,DL}$ and $U_{M,DL}$ for electric and magnetic meta-atoms resonating at downlink frequencies and $U_{E,UL}$ and $U_{M,UL}$ for electric and magnetic meta-atoms,

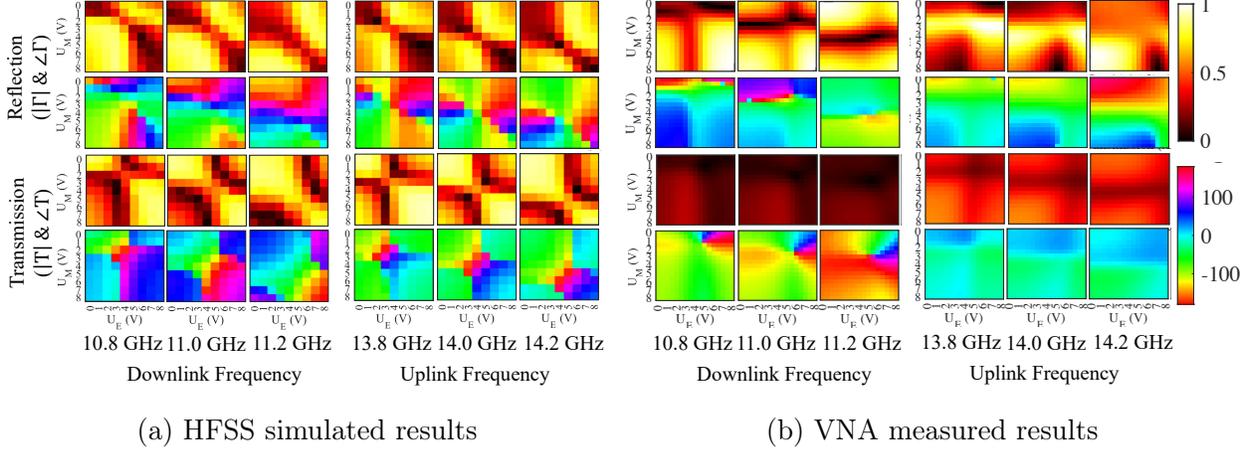


Figure 4.5: Simulated and VNA-measured Huygens pattern for Design 1 at downlink frequencies (10.8 to 11.2 GHz) and uplink frequencies (13.8 to 14.2 GHz).

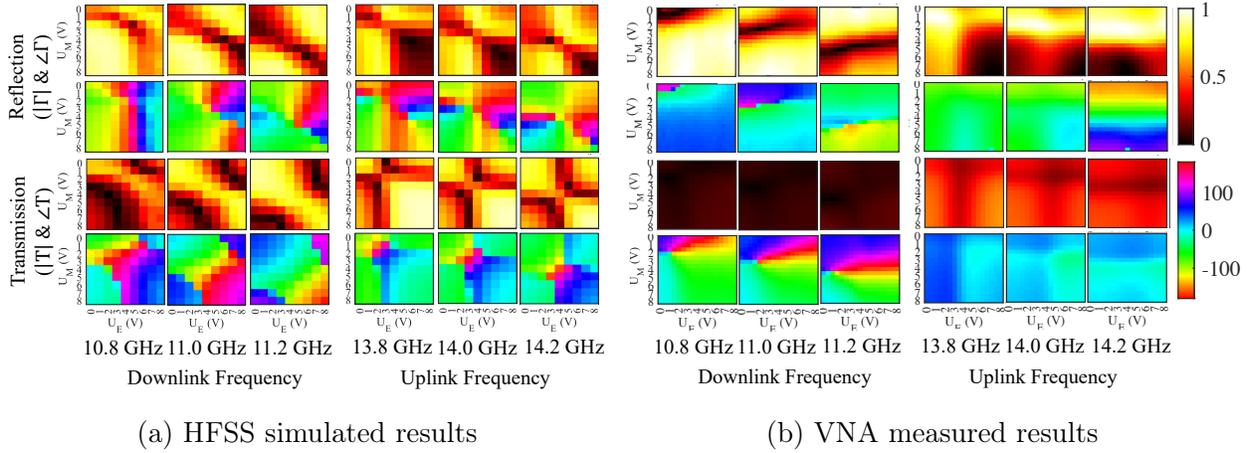


Figure 4.6: Simulated and VNA-measured Huygens pattern for Design 2 at downlink frequencies (10.8 to 11.2 GHz) and uplink frequencies (13.8 to 14.2 GHz).

resonating at uplink frequencies, to obtain transmission and reflection coefficient patterns, known as *Huygens patterns*. Figure 4.5 and Figure 4.6 present the Huygens patterns simulated with HFSS and measured from the VNA at downlink (10.8 to 11.2 GHz) and uplink (13.8 to 14.2 GHz) frequencies for designs 1 and 2, respectively. In simulations, both designs demonstrate a full transmission/reflection phase coverage of 360° with near-lossless amplitude at both downlink and uplink frequencies.

However, real-life prototype measurements deviated significantly from simulations. For design 1, at downlink frequencies, VNA-measured reflection patterns resembled simulated

transmission patterns, and vice versa. While full 360° phase coverage was achieved, the near-zero transmission amplitude indicated high loss. At uplink frequencies, no clear Huygens pattern was observed. Design 2 performed even worse, with the reflection phase at downlink frequencies failing to demonstrate full 360° phase shifts.

There are three plausible reasons for these testing failures. First, HFSS simulations might not have fully captured unwanted resonances or harmonics from the meander lines and through-hole vias. Second, as the wave propagates through the surface, it passes through two resonators in sequence. Each resonator has its own transmission loss, and going through them twice may have drastically lowered the efficiency. Third, while the smaller and larger electric meta-atoms have the same current oscillation (varactor diode direction) for both designs, having a different current direction between uplink and downlink electric meta-atoms may lead to a different result. These discrepancies between simulation and real-world performance highlight the challenges in implementing dual-frequency HMS surfaces and the need for further refinement in both design and modeling approaches.

4.3 Wall-E: Design and Analysis

We explore the key choices in our design space: we first discuss surface-enhanced LEO networking that leverages mono-resonant structures and their shortcomings in realizing a directional, highly-mobile link. Then, we explain our unique dual-band design and illustrate its key properties in fast link establishment and mobility management.

4.3.1 Building the Surface: Meta-Atoms

We now explore novel directions in the design space of the Huygens unit cell, composed of a magnetic side and an electric side, which we discuss in turn. Figure 4.7(a) illustrates a magnetic meta-atom structure that operates only in one frequency band. Here, the magnetic field of an incident EM wave induces a rotating current (denoted by green arrows) within

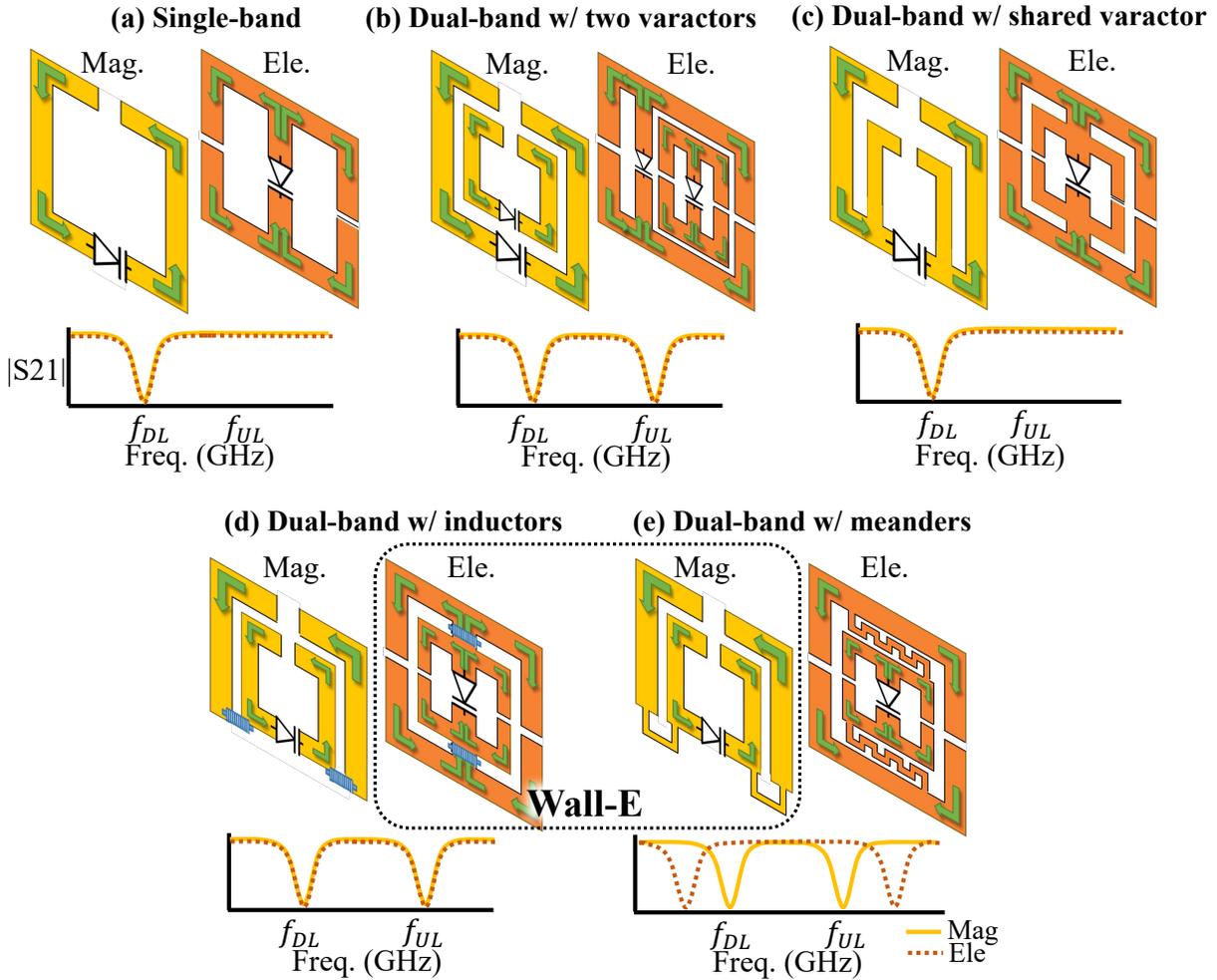


Figure 4.7: The magnetic and electric meta-atom design considerations (*top*: design schematics where the magnetic and electric meta-atom are colored in yellow and orange, respectively; *bottom*: transmission responses in magnitude $|S_{21}|$). The electric and magnetic meta-atom inside a dotted black line are the designs selected for Wall-E.

the metallic loop (colored in yellow), which in turn produces its own magnetic field. To manipulate the field response, the meta-atom is integrated with a varactor diode, a voltage-dependent capacitor. The magnetic meta-atom is in essence a resonator consisting of both inductance and capacitance. Hence, the resonance response can be controlled via a varactor. Thus, a naïve approach to enabling bi-resonant unit cells would be to include two co-located metal rings (the inner ring optimized for the higher uplink frequency of 15 GHz, and the outer optimized for the lower downlink frequency of 10 GHz), as shown in Fig. 4.7(b). Although simple, this approach would require two separate varactors, increasing cost, insertion loss, and biasing complexity.

Instead, we want to control both outer and inner rings simultaneously, using a single shared varactor (Fig. 4.7(c)). However, we find that with such a structure, only the outer ring oscillates—thereby, the meta-atom effectively operates at only a single frequency. In order to allow the passage of low frequency signals into the inner ring, we load the connecting part with two RF chokes, which blocks the signal at higher frequency (for the inner ring) and passes the signal at lower frequency (for the outer ring). We can design the choke in two ways: mount coil inductors (Fig. 4.7(d)) or by bridging the outer and inner rings with a thin *meander line* copper trace (Fig. 4.7(e)). By adjusting meander width and length, we can apply a proper inductance value to choke off signals. Since coil inductors increase insertion loss, we finally choose Fig. 4.7(e) as our magnetic side design candidate.

Figure 4.7(a) also shows the electric side, resonating in one frequency band only. The electric field of an incident wave induces a rotating current within the metallic loop (colored in orange), which in turn produces its own electric field. Similar to the magnetic meta-atom, Fig. 4.7(c) shows two electric meta-atoms with a shared varactor. To properly control two rings using one varactor, we again connect two rings with RF chokes. As shown in Fig. 4.7(d), the outer ring has oscillating currents at a lower frequency (downlink), and the inner ring has its own oscillating current at a higher frequency (uplink). However, unlike the magnetic meta-atom, we do not use meander traces as a RF choke for the electric meta-

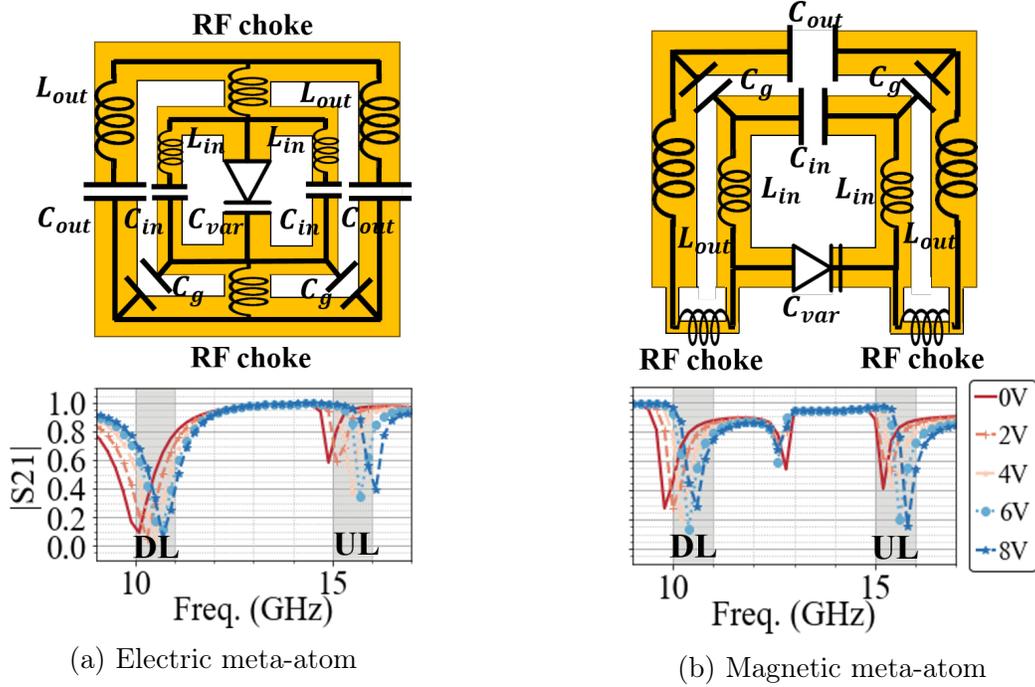


Figure 4.8: Equivalent circuit and its transmission response in magnitude across frequencies and voltages. The downlink and uplink frequency regions are colored in grey.

atom, because the meander trace would need to be placed in the gap between the two rings due to the different structure of the electric meta-atom. Increasing this gap, however, would create a huge frequency difference between two rings as shown in the transmission response of Fig. 4.7(e). Hence, we select the Fig. 4.7(d) as our preferred electric side candidate.

Equivalent Circuit. Fig. 4.8 illustrates the candidate design’s equivalent circuit diagram, with the corresponding *magnitude of transmission* coefficient $|S_{21}|$, across different frequencies, and across different applied varactor control voltages. By definition, HMS currents oscillate at a resonant frequency $f = 1/(2\pi\sqrt{LC})$ where L is the inductance and C is the capacitance of the meta-atom. In Fig. 4.8, we see that each electric and magnetic meta-atom operates at two resonant frequencies, one at downlink and another at uplink. The resonant frequency for the downlink is largely affected by the outer ring’s inductance L_{out} and capacitance C_{out} while the inner ring’s inductance L_{in} and capacitance C_{in} mainly determines the uplink’s resonant frequency. By increasing the voltage to the varactor, we decrease the total

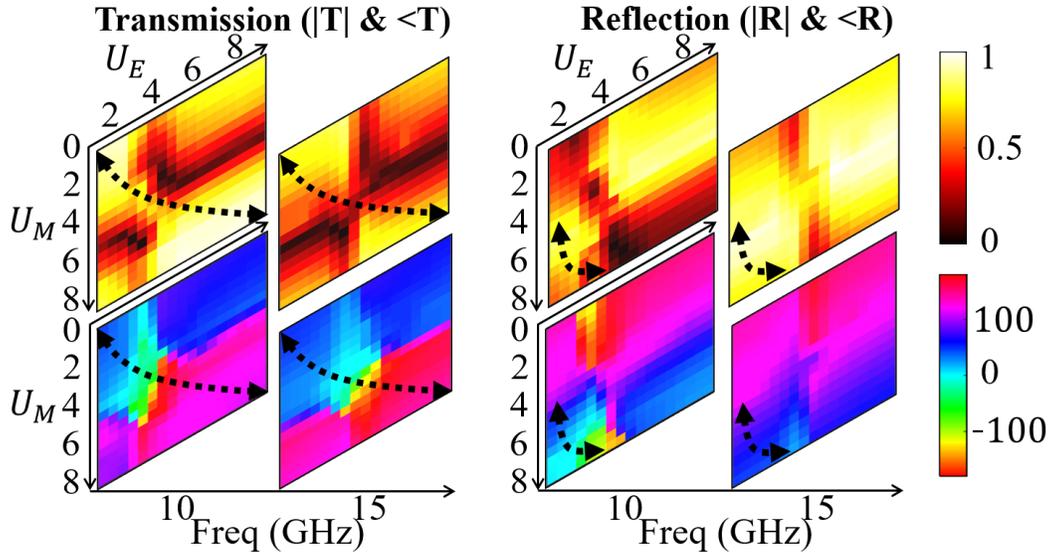


Figure 4.9: Huygen’s transmissive and reflective pattern in magnitude and phase at downlink and uplink frequency with different voltages applied to electric meta-atom U_E and magnetic meta-atom U_M . The path denoted by the black dotted curve shows 360°, high amplitude phase coverage.

capacitance of the meta-atom, which, in turn, shifts the resonant frequencies, meaning that on each side, we can control both the outer and inner rings with just a single varactor.

Huygen’s Pattern. When we place the electric and magnetic meta-atoms together as shown in Fig. 4.10 and sweep the voltage across two varactors U_E and U_M , we obtain the transmission and reflection coefficient pattern, so called *Huygen’s pattern* as depicted in Fig. 4.9. This pattern demonstrates a full transmission/reflection phase coverage of 360° with near-lossless amplitude on the area marked by the black dotted curve. While a single-band HMS obtains the Huygen’s pattern at only one frequency, our design achieves this at *both* frequencies, enabling bi-directional control of an FDD signal.

4.3.2 Establishing a Surface-Satellite Link

Owing to the mobility of the LEO satellite as well as the end users, beam alignment plays a key role in maintaining the link quality of mobile satellite communication networks. We note that the coarse trajectory of the satellite is known a priori (and hence can be incorporated in

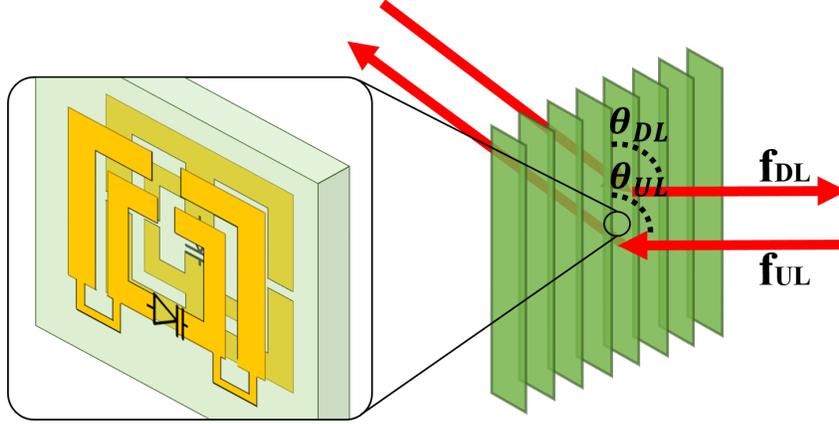


Figure 4.10: Wall-E’s bi-directional beam-steering in FDD communication. Due to its angular reciprocity, the steering angle of downlink θ_{DL} is equal to the steering angle of uplink θ_{UL} .

beam adaptation protocols); yet, the exact real-time location of satellite cannot be perfectly predicted due to the numerous factors like turbulence and uneven gravitational forces [100, 108]. More importantly, the end point user is often mobile adding to the complexity of the three-party beam search between the satellite, user, and the surface. Conventional beam alignment protocols implement a trial-and-error scheme and test different potential directions sequentially. Extending such schemes to surface-enhanced satellite networks yield an increased delay as the beam training should be repeated at two different spectral bands. In fact, [53] demonstrates that a simultaneous uplink and downlink beamforming design in RIS-assisted FDD systems achieves more than 1.4 times transmission rate over a one-way beamforming design [165, 90, 69].

On the other hand, Wall-E can simultaneously steer the downlink and uplink beams at the same angle due to *angular reciprocity*. Specifically, assume a certain biasing voltage configuration applied to the surface such that creates a transmissive steering angle of θ_{DL} for the incident downlink signal at 10 GHz, as shown in Fig. 4.10. Due to the angle reciprocity, an uplink signal impinging the surface at θ_{UL} will be redirected toward the satellite location. Hence, angular reciprocity facilitates fast beam alignments in FDD satellite networks

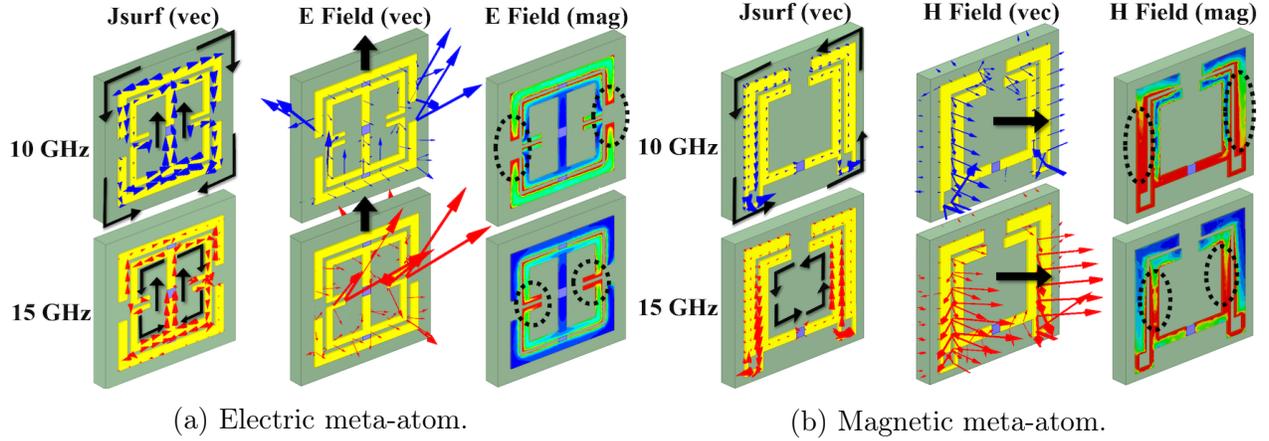


Figure 4.11: An electric meta-atom’s surface currents and electric fields and a magnetic meta-atom’s surface currents and magnetic fields at 10 GHz and 15 GHz.

as the surface configuration optimized for downlink transmissions works under the uplink communication and vice versa.

4.3.3 Enhancing Satellite-Satellite Handover

The fast movement of LEO satellites (around 7.5 km/s velocity relative to a reference point on the ground [17]) can cause multiple handovers resulting in an increase of RTT and a significant throughput drop. Even though other access networks (such as cellular networks) also experience handover, the impact of handovers on the transport layer and quality of service is relatively small, because of their relatively shorter RTT and thereby faster link recovery. We argue that an RIS-enhanced satellite network can substantially alleviate this problem. In particular, Wall-E supports soft handovers by allowing two (or multiple) satellites impinge on the surface at the same time. By carefully choosing the voltage configuration at each meta-atom, Wall-E achieves beam combining and steering. In this case, as the primary satellite fades away due to mobility, the secondary satellite will ensure a non-interrupted link. We highlight that such flexible handovers is owed to the on-demand waveform engineering at Wall-E.

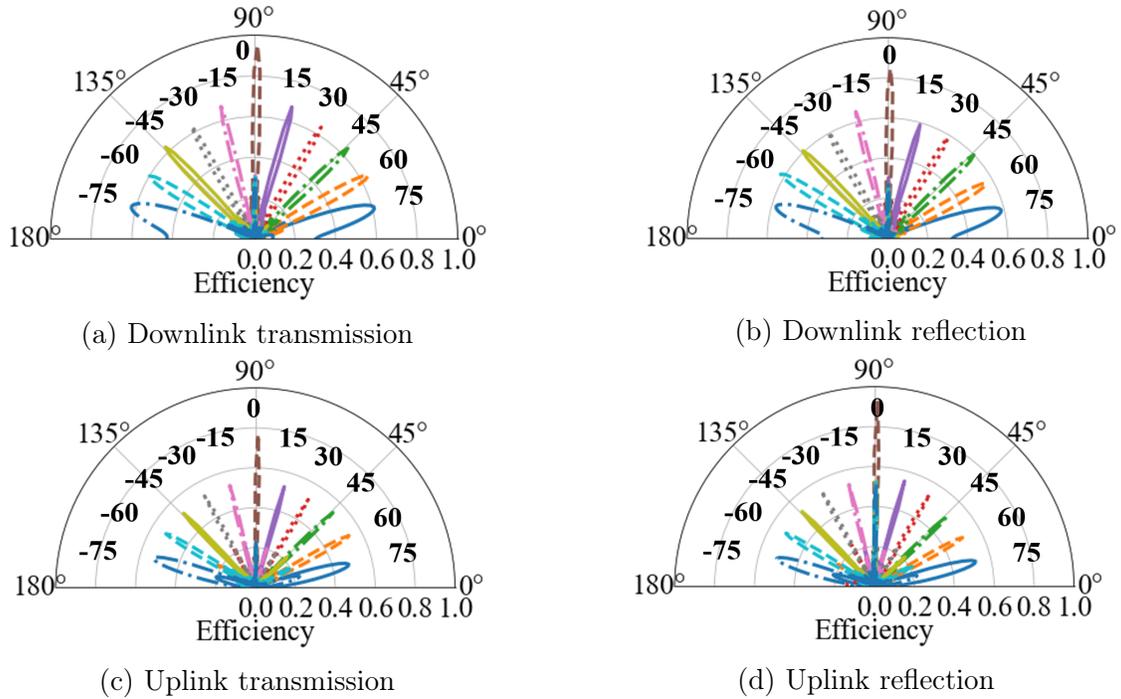


Figure 4.12: The transmission and reflection efficiency as Wall-E steers the downlink and uplink beam.

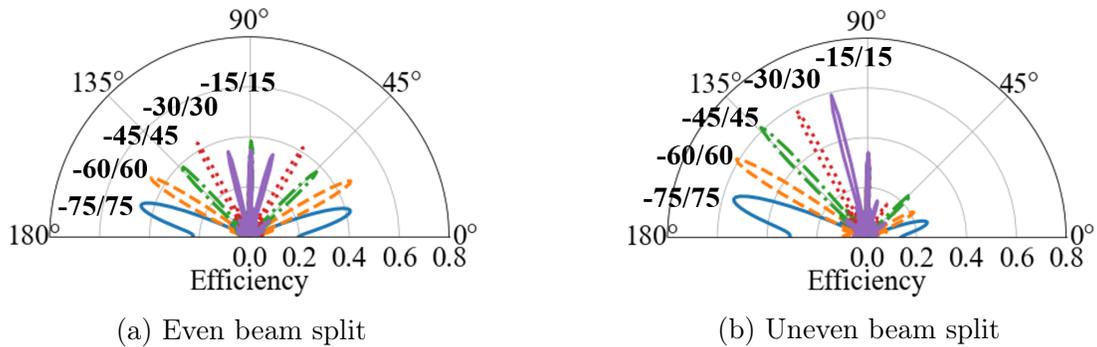


Figure 4.13: The transmission efficiency of a beam splitted by Wall-E. The power split is even for (a) and uneven for (b).

4.4 Feasibility

To project the feasibility of Wall-E, we simulate its performance with HFSS simulation. We also model our varactor based on its Simulation Program with Integrated Circuit Emphasis (SPICE). In the future, we will fabricate and implement Wall-E and experiment with actual satellite signals.

Near-Fields. Figure 4.11 illustrates the electric meta-atom’s surface currents and electric fields (Fig. 4.11(a)) along with the magnetic meta-atom’s surface currents and magnetic fields (Fig. 4.11(b)) at 10 GHz and 15 GHz. For both electric and magnetic meta-atom, the surface currents J_{surf} oscillate on the outer ring at 10 GHz while they oscillate on the inner ring at 15 GHz. We denote the direction of J_{surf} in black arrows, which conforms to Fig. 4.7. Similarly, the fields are excited by the outer ring in 10 GHz by the inner ring in 15 GHz. Fig. 4.11 confirms the bi-resonate nature of the Wall-E meta-atoms.

Radiation Efficiency. In this section, we demonstrate a high efficiency of Wall-E. Specifically, we calculate the efficiency as a magnitude of an array factor at a desired angle: $AF = a_0 + a_1 e^{jk d \cos \theta} + \dots + a_{N-1} e^{jk(N-1) d \cos \theta}$ where $k = 2\pi/\lambda$, d is a meta-atom spacing, θ is a steering angle, and a is a complex value chosen from Fig. 4.9. Fig. 4.12 shows the efficiency of Wall-E as it steers the beam with the step of 15-degree. Specifically, Fig. 4.12(a) demonstrates the efficiency of downlink transmission, which ranges from 62% to 94%. Similarly, Fig. 4.12(b) illustrates the downlink efficiency as Wall-E reflects the beam, which ranges from 60% to 85%. Fig. 4.12(c) and Fig. 4.12(d) reveal the efficiency of 50% to 80% for uplink transmission and reflection. Moreover, we highlight Wall-E’s beam-split performance in Fig. 4.13 for a soft-handover. Here, we split the beam 150, 120, 90, 60, and 30 degrees apart. Specifically, the power is evenly divided for each beam on Fig. 4.13(a), and it is unevenly splitted for Fig. 4.13(b) (1/3 on left and 2/3 on right). The result demonstrates that Wall-E can tailor the beams in a flexible manner, which enables a highly-efficient relay and hand-over in FDD communication.

Link Budget. In this section, we analyze our back of the envelope calculation for closing a 1,150-km air-to-ground link. We formulate a link budget in decibel as follow:

$$P_{rx} = P_{tx} + L_{d_1} + L_{window} + G_{WallE,Rx} + L_{d_2} + G_{WallE,Tx} + G_{rx} \quad (4.1)$$

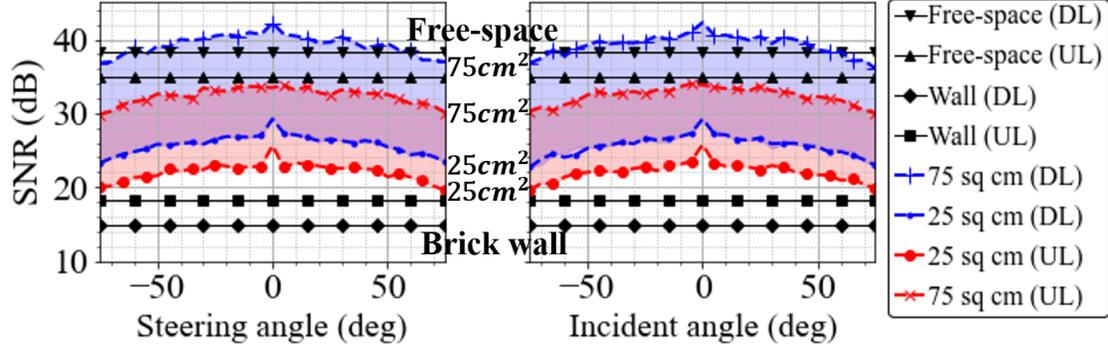


Figure 4.14: Simulated SNRs of transmissive links with varying surface size, steering and incident angle compared to SNRs of a free space path and wall penetration without Wall-E.

where P_{tx} is a transmit power, including the transmitter gain. The Maximum Transmit EIRP of a most powerful satellite is 66.89 dBW, which is equivalent to 97 dBm [11]. We assume that the transmit power is 97 dBm for downlink. L_{d_1} is a free-space path loss between the satellite and Wall-E. Since an orbital height from Earth is approximately 1150 km [11], the free-space path loss L_{d_1} is -173.7 dB for downlink and -176.6 dB for uplink. L_{window} is a -4 dB loss of window where Wall-E is placed. Assuming 5 m distance between Wall-E and user, L_{d_2} is -66.4 dB. G_{rx} is the receiving gain, equivalent to the gain of the user in downlink. We assume that G_{rx} is 25 dB. Lastly, $G_{WallE,Rx}$ and $G_{WallE,Tx}$ is Wall-E's Rx and Tx gain, respectively. Each is calculated based on the effective aperture, $A_e = \frac{\lambda^2}{4\pi}G$. Specifically, the surface gain $G_{WallE} = a_\theta 4\pi A_e / \lambda^2$ where a_θ is the radiation efficiency of Wall-E at a steered angle θ . Finally, we obtain SNRs in decibel by subtracting the noise power from the signal power.

Figure 4.14 demonstrates the simulated SNR as Wall-E steers the beam in two scenarios. The first scenario has the incident beam perpendicular to the surface, and the surface steers from -75 to 75 degrees. On the other hand, the second scenario varies the angle of the incident beam from -75 to 75 degrees while Wall-E steers the beam in a perpendicular direction. For each scenario, we vary the size of Wall-E and compare the simulated results against the free-space path and brick wall blockage in the absence of Wall-E. For both scenarios, the larger the surface is, the higher the SNR is. In particular, the SNR of $75cm^2$

sized Wall-E is higher than the SNR of the free-space path for over a 100 degree field of view. Compared to the brick wall blockage scenario, 75cm^2 sized Wall-E provides approximately 24 dB higher SNR.

4.5 Related Works and Discussion

Dual-Band Metasurfaces. Dual-band metasurfaces have recently gained attention however the existing architecture fall short in meeting at least one of our requirements, namely, flexible reconfiguration, transmissive/reflective modes, and 360-degree coverage. [61] introduces a dual-band metasurface for S- and C-bands, which provides wide-band operation with high transmission efficiency. However, the surface lacks dynamic configuration which makes it inapplicable in our highly mobile application. [120] proposes a dual-band tunable metasurface that operates in C- and Ku- bands using PIN diodes. However, this design supports only the reflection mode; thereby, it's not suitable for through-wall applications. Further, PIN diodes limit the phase shifting resolution and hence the steering efficiency. In contrast, [118] employs varactors to achieve a continuous phase control. Unfortunately, this design is also reflection-only and is limited in angular coverage. To the best of our knowledge, Wall-E is the first design of a dual-band reflective/transmissive reconfigurable metasurface with a 360° phase coverage and high efficiency.

Besides the surface design itself, recent works [96, 171, 18] have developed a system model for intelligent reflecting surface (IRSs)-assisted channels and showed that a IRS can improve the link budget of ground-to-satellite links by more than 6 dB. While existing works only focus on modeling reflective surfaces, our paper demonstrates the link budget improvement for both reflective and transmissive RISs.

Reflectarray Antennas. Prior works [41, 40, 112, 73] have proposed the use of reflectarray antennas for space communication, where the reflectarray is placed on the satellite and is excited via the feed horn. Such an architecture can realize flexible steering as the

reflected signal can be dynamically steered according to the array configuration. Further, [107, 95] explored the multibeam reflectarrays for the multispot coverage from the satellite. Metasurfaces and reflectarrays are both spatially-fed structure composed of small elements. However, we propose using a metasurface as an intermediate node (hence not co-located with the satellite nor end user) to increase path diversity. In doing so, Wall-E’s capability to support both transmission and reflection plays a key role.

Wall-E for Satellite Link Aggregation. While gaining a lot of attention, the bandwidth of a single satellite path is unlikely to provide low-latency links comparable with a fiber path [62]. However, aggregating abundant paths from many satellites within coverage zone of a user can, in principle, offer lower latency than a fiber path. We highlight that Wall-E can play a crucial role in realizing satellite link aggregation by combining and steering the signals from multiple satellites into a desired direction. In the future, we will explore novel scheduling algorithms for optimum coordination between multiple parties (all nearby LEOs, RIS, and user) and extend the prior efforts [143, 144, 45] on link satellite scheduling that do not address multi-satellite and RIS-enhanced networking.

4.6 Conclusion

Wall-E is an electronically tunable surface programs dual-frequency radio at 10 and 15 GHz, refracting the satellite link directly into homes and to satellites. This work shows that Wall-E design ensures angular reciprocity, allowing simultaneous steering of downlink and uplink beams at the same angle, thus accelerating the beam alignment process.

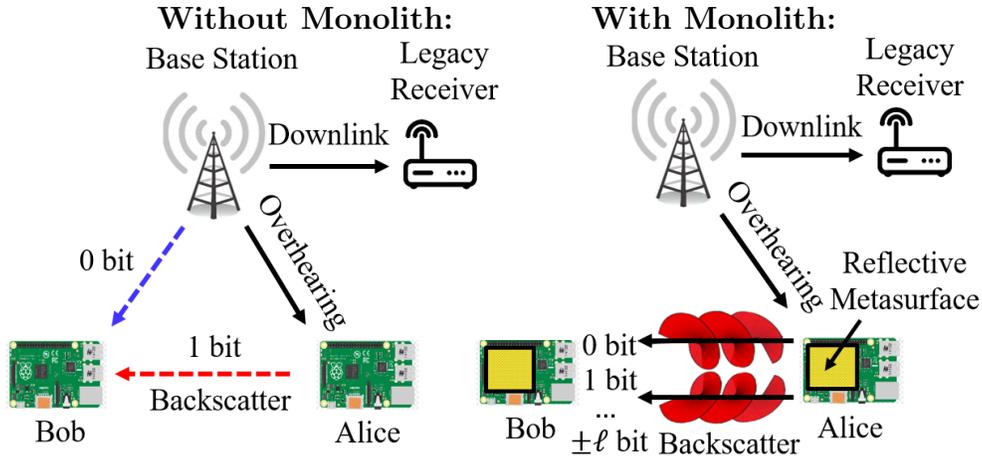
Chapter 5

Metasurface-Enabled Spatial

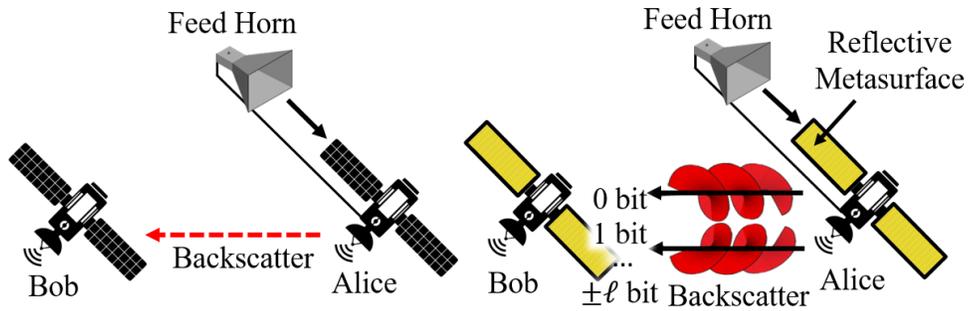
Multiplexing for Rank-Deficient Satellite

Networks

The multiple-input multiple-output (MIMO) system enables a plethora of applications that require high traffic volume, such as 5G services in urban canyons and indoor VR/AR, by establishing independent parallel channels between multiple transmit and receive antennas. However, the capacity offered by spatial MIMO systems largely relies on multipath components, and so the practical applicability of such systems is challenging in many *rank-deficient* application scenarios such as farm fields in smart agriculture and satellite channels in space communication. Further issues surrounding the power consumption of such MIMO systems, including that of power-hungry radio frequency (RF) chains whose power requirements scale proportionally with the number of antennas, constrain the applicability of spatial MIMO systems even further. Today's emerging wireless applications in rank-deficient environments, such as mega-constellation satellites [62], demand high bandwidth, implying that power availability at edge devices is becoming a more critical issue which ultimately will restrict such applications from adopting spatial MIMO systems.



(a) Ambient backscatter system for edge devices.



(b) Reflectarray antenna for satellite communication.

Figure 5.1: Various use cases for Monolith. The sender (Alice) backscatters the signals, and the reader (Bob) decodes the signals. With Monolith, a metasurface artificially creates multiple orthogonal beams and spatially multiplexes them.

To address these issues, we introduce **Monolith**, a reconfigurable metamaterial surface that backscatters the signal into several multiplexed orthogonal beams. Each of the multiplexed beams carry a “unique” orbital angular momentum (OAM) state which enables orthogonality for beams with different OAM states. OAM state can be equivalently thought of as the number of phase rotations applied to the incoming signal, and so a beam with non-zero OAM state takes on a helical shape in the direction of propagation. Since Monolith itself generates multiple orthogonal beams, it eliminates the need of rich scattering environ-

ments for spatial multiplexing. We believe Monolith will expand the applicability of MIMO in the following rank-deficient application scenarios:

1. Ambient backscattering for low-power devices Ambient backscattering [89, 31] offers ultra-low energy consumption by enabling sensory devices to pick up an ambient signal, harvest it as the source of power, and then reflect it with encoded data, such that these devices can communicate without generating and transmitting their own signals. Eliminating the need for active transmitter and power-consuming RF components significantly lowers the energy burden and therefore enables edge devices in rural and remote areas to operate on a smaller battery or even a batteryless device. While state-of-the-art ambient backscattering systems are extremely low bit-rate, Monolith enhances the link capacity of ambient backscattering systems by adopting OAM spatial multiplexing. In conventional ambient backscatter networks (Fig. 5.1(a) on left), the sender indicates either a ‘0’ or a ‘1’ bit by switching its antenna between reflecting and non-reflecting states. In Monolith (Fig. 5.1(a) on right), on the other hand, the sender converts the ambient signal into several multiplexed orthogonal beams wherein each beam supports a different data bit. The receiver later separates the multiplexed beams.

2. Reflectarray antenna system for satellites To achieve a high gain, low profile, and light-weight communication system, satellites deploy a reflectarray antenna that consists of a feed horn and a planar structure of reflecting elements [40, 112, 95]. Specifically, the horn excites the array, which in turn steers the reflected beam according to its configuration. To maximize the link capacity of satellites, Monolith utilizes an OAM metasurface as the reflecting antenna array. As shown in Fig. 5.1(b), the surface transforms the signals from the horn into multiple orthogonal beams and spatial multiplexes them on the space. We note that the use of extremely high frequency in satellite communication [78] mitigates the divergence of OAM beams, rendering OAM-MIMO to be more practical and deployable. We will further discuss the beam divergence in §5.3.

In this chapter, we present an in-depth analysis on OAM multiplexing and demonstrate that OAM-based MIMO systems achieve significantly higher wireless capacity than the traditional MIMO system in rank-deficient wireless environments. Our further studies elaborate on Monolith’s system architecture, ranging from its surface design to its encoding and decoding schemes, along with simulated results that project the feasibility of Monolith.

5.1 Primer: Orbital Angular Momentum

From Maxwell’s theory, it is known that electromagnetic (EM) waves carry both linear momentum and angular momentum. The angular momentum has two independent attributes: a spin angular momentum (SAM) associated with the polarization of the EM wave, and an orbital angular momentum (OAM) associated with the spatial distribution of the EM wave fields [163]. In particular, OAM beams carry a twisted/helical phase front that rotates as the beam propagates. This helical phase can be manipulated by generating phase twists across multiple transmitting antennas. Let us define the number of phase twists as a *state*. By increasing the number of phase twists, OAM can theoretically have “infinite” states, each corresponding to a distinct topological charge [151]. Moreover, OAM beams with different states are mutually orthogonal in spatial domain, thus offering a possibility of unlimited, parallel channels for data transmission. Figure 5.2 illustrates an example metasurface that reflects and converts a Gaussian beam with no OAM into multiplexed OAM beams with five different OAM states ($l = +2, +1, 0, -1, \text{ or } -2$) along with the phase and intensity profile of each OAM beam. The intensity nulls at the center, and it expands as the absolute value of OAM state increases.

In this section, we address a recent controversy over OAM and investigate the advantages of OAM-based multiplexing over the conventional MIMO multiplexing.

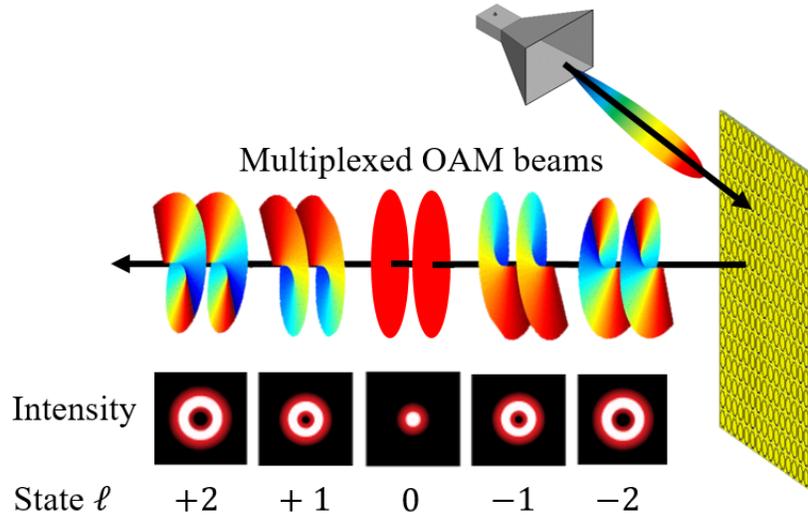


Figure 5.2: Mixed-state OAM generation using Monolith.

5.1.1 Is OAM a New Concept?

After Tamburini *et al.* [132] argued that the OAM is a new degree of freedom that allows an unlimited use of wireless channels on the same frequency, there was controversy over whether OAM enables a new and distinct physical degree of freedom [132, 131] or is simply a subset of MIMO [130, 38]. OAM beams with different states are intrinsically orthogonal to each other, providing a new opportunity to spatially multiplex different beams along the same physical paths. Since MIMO is a general technique dealing with the use of multiple antennas and multipath propagation, OAM is in fact a special form of MIMO [38] and does not provide a new degree of freedom. Unlike traditional MIMO, OAM does not need a large number of unrelated paths for spatial reuse (and indeed, cannot leverage physical paths reflecting off objects in the environment) because different OAM beams themselves are orthogonal. Therefore, OAM-MIMO communication fills the need for spatial multiplexing in the absence of a rich scattering environment, while conventional MIMO retains its utility in the presence of a rich scattering environment.

Earlier studies [38] on OAM generation utilize a uniform circular array (UCA) where each adjacent array element adopts excitation feed with equal amplitude and $2\pi l/N$ phase

difference. However, a UCA requires a large number of antennas to generate high-order OAM state and has a large beam divergence angle. Also, it requires power-hungry phase shifters, which is not suitable for low-power devices. Very recently, artificially-engineered surfaces, called *meta-surfaces*, have been studied to manipulate the EM wave properties [26, 23]. Compared to UCAs, metasurfaces have advantages of low power, high gain, and flexible capability of modulating EM waves. Recently, Ref.[85] proposed a 1-bit reconfigurable metasurface to generate different OAM beams by loading the PIN diode (*i.e.* on-off switch) on each element of metasurface. Similarly, [55, 60] demonstrated state-reconfigurable OAM metasurfaces using varactor diodes for finer phase control. In this paper, we explore varactor-loaded metasurfaces for a better control of OAM phase profiles.

5.1.2 Understanding the Benefits of OAM

In this section, we explain metasurface-based OAM multiplexing's capacity gains using standard MIMO theory. We demonstrate that OAM-based MIMO communication achieves a significant capacity gain over traditional MIMO communication when the rank of channel is low.

Channel Model In free-space MIMO systems, the channel matrix is expressed as $\mathbf{H}_{\text{MIMO}} = [h_{mn}^{\text{MIMO}}]_{M \times N}$, where h_{mn}^{MIMO} is channel response between n^{th} transmitting antenna and m^{th} receiving antenna. It takes the form:

$$h_{mn}^{\text{MIMO}} = \beta \frac{\lambda}{4\pi d_{mn}} e^{-j \frac{2\pi}{\lambda} d_{mn}} \quad (5.1)$$

where β is a constant containing attenuation and phase associated with the antenna system configuration, λ is the carrier wavelength, and d_{mn} is the distance the signal propagates. Similarly, in OAM-based MIMO systems, the channel matrix is expressed as $\mathbf{H}_{\text{OAM}} = [h_{mn}^{\text{OAM}}]_{M \times N}$, where h_{mn}^{OAM} can be modeled as MIMO channel response superimposed by a spiral phase fac-

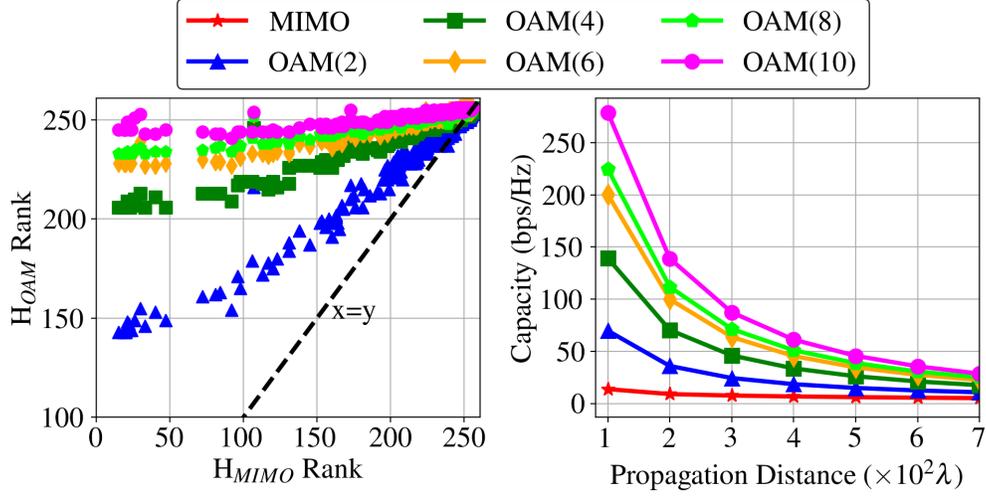


Figure 5.3: *Left*: Channel rank enhancement by OAM. *Right*: Capacity comparison between OAM and MIMO.

tor. It takes the form:

$$h_{mn}^{\text{OAM}} = h_{mn}^{\text{MIMO}} \cdot e^{j\tilde{l}_n \tilde{\varphi}_{mn,l}} \quad (5.2)$$

where \tilde{l}_n contains all Q number of OAM states generated by the n^{th} transmit antenna, and $\tilde{\varphi}_{mn,l}$ denotes the cumulative phase of the l -labeled OAM state after propagation from n^{th} transmit to m^{th} receive antenna.

Our Scenario In our system, the transmitting and receiving elements are *meta-atoms*, and they are arranged in a uniform rectangular array pattern on a substrate we refer to as a *metasurface* (details are provided in Sec. 5.2, Fig. 5.4 depicts this scenario for a 16×16 metasurface system). Let us assume that both the transmitting and receiving metasurfaces consist equal number of meta-atoms. Then, the propagation distance ($d_{(p,q),(s,t)}$) between the $(s,t)^{\text{th}}$ transmit and $(p,q)^{\text{th}}$ receive meta-atoms is:

$$d_{(p,q),(s,t)} = \sqrt{D^2 + \Delta^2 \{(p-s)^2 + (q-t)^2\}} \quad (5.3)$$

where D is the distance between the center of transmitting and receiving metasurfaces, and Δ is the spacing between consecutive meta-atoms. (p,q) and (s,t) indicate the (row,

column) locations of the receive and transmit meta-atoms in their respective metasurfaces. The cumulative OAM phase factor corresponding to these meta-atoms is then:

$$\tilde{\varphi}_{(p,q),(s,t),l} = \phi_{(p,q),(s,t)} + \frac{2\pi}{\lambda} \tilde{l}_{(s,t)} D \quad (5.4)$$

where $\phi_{(p,q),(s,t)}$ is the azimuthal angle between the $(s,t)^{th}$ transmit meta-atom and $(p,q)^{th}$ receive meta-atom along the axis of propagation, which depends only on the relative locations of these meta-atoms. It can be calculated as [94]:

$$\phi_{(p,q),(s,t)} = \begin{cases} \tan^{-1}(|p-s|/|q-t|), & p > s, q > t \\ \pi - \tan^{-1}(|p-s|/|q-t|), & p > s, q < t \\ \pi + \tan^{-1}(|p-s|/|q-t|), & p < s, q < t \\ 2\pi - \tan^{-1}(|p-s|/|q-t|), & p < s, q > t \\ 0, & p = s, q \geq t \\ \pi/2, & p > s, q = t \\ \pi, & p = s, q < t \\ 3\pi/2, & p < s, q = t \end{cases} \quad (5.5)$$

Using Eqs. 5.1 and 5.3, we obtain \mathbf{H}_{MIMO} , and using Eqs. 5.1–5.5, we obtain \mathbf{H}_{OAM} . We next demonstrate benefits of OAM.

Simulation Results In MIMO communication systems, the rank of the channel matrix is an indicator for the number of data streams that can be spatially multiplexed on the channel. Therefore, higher rank channels provide more wireless capacity. To understand the impact of OAM on MIMO channel rank, we first investigate in Fig. 5.3 (*left*) how the rank of \mathbf{H}_{OAM} compares with that of \mathbf{H}_{MIMO} . For this evaluation, we consider an $M = N = 256$ metasurface system. The carrier frequency and corresponding wavelength (λ) are

set to 2.4 GHz and 124.91 mm respectively, the meta-atom spacing Δ is 20 mm, and the gain coefficient β is 40 dB. The distance D between transmit and receive metasurfaces is varied from $10^2\lambda$ to $10^5\lambda$ in steps of $10^2\lambda$, where each step results in a different channel matrix. OAM(Q) indicates Q different OAM states are generated by transmit meta-atoms (*i.e.*, $\tilde{l}_n = n\%Q$). The figure shows that OAM-based MIMO systems achieve higher channel rank than traditional MIMO system under similar channel conditions, and that the \mathbf{H}_{OAM} rank increases with increasing the number of OAM states. This rank enhancement is a result of the additional spiral phase factor (Eq. 5.2) conditioned in OAM-based MIMO systems.

We next compare in Fig. 5.3 (*right*) the end-to-end channel capacity performance of OAM-based MIMO systems against the traditional MIMO system. The capacity is computed as:

$$C = \sum_i^{\gamma} \log_2 \left(1 + \frac{P_i}{\sigma^2 / \delta_i^2} \right) \text{ bps/Hz} \quad (5.6)$$

where γ is the channel matrix rank, σ^2 is the channel noise variance, and δ_i is the i^{th} singular value obtained via singular value decomposition of the channel matrix. The total transmit power budget is set to one Watt, allocated proportionally to all transmit meta-atoms (P_i). The figure shows that OAM-based MIMO systems achieve significantly higher wireless capacity than the traditional MIMO system. At a propagation distance of 200λ , traditional MIMO achieves a capacity of 8.98 bps/Hz, whereas OAM-MIMO with 2, 4, 6, 8, and 10 states achieves 35.90, 70.43, 99.86, 111.71, and 138.91 bps/Hz capacity performance respectively.

5.2 Monolith: Design and Analysis

We now introduce Monolith, a metamaterials-based programmable surface design that transforms an incident Gaussian beam into multiple OAM orthogonal beams with different topological charges. We will describe our surface design as well as its encoding and decoding scheme.

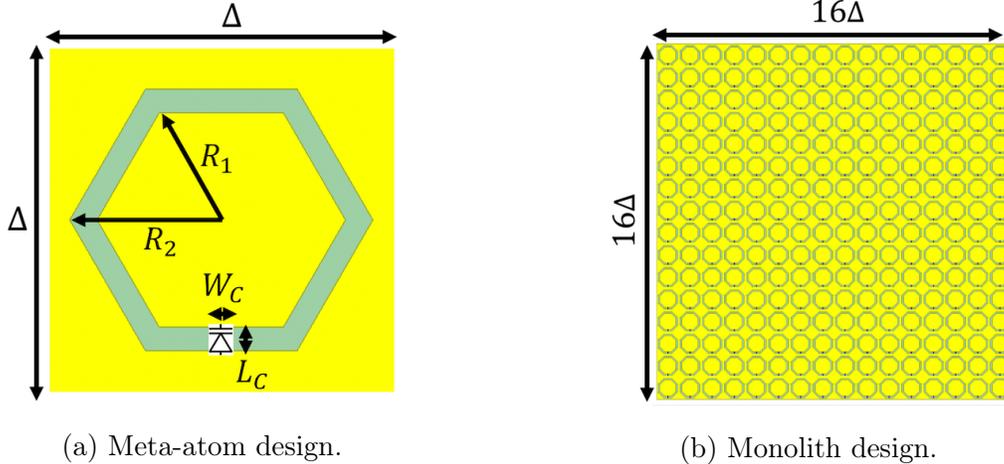


Figure 5.4: Schematic model of meta-atom and metasurface. Our meta-atom consists of several design parameters, including Δ , R_1 , R_2 , W_c , and L_c .

5.2.1 Surface Design

Figure 5.4 shows Monolith design architecture, consisting of its core building blocks, *meta-atoms*. Each meta-atom is an independently controllable scatterer that locally alters the phase of the incident wave. Here we first describe our meta-atom design and then present how we leverage Monolith to generate multiple OAM beams.

Meta-atom Figure 5.4(a) shows our meta-atom (hexagonal structure), made up of copper material that lays on top of a dielectric substrate (green/light shade layout). A varactor diode (*i.e.*, a voltage-dependent capacitor) connects the inner and outer metal patches of the meta-atom. By controlling voltage levels across the varactor diode, we can change its capacitance and in turn alter the meta-atom's response to incident wave's EM field. To understand this phenomenon more clearly, let us model the load impedance across the meta-atom as $Z_{eff} = j\omega L + \frac{1}{j\omega C}$, where L and C are the inductance and capacitance of the meta-atom, respectively. By applying a certain voltage value across the varactor, we alter C and obtain a new load impedance Z_{eff}^{new} . This new load impedance then results in a new electromagnetic response to the incident wave. Figure 5.5 illustrates the magnitude (*Left*) and phase (*Right*) response of EM waves reflected from the meta-atom as we apply different voltage values. By applying voltage from 0 to 10V, our meta-atom can provide up to 2π phase shift with a

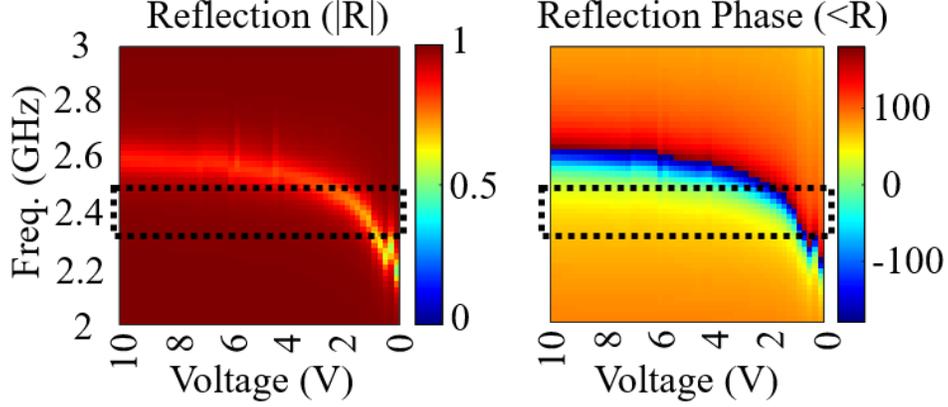


Figure 5.5: *Left*: simulated magnitude; *right*: phase of reflected coefficient for the meta-atom across different voltage values from 0 to 10 V.

near-perfect reflection magnitude within 2–3 GHz. Using this technique, we will arbitrarily configure each meta-atom with a desired phase shift and form an OAM phase distribution across the entire metasurface.

Generating OAM Beams Now, we will describe how we leverage the ensemble of many meta-atoms to generate OAM beams. A single metasurface layer is made up of 256 meta-atoms arranged in a 16×16 uniform rectangular array fashion as shown in Fig. 5.4(b). To generate each OAM beam as represented in Eq. 5.2, we configure the helical phase distribution ($e^{jl\varphi}$ term) by applying an appropriate voltage distribution across the surface. Specifically, we calculate a required phase shift per meta-atom for an OAM state l as:

$$\varphi(x, y) = l \cdot \tan^{-1}(y/x) - 2\pi/\lambda \cdot r_{xy} \quad (5.7)$$

where (x, y) indicates the coordinate location of each meta-atom in the surface plane with its center as an origin. λ is a carrier wavelength, and r_{xy} is the distance between the center of a transmitter and the coordinate (x, y) , which is equivalent to $\sqrt{x^2 + y^2 + d^2}$. Here, d is the distance between an incident wave and the center of the metasurface. After acquiring the required phase distribution from Eq. 5.7, we use a genetic optimization algorithm to find

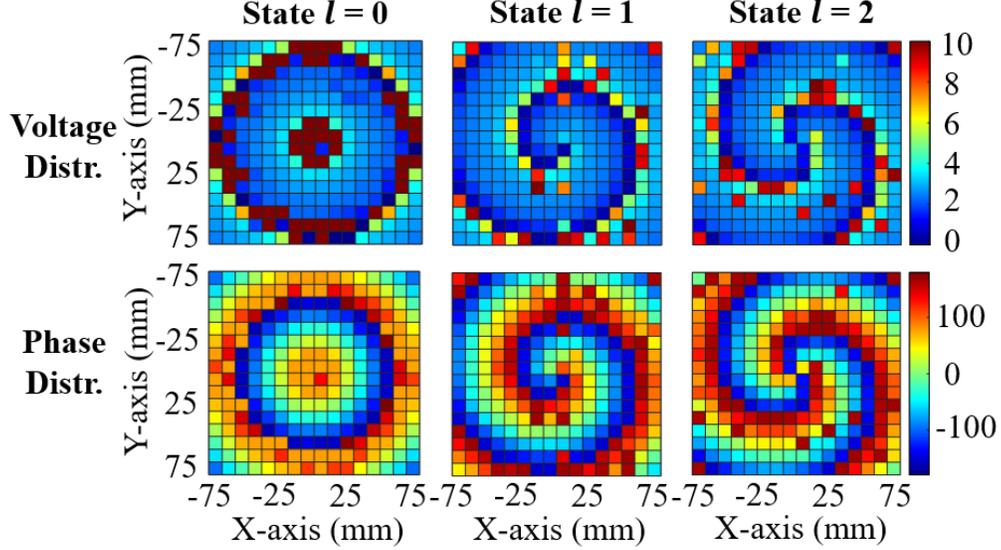


Figure 5.6: *Upper*: an applied voltage distribution; *lower*: phase distribution by applying the voltage distribution. X- and Y-axis represent the coordinate of each meta-atom with the meta-atom at the center being (0,0) mm.

a corresponding voltage distribution:

$$\Theta^* = \underset{x,y}{\operatorname{argmin}} \sum (|F(V(x,y)) - \varphi(x,y)|) \quad (5.8)$$

where $F(V(x,y))$ represents the voltage-to-phase mapping function based on Fig. 5.5. Applying the foregoing OAM beam generation analysis, we present the phase and voltage distributions for each OAM beam with its topological charge $l = 0, 1$, or 2 in Fig. 5.6. Since each surface layer is dedicated for a single OAM state, we stack several layers of the surface to multiplex different beams. We will explore an optimal stacking configuration in our future work.

Simulation Results To evaluate Monolith’s ability to generate different OAM states, we present our HFSS simulation results on Monolith. Figure 5.7 shows that Monolith can support an OAM state l of $-2, -1, 0, 1$, and 2 . We also observe that the beam with state $l = 0$ shows a high intensity at the center while the magnitude of the beams with non-zero OAM state nulls at the center. This nullified area expands as $|l|$ increases. Furthermore,

Monolith creates one phase twist for OAM state +1/-1 and two phase shifts for OAM state +2/-2. Every beam with a positive-valued state complements its respective negative-valued state. These results thus comply with our theory, and we verify orthogonality of OAM beams with different states (§5.1).

5.2.2 Data Encoding and Decoding

In this section, we describe Monolith’s data encoding and decoding schemes.

Encoding In traditional ambient backscatter networks, the received signal at the reader can be expressed as:

$$y(t) = x(t) + \alpha bx(t) + w(t) \quad (5.9)$$

where $x(t)$ is an ambient signal from a base station, $w(t)$ is noise, α is an attenuation factor of backscattered signals relative to ambient signals, and b is a binary bit transmitted by the backscattering transmitter [89]. Similarly, we can represent the received signal of our OAM-based system as:

$$y(t) = x(t) * h_{env,r}(t) + (x(t) * h_{env,t}(t)) * h_{oam}(t) \cdot B + w(t) \quad (5.10)$$

where $h_{env,r}(t)$ and $h_{env,t}(t)$ are wireless channels between the base station and the reader and between the base station and the tag, respectively. Here, $*$ operator indicates convolution. $h_{oam}(t)$ is a multi-state OAM channel defined in Eq. 5.2, and B is a vector of the phase modulation $e^{j\theta_{l_1}(t)}$, $e^{j\theta_{l_2}(t)}$, \dots , $e^{j\theta_{l_Q}(t)}$. Each phase modulation $e^{j\theta_{l_n}(t)}$ corresponds to an OAM state l_n , and Monolith manipulates the phase $\theta(t)$ according to the data that is being modulated. We note that OAM state of 0 (*i.e.* a Gaussian beam without OAM) is not multiplexed and thus it is not part of H_{OAM} .

Decoding For traditional ambient backscatter networks, the receiver distinguishes reflected and non-reflected states using two power levels, $|1 + \alpha|^2 P$ and P . By doing so, they decode

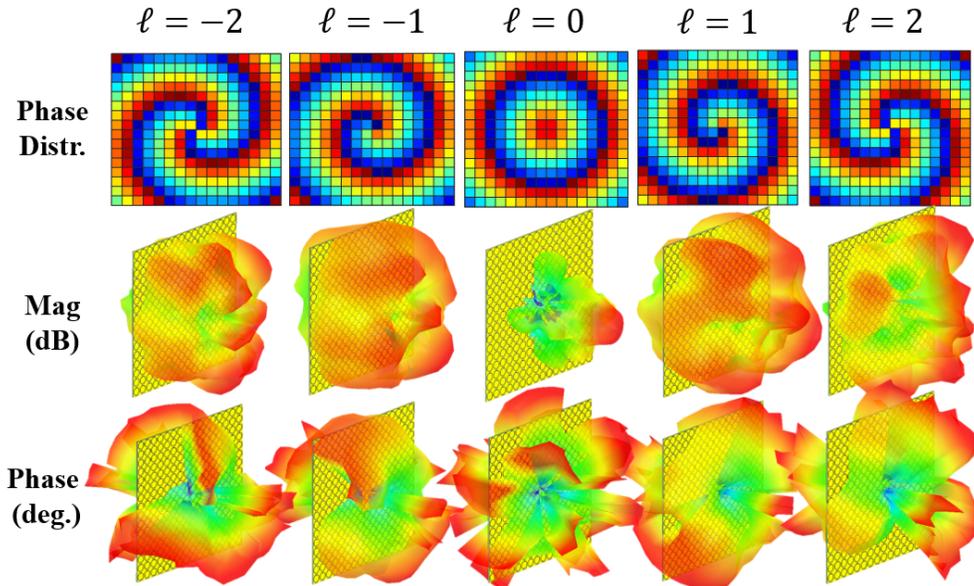


Figure 5.7: *Upper*: a phase distribution of an OAM state from -2 to 2 ; *middle*: a magnitude of the generated OAM beams; *lower*: a phase of the generated OAM beams.

the information from backscattered signals [89]. Monolith, however, separates the signal in two stages. First, we separate h_{env} from received signals to extract h_{OAM} . Since h_{env} is a Gaussian beam without OAM, we can demultiplex it by applying a simple spatial filter. Then, for a mixed-state channel h_{OAM} , we transform each OAM beam back to a Gaussian beam using a mode-conversion spatial filter and successively cancel out from the mixed-state OAM channel until we decode all beams.

5.3 Challenges and Discussion

Beam Divergence OAM beams with higher-order states provide more capacity gains, but such beams physically diverge quickly, making it difficult for a receiver to fully capture OAM-multiplexed beams. To minimize the divergence of OAM, we must optimize many factors, including the carrier frequency, propagation distance, beam waist at the transmitter, and OAM state indices. Otherwise, the required receiver size becomes too large, rendering the OAM-MIMO system impractical. An alternative solution is to design a demultiplexing algorithm that captures and decodes only the partial area of the received beam where the

orthogonality of different OAM states is preserved. This is feasible because the phase pattern repeats over different topological areas, and this repetition tends to occur more frequently when the state indices are low as shown in Fig. 5.6. By doing so, we can minimize the receiver size—we leave this for future work.

Beam Distortion Since scatterers deteriorate the orthogonality of OAM beams, it is challenging to realize OAM-MIMO through such environments. We can mitigate the noise by using multi-plane-light converters (MPLCs) as a spatial filter. Specifically, MPLC shapes the EM wave at multiple propagating distances to accomplish mode conversion between OAM and Gaussian beams at different locations [151]. By doing so, the patterns are updated using the wavefront matching method, which partially undo the effects of scattering.

Beam Steering In line-of-sight (LoS) communication, alignment between the transmitter and receiver is crucial for sufficient signal power. Such alignment is even more important for OAM multiplexing systems because misalignment may cause crosstalk among channels. To mitigate misalignment, [155, 76] have explored the simultaneous generation and tunable steering of multiplexed OAM beams. Specifically, [60, 76] steer OAM beams by applying a proper phase distribution over different meta-atoms.

Backscatter Communication Prior works on backscatter communication have mainly focused on sensor applications that require less than one Mbps. In particular, ambient backscatter communication [89, 31] was introduced to enable low-power, low data-rate device-to-device communication by backscattering ambient signals. While [109] has explored a way to increase the transmission range of such signals, our work in contrast focuses on increasing the link capacity of ambient backscattering via low-power OAM multiplexing techniques.

Optical Communications Since the path between two satellites is free-space, inter-satellite networks commonly deploy free-space optical communication (FSO), which uses optical links to exchange data [78]. Specifically, [40, 112, 95] have used reflectarray antennas, which are placed on the satellite and excited via the feed horn. The use of extremely high frequency

mitigates the beam divergence, rendering OAM-MIMO to be easily deployable without a complex receiver design.

5.4 Conclusion

This chapter proposes a holistic OAM-based ambient backscatter design for rank-deficient wireless environments. We theoretically analyze the OAM-MIMO multiplexing and provide a preliminary study on generating the OAM beam with different states.

Part III

Intelligent Massive IoT Networks

Chapter 6

Cross-Link Channel Prediction in Massive IoT Networks

While conventional wisdom holds that the channels of the nodes that are at least half a wavelength apart are independent due to link-specific signal propagation paths [138], with enough background data and measurements of a wireless environment, we find that it is possible to predict the CSI of a link that has not been recently observed. Fig. 6.1 illustrates

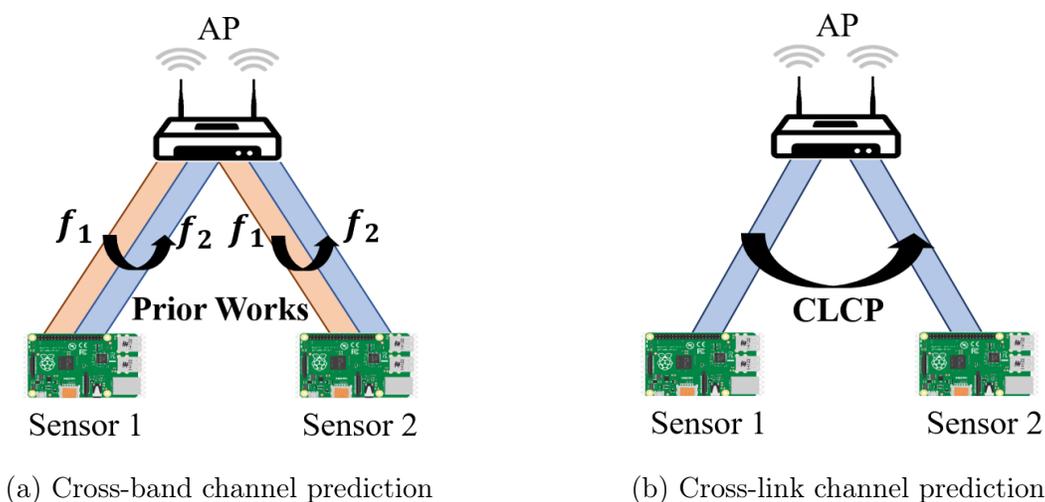


Figure 6.1: *Left:* Previous work on cross-band channel prediction infers a downlink channel at frequency f_2 using the uplink channel at frequency f_1 on the *same* link ([142, 14]). *Right:* CLCP infers the channel to Sensor 2 using channel measurements from Sensor 1.

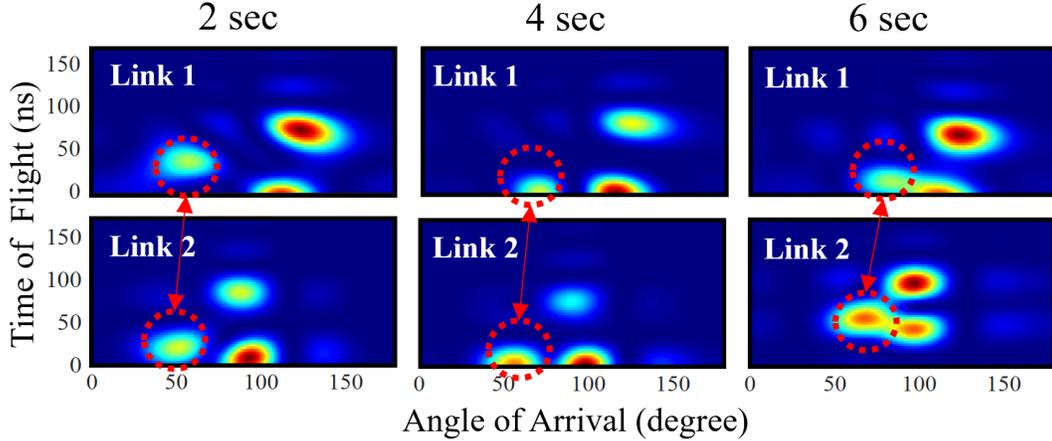


Figure 6.2: *CLCP's mechanism*: Time of flight and angle of arrival for two nearby IoT sensors (*upper* and *lower*, respectively). While each sensor (link) has a distinct set of static wireless paths, their parameters both indicate reflections off the same moving object, highlighted in red dotted circles.

our high-level idea: unlike previous works [142, 14] that use CSI measurements at frequency f_1 to infer CSI at f_2 for a single link, our approach exploits the cross-correlation between different links' wireless channels to leverage traffic on one sensor's link to predict the wireless channel of another.

To support our idea, we measure the wireless channels from two nearby sensors that are roughly 30 cm apart in the presence of a moving human. Figure 6.2 visualizes these channels using two wireless path parameters, Time-of-Flight (ToF) and Angle-of-Arrival (AoA)¹. While Sensor 1's channel (*upper*) is independent from Sensor 2's channel (*lower*), ToF and AoA from both reflect the same moving body in the environment, indicated by the dotted red circles, while other major paths remain unchanged. This suggests the existence of a function that correlates the occurrence of wireless links of stationary sensors in the presence of moving reflectors.

We propose the *Cross-Link Channel Prediction* (CLCP) method, a wireless channel prediction technique that uses multiview representation machine learning to realize this vision.

¹We note that the spectrum representing Link 2 at 6-sec abruptly changes due to a carrier frequency offset resulting from non-synchronized local oscillators. This figure shows the raw observation without any error correction.

In summary, a CLCP AP uses uplink channels observed from the last transmission to predict a large number of unobserved wireless links. Here, the AP treats each channel reading like a photo of an environment taken at a particular *viewpoint*, and combine multiple different views to form a joint representation of the environment. It then exploits this representation to predict unobserved wireless CSI readings. Since CLCP has no dedicated channel sounding or extra pilot signal, the aggregated overhead no longer scales with the number of radios. However, since traffic patterns are not regular in reality, the number of channels and links observed at the time of prediction is likely to change. Such dynamics in input data often lead to an explosion in the number of trainable parameters, making the learning process intractable [162, 153]. CLCP hence adopts a special model design (§6.3.1) and training paradigm (§6.3.1) that allows it to construct a joint representation from any combination of multiple views. Moreover, we often treat deep learning models as *black boxes* whose inner workings cannot be interpreted. As a result, designers cannot differentiate whether the trained model truly works or is simply overfitting. To validate CLCP’s learning mechanism, we visualize a fully trained feature representation and interpret it using the wireless path parameters, ToF and AoA (§6.3.2).

Our implementation and experimental evaluation using 802.11ax validate the effectiveness of CLCP through microbenchmarks and head-to-head performance comparison against OptML [14] and R2F2 [142] cross-band channel prediction methods (§6.5). CLCP provides a 2× throughput gain over baseline 802.11ax and a 30% throughput gain over R2F2 and OptML in a 144-link testbed. Moreover, by eliminating the need for channel sounding, it increases sleep time by 65%.

6.1 Cross-Link Correlation Model

To substantiate the concept of cross-link correlation and in particular in Fig. 6.2 using a simple model, let us consider a receiver estimating channels from two separate transmitters

(Fig. 6.3). All the nodes—the two transmitters T_1 and T_2 and the receiver R —lie on the y -axis. Let us assume that only two paths originate from every transmitter: the line-of-sight (LoS) path parallel to the y -axis and the path reflected by the red obstacle on the right. Then, in the plane wave approximation (we neglect here effects over the amplitudes), the channels between the two transmitters T_1 and T_2 and the receiver can be expressed as the sum of the LoS and the reflected paths as follows:

$$\begin{aligned} H_1 &= H_{1LoS} + H_{1R}(X) \\ H_2 &= H_{2LoS} + H_{2R}(X) \end{aligned} \tag{6.1}$$

where the four paths are defined by:

$$\begin{aligned} H_{1LoS} &= e^{-jkR} \\ H_{2LoS} &= e^{-jk(R-T)} \\ H_{1R}(X) &= \Gamma \cdot e^{-jk2\alpha(X)} \\ H_{2R}(X) &= \Gamma \cdot e^{-jk2\beta(X)} \end{aligned} \tag{6.2}$$

In Eq. (6.2), k is the wave number at the generic frequency and Γ is the reflection coefficient of the red obstacle, $\alpha(X)$ and $\beta(X)$ are half of the total reflected path lengths, which are defined as follows:

$$\begin{aligned} \alpha(X) &= \sqrt{X^2 + \frac{R^2}{4}} = \sqrt{X^2 + P^2} \\ \beta(X) &= \sqrt{X^2 + \frac{(R-T)^2}{4}} = \sqrt{X^2 + Q^2} \end{aligned} \tag{6.3}$$

where $P = R/2$ and $Q = (R - T)/2$.

We focus on the reflected paths, which depend on the distance X between the red obstacle and the y -axis. If the distance of the obstacle increases by ΔX , we can express the new

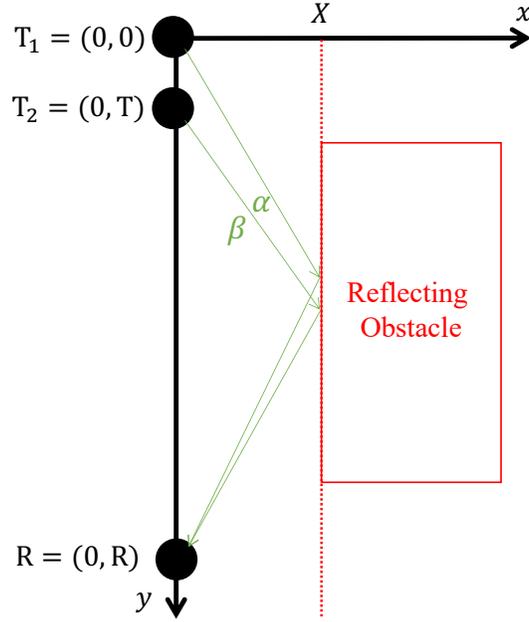


Figure 6.3: Cross-link correlation modeling under a multipath propagation scenario with two transmitters T_1 and T_2 , one receiver R , and a reflecting obstacle lying on the y -axis at X . Here, α and β are reflected path lengths.

reflected channels by considering their first order Taylor expansions, *i.e.*:

$$\begin{aligned}
 H_{1R}(X + \Delta X) &= H_{1R}(X) + \frac{dH_{1R}(X)}{dX} \Delta X \\
 &= H_{1R}(X) \left[1 - \frac{2jkX\Delta X}{\sqrt{X^2 + P^2}} \right] \\
 H_{2R}(X + \Delta X) &= H_{2R}(X) + \frac{dH_{2R}(X)}{dX} \Delta X \\
 &= H_{2R}(X) \left[1 - \frac{2jkX\Delta X}{\sqrt{X^2 + Q^2}} \right].
 \end{aligned} \tag{6.4}$$

If the two transmitting nodes are sufficiently close with respect to their distance to the receiver (which means $T \ll R$), we can consider $P \approx Q$ and define:

$$\Theta(X) = \left[1 - \frac{2jkX\Delta X}{\sqrt{X^2 + P^2}} \right] \approx \left[1 - \frac{2jkX\Delta X}{\sqrt{X^2 + Q^2}} \right]. \tag{6.5}$$

Rewriting the reflected channels corresponding to $X + \Delta X$ as:

$$\begin{aligned} H_{1R}(X + \Delta X) &= H_{1R}(X)\Theta(X) \\ H_{2R}(X + \Delta X) &= H_{2R}(X)\Theta(X) \end{aligned} \tag{6.6}$$

we can now compare the two 2-path channels when the red obstacle is located in X to those when the red obstacle is in $X + \Delta X$:

$$\begin{aligned} \tilde{H}_1 &= H_{1LoS} + H_{1R}(X + \Delta X) = H_{1LoS} + H_{1R}(X)\Theta(X) \\ \tilde{H}_2 &= H_{1LoS} + H_{2R}(X + \Delta X) = H_{2LoS} + H_{2R}(X)\Theta(X) \end{aligned} \tag{6.7}$$

If the number of measurements (thanks to the knowledge of the channels for different points in the frequency domain) allows separating the LoS contribution from the reflected one when the obstacle is located in X , then we can estimate H_{1LoS} , $H_{1R}(X)$, H_{2LoS} and $H_{2R}(X)$ by measuring H_1 and H_2 . Then, when the obstacle moves to $X + \Delta X$, it is possible to derive $\Theta(X)$ from \tilde{H}_1 by separating it into H_{1LoS} and $H_{1R}(X + \Delta X)$. This simple model demonstrates that, at least an approximate fashion, there is a correlation between different wireless links in the presence of moving reflectors, making it possible to provide a simple but principled estimate of \tilde{H}_2 without actually measuring it.

6.2 Primer: ML Background

In the context of learning, cross-band channel prediction [142, 14] is a markedly different problem than cross-link channel prediction. In the former, uplink and downlink channels share exactly the same paths in the wireless channel. Therefore, the learning task is simply to map the (complicated) effect of changing from one frequency band to another, given a fixed set of wireless channel paths. For CLCP, the channels of nearby radios have distinct paths and the learning task is to elucidate the correlations between the two links.

To correlate the occurrence of distinct wireless links given they share *some* views on the wireless environment, CLCP must **1)** discard radio-specific information from raw observations and extract a feature representation that conveys information on moving reflectors; **2)** integrate the extracted representation with radio-specific properties (*e.g.*, signal paths and noises) to synthesize unobserved channels. However, radio-specific and environment-specific information superimpose in channel readings and thus are not easily separable. We exploit representation learning to achieve these tasks. Specifically, an *encoder* network of our model accomplishes the first task, and a *decoder* network achieves the second task. Before we explain our design in detail, we first provide the background on representation learning.

Variational Autoencoder (VAE). The goal of a VAE [81] is to compress input data into a low-dimensional feature while preserving maximum information about the original data. In a nutshell, its *encoder* neural network compresses the input data into a representation z , also known as a *latent variable*, and its *decoder* decompresses z back to the original input. Its loss function, known as an *evidence lower bound* (ELBO), consists of a *reconstruction term* and a *regularization term*. The reconstruction term ensures that the decoder output resembles the encoder input, while the regularization term encourages the latent variable to follow a normal, Gaussian distribution, preventing overfitting and enhancing generalizability beyond the training set [121, 13]. By minimizing this loss, the VAE finds a distribution that best describes the data.

Multiview Representation Learning. *Multiview (multimodal) representation learning* [153, 113, 127] has proven effective in capturing the correlation relationships of information that comes as different *modalities* (distinct data types or data sources). For instance, different sets of photos of faces, each set having been taken at different angles, could each be considered different modalities. This model learns correlations between different modalities and represents them jointly, such that the model can generate a (missing) instance of one modality given the others. Like VAEs, it encodes a primary view into a low-dimensional feature that contains useful information about a scene. However, instead

of simply reconstructing the input, multi-view learning decodes this latent variable, which describes the scene, into a *secondary view*.

A more advanced form of multiview learning adopts multiple different views as input data and encodes them into a joint representation. By analyzing multiple information sources simultaneously, we learn a better, comprehensive feature representation. Past works [128] use multiview learning to synthesize unseen images at an arbitrary angle given the images of a scene taken at various angles. Likewise, we treat each wireless link like a photo of a scene taken from a particular viewpoint. We observe wireless links from various viewpoints and combine them to form a joint representation of the channel environment. We then exploit this joint representation to predict the link at unobserved viewpoints.

6.3 CLCP Design

Our system operates in a sequence of steps shown in Fig. 6.4: to begin with, an AP acquires *channel state information* (CSI) and *buffer status report* (BSR) from all users, which it then uses to schedule an uplink OFDMA packet. Once scheduled, the AP triggers the uplink transmission. When the acquired CSIs become outdated, the AP estimates the channels from the most recently received OFDMA packet and uses them to predict (1) the remaining bands of estimated channels using *cross-band channel prediction* (CBCP) and (2) full-band CSIs of unobserved links using *cross-link channel prediction* (CLCP). The AP then schedules the uplink packet using the predicted CSIs and sends a trigger frame (TF) to re-initiate the uplink transmission.

(1) Opportunistic Channel Observation. OFDMA divides the entire bandwidth into multiple subchannels known as a resource unit (RU). The AP assigns RUs to individual users, enabling one OFDMA packet to contain channel estimates from multiple users. Our goal is to leverage multiple subchannels already present in existing OFDMA transmissions to predict the channels of *other* users. In Fig. 6.4, user 3, 4, and 6 simultaneously transmit

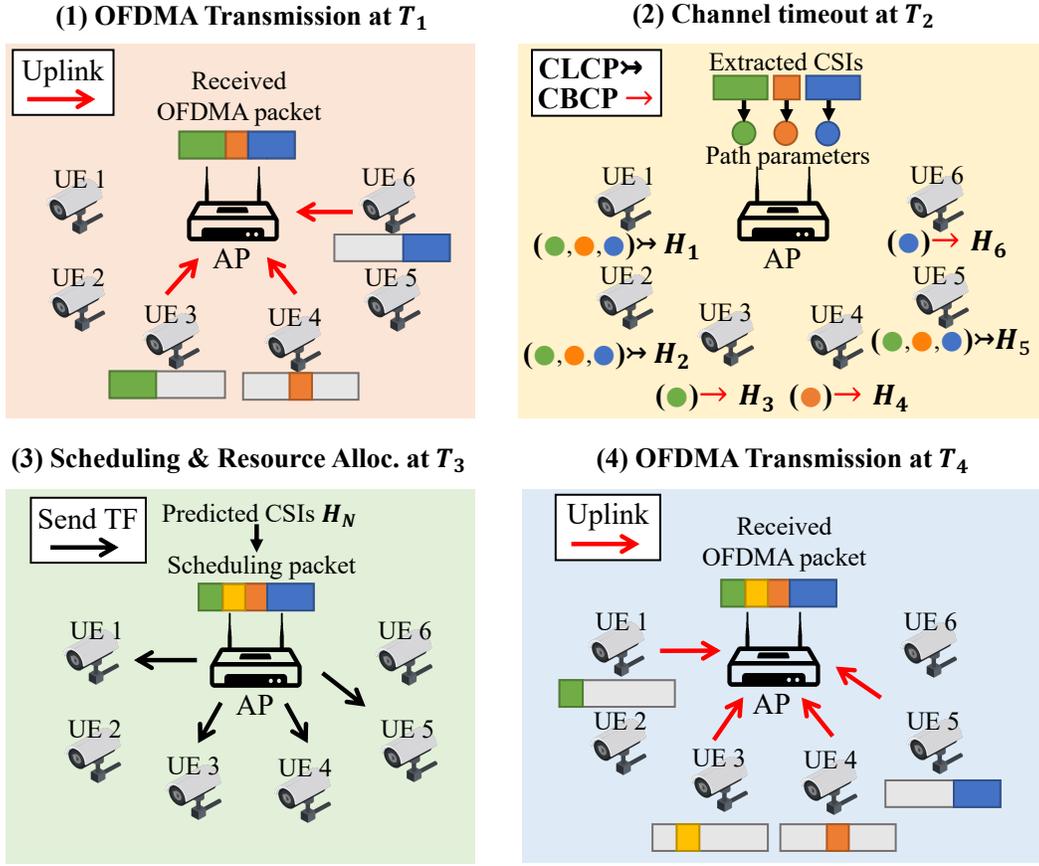


Figure 6.4: *System overview for uplink transmission:* (1) an AP receives uplink traffic from multiple users simultaneously; (2) when channels become outdated, the AP extracts the path parameters of partial CSIs estimated from the latest OFDMA packet and predicts unobserved links using CLCP and unobserved bands using CBCP in a server; (3) Then, the AP schedules uplink traffic based on predicted CSIs and triggers the users. (4) Finally, the AP receives the scheduled packet.

uplink signals in their dedicated RUs, which adds up to a full bandwidth channel. After the acquired CSIs expire, the AP extracts the subchannels from these users. It then uses this information to predict the remaining channels of these users *and* the whole channels of unobserved users 1, 2, and 5. This way, we completely eliminate the need for channel sounding.

(2) Channel Prediction. Once the AP estimates the subchannels, it directly routes the channel estimates to a backend server through an Ethernet connection. At the server-side,

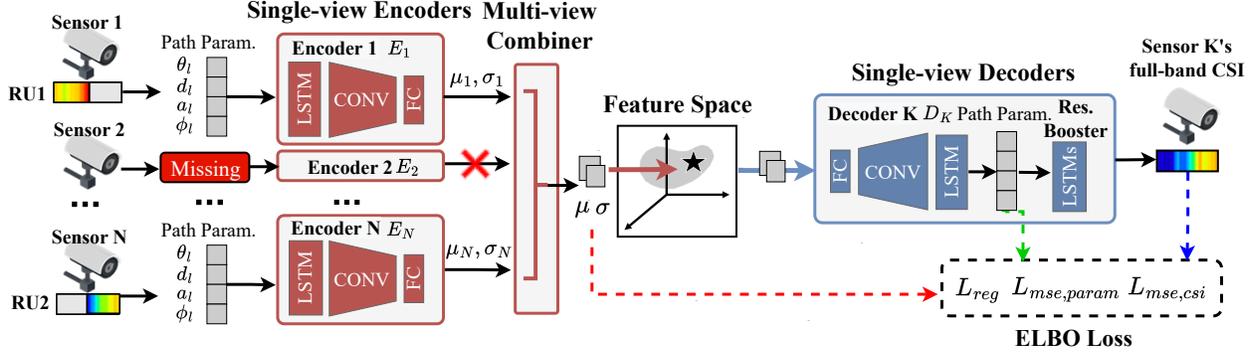


Figure 6.5: CLCP ML model with N measured channels, each represented as a set of wireless path parameters $\{\theta_l, d_l, a_l, \phi_l\}_{l=0}^L$ with L paths estimated from measured channels. Each set of the parameters is served by a **Single-view Encoder** network E_i ($i \in [1, N]$) that compresses the measured wireless path information of its dedicated radio and outputs variational parameters μ_i and σ_i . The **Multi-view Combiner** integrates all variational parameters into μ and σ , based on which **Single-view Decoder** networks D_K generate a set of path parameters that are unobserved. If any input channel is not observed, CLCP drops the respective encoder network (E_2 , for example).

path parameters are extracted from observed channel estimates and are then fed to CLCP for channel prediction (§5.2).

(3) **Scheduling and Resource Allocation (SRA)**. The AP schedules the upcoming OFDMA transmission using predicted CSIs (§6.3.3) and BSRs. It is worth noting that OFDMA scheduling requires full-bandwidth channels in order to allocate a valid combination of RUs with varying subcarrier sizes [147] and to find a proper modulation and coding scheme (MCS) index for each RU. Moreover, unlike 802.11n and 802.11ac, 802.11ax provides support for uplink MU-MIMO, which requires CSI to find an optimal set of users with low spatial channel correlation and appropriate decoding precedence. After scheduling an uplink packet, the AP encloses this information in a *trigger frame* (TF) and broadcasts the frame.

(4) **Uplink data transmission**. After receiving the TF, the corresponding users transmit data accordingly.

6.3.1 Cross-Link Prediction Model

Our CLCP ML model is summarized in Fig. 6.5: there is a *single-view encoder* network $q_\phi(z|h)$ (§6.3.1) dedicated to each single *view* (*i.e.*, channel h of each radio). Every encoder outputs the distribution parameters, μ and σ , and a *multi-view combiner* (§6.3.1) fuses the parameters from all encoders into a joint representation z . If a channel is not observed, we drop its respective encoder network (*e.g.* E_2 in Fig. 6.5). A decoder network $p_\theta(h|z)$ (§6.3.1), dedicated to each *target* radio whose CSI we seek to synthesize, samples the joint representation z to predict a *cross-link* channel.

A key challenge is that across different prediction instances, the combination of observed channels varies, as we utilize channels from existing OFDMA transmission. In Fig. 6.5, the input channels consist of two RUs, one from Sensor 1 and the other from Sensor N . In the next prediction instance, the OFDMA packet is likely to contain a different RU combination, possibly from different radios. Such difference in the number of input channels complicates the learning process. We will address how CLCP is robust against observations that vary in frequency (§6.3.1) and link (§6.3.1).

Path Parameter Estimator

To reduce the learning complexity of CLCP, we extract the geometric information *i.e.*, wireless path parameters, from raw CSI and use them directly as input data. This approach is particularly effective because the path parameters are *frequency-independent*, which means that CLCP is robust against observations from different frequency bands. We can represent the channel h observed at antenna M_i as a sum of L paths where each path is characterized by an arrival angle θ_l , a time delay d_l , an attenuation a_l , and a reflection ϕ_l :

$$h_{M_i} = \sum_l^L (a_l e^{\frac{-j2\pi d_l}{\lambda} + j\phi_l}) e^{\frac{-j2\pi i k \cos(\theta_l)}{\lambda}} \quad (6.8)$$

where λ and k are wavelength and antenna distance. To extract the 4-tuple of parameters $\{(\theta_l, d_l, a_l, \phi_l)\}_{l=0}^L$, we use the mD-Track path estimation algorithm [157]. We select the number of estimated paths L as four. Also, as a lesser number of tones limits path estimation accuracy, we use raw CSI instead of path parameters during operation at or under 56 tones (20 MHz legacy 802.11n devices, *e.g.*). For simplicity, we denote the 4-tuple as \ddot{h} .

CLCP’s Single-View Encoder

Like VAE, a single-view encoder $q_\phi(z_n|\ddot{h}_n)$ compresses its dedicated channel \ddot{h}_n into Gaussian distribution parameters, μ_n and σ_n . We can then compute the distribution that best describes its dedicated viewpoint as $z_n = \mu_n + \sigma_n \odot \epsilon$ where \odot refers to the element-wise product and $\epsilon \sim N(0, \mathbf{I})$. The encoder consists of a long short-term memory (LSTM) layer followed by two stacked convolutional (CNN) layers and fully connected (FC) layers. Each convolutional layer employs 1D kernels as filters and is followed by a batch norm layer that normalizes the input at each layer. To give a high-level insight, the LSTM networks are used to process temporal information [141], such as the ToF, while CNN layers handle spatial information [52], particularly the AoA. Also, we use a multi-layer CNN to gradually compress data in the encoder. This architectural choice has been empirically tested, following conventional practices in the field. We learn all layer weights of the encoders and decoders end-to-end through backpropagation. If an encoder has no input due to an unobserved link (*e.g.*, Sensor 2), CLCP drops the respective encoder networks (E_2 in Fig. 6.5).

CLCP’s Multi-view Combiner

A naïve approach to handle a varying number of inputs is to train an encoder network for each combination of inputs. However, this approach leads to a significant increase in the number of trainable parameters, making it computationally intractable for complex multi-view datasets. Instead, CLCP constructs one encoder per each user and efficiently fuses the output of all encoders into a joint representation. We model this multiview combiner

after the *product-of-experts* (PoE) [66, 153] whose core idea is to combine several probability distributions (*experts*), by multiplying their density functions. With any combination of input channels $\ddot{\mathbf{H}} = \{ \ddot{h}_i \mid \text{channel of } i^{\text{th}} \text{ radio} \}$, we formulate the multiview combiner as:

$$q_\phi(z|\ddot{\mathbf{H}}) \propto p(z) \prod_{\ddot{h}_n \in \ddot{\mathbf{H}}} \tilde{q}(z_n|\ddot{h}_n) \quad (6.9)$$

where $p(z)$ is a Gaussian distribution, and $\tilde{q}(z_n|\ddot{h}_n)$ is an encoder network dedicated to the n^{th} radio. Here, the multiview combiner approximates the joint representation z by multiplying all individual distributions z_n , parameterized by the encoder outputs, μ_n and σ_n .

PoE assumes that multi-view inputs are conditionally independent given the distribution z , which represents dynamics in the wireless environment. This assumption is valid because the signal paths of different radios experience independent fading in general. In other words, the propagation of one radio signal does not directly affect the propagation of another radio signal. However, when external factors like human mobility are introduced into the environment, the changes in the propagation of the two signals can become correlated. The foregoing assumption allows factorization of the encoders in CLCP. With this factorization, we can simply ignore the encoder with no input channel.

CLCP's Single-View Decoder

Our single-view decoder $p_\theta(\ddot{h}_n|z)$ is another DNN, whose input is the joint representation z . The goal of each decoder is to synthesize an accurate channel estimate of its dedicated radio. The decoder architecture is in the exact opposite order of the encoder architecture. The decoder predicts the path parameters of a target radio, which are then reconstructed into a channel estimate h_{interm} according to Eq. (6.8). In practice, estimating path parameters induces some loss of information. To compensate for this loss, we add an extra layer, a

resolution booster, that consists of stacked BiLSTMs and takes in the intermediate channel h_{interm} and generates a final channel estimate h_{pred} .

Objective function

We define our loss function as:

$$\text{ELBO} = \mathbb{E}_{q_\phi(z|\mathbf{H})} [\log p_\theta(\check{h}|z)] - \beta D_{\text{KL}}(q_\phi(z|\mathbf{H})||p(z)) \quad (6.10)$$

where β is a weight for balancing the two terms in the ELBO. Like VAE (§6.2), the second term represents the regularization loss (L_{reg}) that makes the joint distribution z close to a Gaussian distribution $p(z)$. For VAE, the first term represents the reconstruction error. However, instead of simply reconstructing the input, our decoders must synthesize unobserved channels. Hence, our first term consists of a mean squared error between the predicted channel h_{pred} and ground-truth channel h_{gt} of a target radio: $L_{mse,csi} = \frac{1}{S} \sum_{s=0}^S (\|h_{s,gt} - h_{s,pred}\|_2)$ where S is the number of subcarriers. We also include a mean squared error between the predicted and ground-truth path parameters as a part of the first term. Since some paths are stronger than others when superimposed, we weigh the error of each path based on its amplitude a as follow: $L_{mse,param} = \sum_{l=0}^L (a_l \|\ddot{h}_{l,gt} - \ddot{h}_{l,pred}\|_2)$. Our path parameter estimator prioritizes the estimation of stronger paths. In cases where the actual number of paths is less than the defined number L , the amplitude of the extra paths is lower than that of the actual paths. Our weighted loss function that considers the amplitude of each path effectively minimizes the impact of erroneously estimated extra paths. Finally, our loss function becomes $L_{mse} = -(\alpha L_{mse,csi} + \eta L_{mse,param})$ where α and η are the weight terms. The weight term applied to the CSI loss further compensates for path estimation errors. By maximizing ELBO, we maximize the lower bound of the probability of generating a comprehensive representation of the wireless environment and predicting accurate channels.

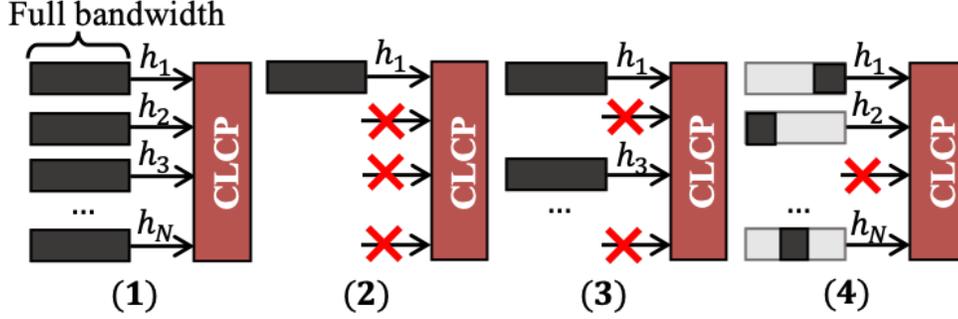


Figure 6.6: Multi-step training paradigm.

Multi-stage training paradigm.

If we train all encoder networks altogether, then CLCP cannot generate an accurate prediction when some links are unobserved at the test time. Conversely, if we individually train each encoder network, CLCP fails to capture the relationship across different links. To address this issue, our loss function consists of four ELBOs (see Fig. 6.6): (1) one from feeding all N full-band channels, (2) the sum of N ELBO terms from feeding each full-band channel at a time, and (3) the sum of k ELBO terms from feeding k randomly chosen subsets of full-band channels (where k is randomly chosen between 2 and N). Lastly, (4) we repeat the first three steps with a random subchannel to mimic channels in the actual OFDMA transmission. We then back-propagate the sum of four ELBOs end-to-end.

In the case of adding a new node, once CLCP’s state is already constructed, only one additional decoder needs to be added. Thus, CLCP does not require a fresh training start.

6.3.2 Model Interpretability

The proximity of low-dimensional features in the feature space indicates their relevance for a given task. For instance, if we encode the channels of two radios into our low-dimensional features and they appear close in the feature space, it suggests that both channels have been affected by the same moving reflectors simultaneously. To understand the learning mechanism of CLCP, we visualize the feature space using t-distributed Stochastic Neighbor Embed-

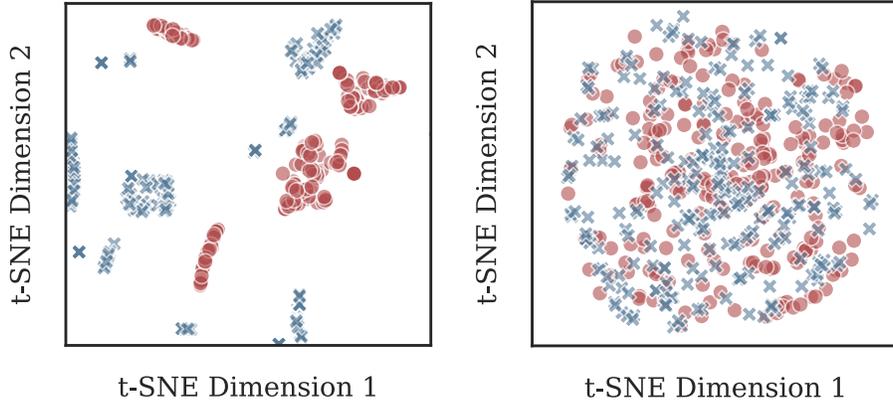


Figure 6.7: 2D t-SNE visualization of the latent space of two adjacent clients before (*left*) and after training (*right*). Encoded CSI instances of Radio 1 and 2 are highlighted in red and blue, respectively.

ding (t-SNE) dimension reduction, a statistical technique for visualizing high-dimensional data by giving each data point a location in a two-dimensional map. We collect channel instances of two radios for three hours and randomly select 250 timestamps, which we then feed into CLCP encoders. Each encoder output μ is represented by a color-coded data point, with red and blue points representing Radio 1 and Radio 2, respectively.

Figure 6.7 depicts visualizations of the feature space of two clients, before and after training. Prior to training, an encoder outputs separate and non-intersecting clusters. Post-training, the red and blue data points coalesce into a singular and Gaussian cluster, revealing that the model learns to correlate two distinct links in a high-dimensional feature map.

Figure 6.8 provides a comprehensive analysis of the fully trained feature embedding, utilizing ToF and AoA path parameters. Closely located low-dimensional features for each radio indicate similar path parameters as shown in the spectra on right. Furthermore, a pattern across different radios is discernible, where the number of strong reflectors in Radio 1 and Radio 2’s spectra are similar when their low-dimensional features are in close proximity. For instance, both Radio 1 and 2 exhibit many reflectors in the upper-left spectrum, and only one in the bottom-left spectrum. This reveals the CLCP’s proficiency in encoding wireless channels and distributing encoded features based on reflector movement. Also, CLCP can

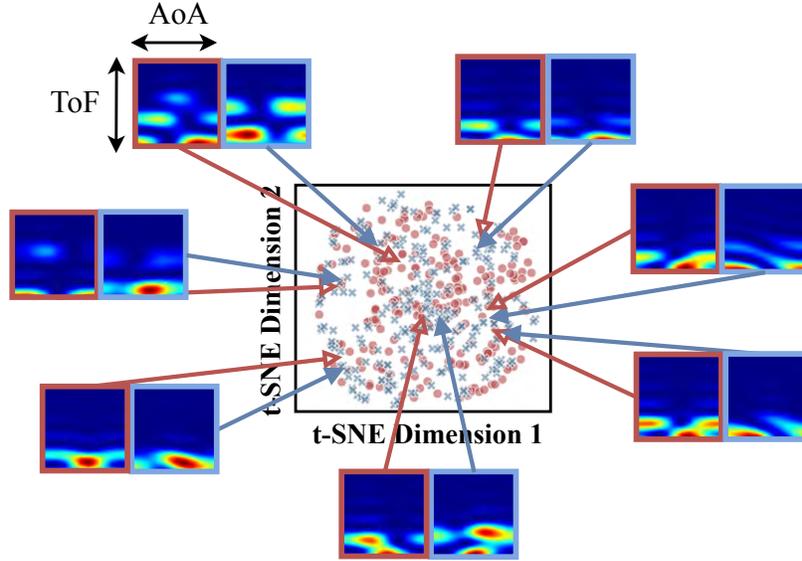


Figure 6.8: A fully trained latent feature embedding alongside exemplary data points and their corresponding path parameters. The x-axis represents the Angle of Arrival (AoA), while the y-axis shows the Time of Flight (ToF). Data points that are positioned close to each other for each radio reveal similar path parameters.

efficiently locate encoded features in the latent space, making accurate generalizations when presented with previously unseen channels.

6.3.3 Scheduling and Resource Allocation

While scheduling and resource allocation have been extensively studied for LTE, 802.11ax's OFDMA structure differs significantly due to its variable RU sizes, requiring a different design. Frequency-selective fading necessitates assigning RUs to users based on channel measurements. However, exhaustively searching all user-RU combinations is computationally heavy, as the search space grows exponentially with user count and RU granularity. Furthermore, allocating RUs solely on channel capacity, without considering variable RU sizes and user data rates, leads to inefficient padding of packets to equalize transmission durations, increasing MAC layer overhead. In this section, we propose an optimized 802.11ax scheduling algorithm that jointly optimizes both channel conditions and user data requirements.

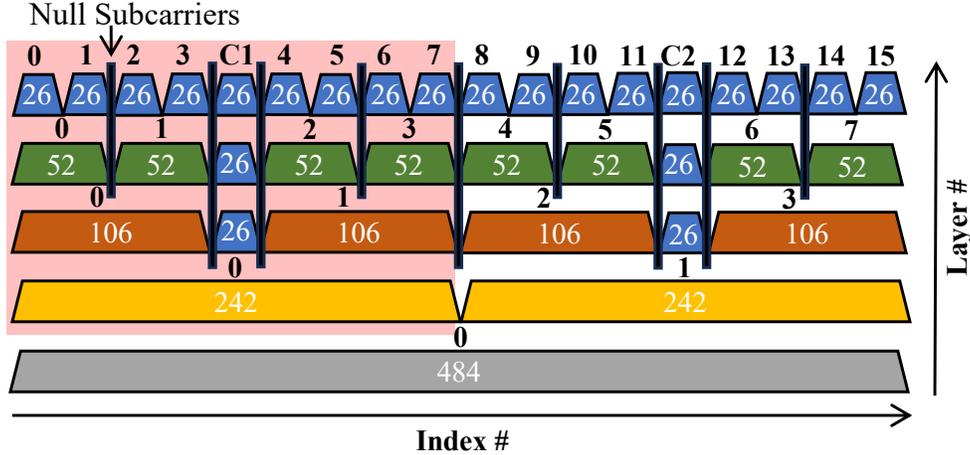


Figure 6.9: 802.11ax resource unit (RU) locations in a 40 MHz channel with 5 layers. RU locations in a 20 MHz channel with 4 layers are highlighted in red. The RU index at each layer increases from left to right. An intermediate 26-tone RU is denoted as C.

Problem Formulation

In 802.11ax OFDMA, the channel bandwidth is segmented into RUs of varying sizes. These RUs range from the smallest unit of 26 tones (equivalent to 2 MHz) up to the largest unit of 996 tones (77.8 MHz). The sizes and locations of these RUs are standardized for channel bandwidths of 20, 40, 80, and 160 MHz. Figure 6.9 illustrates the RU locations for a 40 MHz bandwidth channel as an example.

Our objective is to maximize channel capacity by selecting a combination of RUs that fully utilize the available bandwidth. We also aim to match RU sizes with user data needs, allocating larger RUs to high-data devices and smaller RUs to low-data devices. This scheduling task is complex, as the number of possible combinations grows exponentially with the number of users and RU size options.

We can formulate the scheduling problem as the following optimization task. The scheduler allocates RUs to certain users whose channels maximize zero-forcing beamforming (ZF-BF) capacity, which we compute by summing achieved rates at each subcarrier s . The

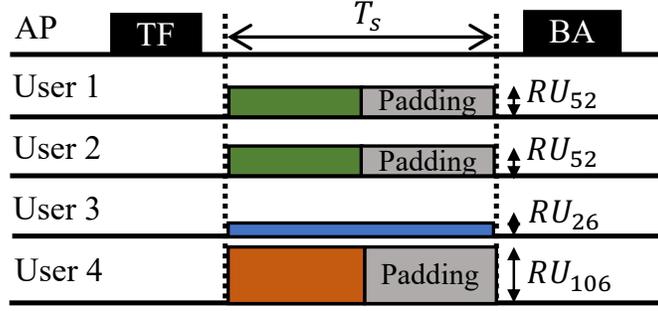


Figure 6.10: OFDMA transmissions of four users with an equal amount of data D in a $20MHz$ bandwidth channel. T_s is a scheduled transmission duration.

capacity of a complete user schedule is the sum of the capacities of all assigned RUs:

$$R_{ZFBF} = \sum_j \sum_{s \in s_j} \sum_{u \in u_j} \log_2(1 + P_{u,s}) \quad (6.11)$$

where $P_{u,s}$ denotes the signal power of subcarrier s of user u that is assigned to the j^{th} RU. For RU with 106 or more subcarriers, multiple users can be assigned. Our optimization objective is $\arg \max_{g \in G} R_{ZFBF}(g)$, where g is chosen from G , the set of all possible user schedules.

Beyond maximizing channel capacity, we must also consider users' buffer status. Ignoring this can lead to inefficiencies. For example, as illustrated in Fig. 6.10, if users with equal data D are scheduled to RUs of different sizes, their transmission durations T_j will vary. However, 802.11ax requires all transmissions in an OFDMA packet to start and end simultaneously. This forces users with smaller RUs to transmit null data to fill the time, wasting capacity and power. Therefore, our scheduler must balance channel capacity maximization with efficient use of allocated resources based on actual user data needs.

Our Algorithm

Our resource allocation and scheduling algorithm, summarized in Fig. 6.11, follows these key steps: (1) we create "buckets" of UE candidates for each RU layer (see Fig. 6.9), based on their buffer status; (2) within each bucket, we select UEs that maximize the capacity for each RU; (3) we then search for the optimal combination of RUs. This approach increases

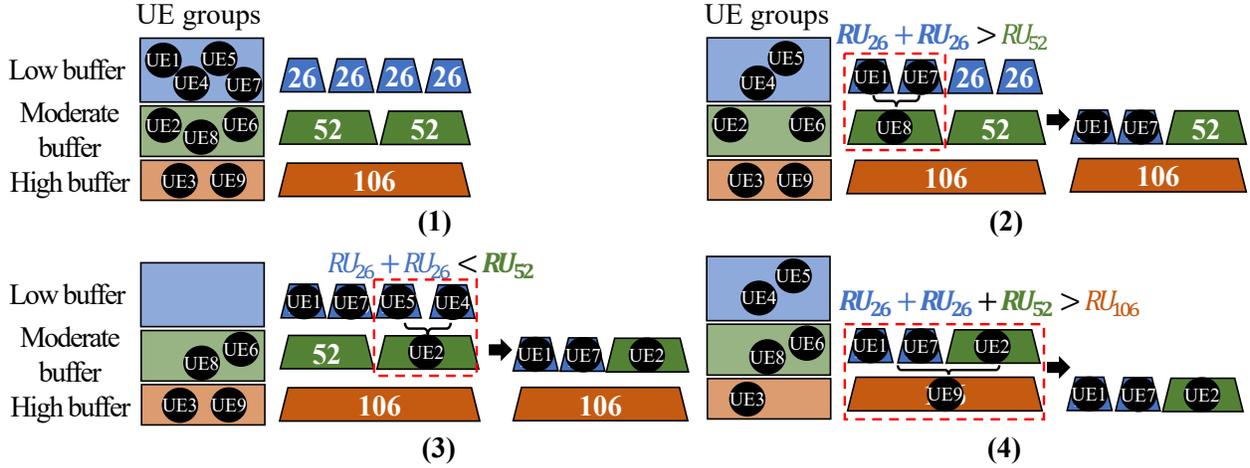


Figure 6.11: *Our scheduling and resource allocation algorithm:* (1) rate adaptation and bucketing; (2)-(3) UE selection and capacity comparison between the upper two layers; (4) UE selection and capacity comparison between previously selected RUs and the next layer.

the likelihood of assigning users with lower buffer occupancy to smaller RUs (higher layers) and users with higher buffer occupancy to larger RUs (lower layers), thereby optimizing the resource allocation. By doing so, we aim to match RU sizes with user data needs while maximizing overall data rates.

(1) Rate Adaptation and Bucketing. The scheduler starts by calculating the Modulation and Coding Scheme (MCS) index for each UE using ESNR-based rate adaptation [58]. It then creates L empty buckets, each representing a specific RU layer, and distributes UEs into these buckets based on the amount of data they need to send. This distribution is done by comparing each UE's data rate (calculated as data length divided by maximum transmission duration²) against the data rates of RUs at different layers and corresponding MCS indices [71]. A UE is placed in one or more buckets dedicated to RU layers where its data rate is lower than the RU's data rate. For example, a UE with a data rate lower than both 26-tone and 52-tone RU rates would be placed in buckets dedicated for these two layers.

(2) UE selection and capacity comparison. Starting from the top-left RU layer, the scheduler selects a UE that maximizes channel capacity for each RU within its corresponding

²In 11ax, the transmission time of the PPDU is limited to 5.484 ms

bucket. As illustrated in Fig. 6.11, it first identifies the UE that maximizes capacity for the first 26-tone RU in the "low buffer" bucket (UE 1). This UE is then removed from the bucket, and the process repeats for the second 26-tone RU (UE 7), as the 802.11ax standard allows only one RU per user. Next, a UE is chosen for the 52-tone RU in the layer below (UE 8). The scheduler then compares the capacity of this 52-tone RU against the combined capacity of the two 26-tone RUs, selecting the configuration with higher capacity (in this case, UE 1 and UE 7). Any unselected UE (like UE 8) is returned to its original bucket.

(3) RU selection and combination. This process iterates through subsequent resource blocks until reaching the lowest layer, completing a combination that covers the entire 160 MHz bandwidth. If the bucket is empty, the scheduler removes the corresponding layer. Our algorithm efficiently reduces the search space while considering both user data needs and frequency-selective fading.

802.11ax supports joint MU-MIMO and OFDMA multi-user transmissions for RUs larger than 106 subcarriers. Our algorithm employs a multi-user greedy approach for MU-MIMO scheduling. It starts by selecting the user with the highest channel capacity, and then iteratively adds users that best match the existing selection. This way, we progressively build better locally optimal solutions at each iteration. After scheduling an uplink packet, CLCP directly obtains CSI for users scheduled for uplink MU-MIMO transmission.

6.4 Implementation

We conduct an experimental study on cross-link channel prediction in a large indoor lab for testbed (a) as shown in Fig. 6.12(a) and an entire floor for testbed (b) as depicted in Fig. 6.12(b).

1. Testbed (a) is equipped with high-bandwidth 802.11ax commodity radios with three APs highlighted in red—Asus RT-AX86U APs with the Broadcom chipset (four antennas) that support 160-MHz bandwidth (2048 subcarriers per spatial stream) at 5

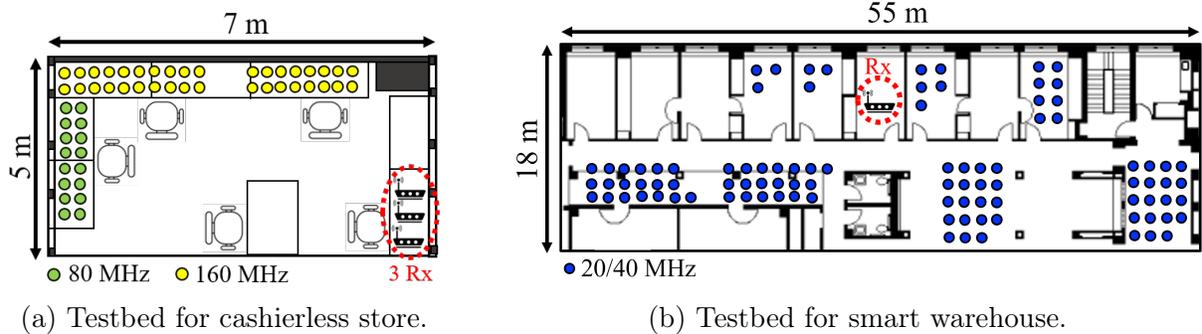


Figure 6.12: CLCP preliminary experimental testbed floor plans and radio hardware with different operating bandwidths.

GHz. The users (Fig. 6.12(c)) include the Asus RT-AX82U router (four antennas), Asus ZenWifi XT8 routers (four antennas), and the smartphones, such as the Samsung A52S (single-antenna) and Xiaomi Mi 11 (two antennas). While the bandwidth of the routers is 160 MHz, the smartphones use 80 MHz bandwidth. We identify a total of 144 separate links (with each spatial stream counted as a separate link). To extract CSIs from commodity 802.11ax devices, we use the AX-CSI extraction tool [51]. Two people moved at a constant walking speed of one to two meters per second.

2. Testbed (b) is equipped with low-bandwidth 802.11n commodity radios. We use 11n WPJ558 with the Atheros QCA9558 chipset (three antennas) for both the AP and users. We identify a total of 95 separate links in NLoS settings, and we extract traces using the Atheros CSI Tool. There were 8 people moving at a constant walking speed.

Each CSI extraction tool measures the amplitude in its own arbitrary unit (AU). The AU of AX-CSI scales between 0 and 2000, while Atheros CSI uses 0 to 500. For trace-driven simulation, we normalize the CSI amplitude outside the stop-band region. All devices

together are generating traffic constantly using `iperf`. Since commodity 802.11ax devices do not allow OFDMA scheduling on the user side, we run a trace-driven simulation using a software-defined simulator. We implement CLCP’s using Pytorch, and the model parameters include batch size of 16 and learning rate of $5e^{-6}$. We employ Adam for the learning rate optimization.

Channel measurement error. To minimize the packet boundary delay during OFDM symbol boundary estimation, the AP averages the phases of multiple CSIs within the channel coherence time. Then, to compensate for the amplitude offset due to the power control uncertainty error, the AP obtains the Received Signal Strength Indicator (RSSI), reported alongside CSI in the feedback, detects the RSSI and phase outliers, and discards the associated packet. Lastly, to minimize carrier frequency offset due to non-synchronized local oscillators, the AP subtracts the phase constant of the first receiving antenna across all antennas. Since phase constant subtraction does not alter a relative phase change across antennas and subcarriers, the signal path information is preserved in CSI.

6.5 Evaluation

We begin by presenting the methodology for our experimental evaluation (§6.5.1), followed by the end-to-end performance on throughput and power consumption (§6.5.2). Lastly, we present a microbenchmark on CLCP (§6.5.3).

6.5.1 Experimental Methodology

Use cases. We evaluate CLCP in two use case scenarios: a cashierless store (Fig. 6.12(a)) and a smart warehouse (Fig. 6.12(b)). **Cashierless stores** typically experience a high demand of data traffic as densely deployed video cameras continuously stream their data to the AP for product and customer monitoring. To reflect a realistic cashierless store application, we configure all users to continuously deliver standard quality video of 1080p using UDP

protocol for trace-driven simulation. Also, we leverage 80 and 160 MHz bandwidth for every uplink OFDMA packet. In **smart warehouses**, IoT devices transmit relatively little data traffic and are widely and sparsely deployed compared to the cashierless store use cases. Hence, each uplink packet has 20 and 40 MHz bandwidth in NLoS settings. CLCP is scoped for static IoT sensor deployments with environmental dynamics, and the movement of the sensor³ is beyond the scope of our work.

Evaluation metrics. To quantify the network performance, we calculate uplink throughput as the total data bits transmitted to the AP divided by the duration, measured every 500 ms. Also, we report the total number of Target Wake Time (TWT) packets to evaluate power consumption. TWT is an 802.11ax mechanism for power-saving, where user devices sleep between AP beacons and only wake up to transmit signals, such as uplink data transmission and channel reports. If a user is not scheduled for uplink transmission or reporting CSIs to the AP, it is not triggered for a TWT packet, increasing device sleep time and conserving IoT device battery life.

Baselines. Our baseline follows sounding protocols in which the AP periodically requests BSRs and CSIs from all users. Upon receiving the a null data packet (NDP) from the AP, all users calculate the feedback matrix for each OFDM subcarrier as follows [71]:

$$\frac{\text{CSI tones} \times \text{CSI bits} \times \text{TxAntenna} \times \text{RxAntenna} \times T_c}{\text{Subcarrier Group} \times \text{Feedback Period}} \quad (6.12)$$

where T_c signifies the wireless channel coherence time. We use 8-bit CSI quantization, a channel coherence time of 15 ms, and a subcarrier grouping of 4. The other control protocols we consider are BSR report (32 bytes), BSR poll (21 bytes), CSI poll (21 bytes), MU-RTS (20 bytes), CTS (14 bytes), TF ($28 + (5 \times K)$ bytes), and BlockAck/BA ($22 + (5 \times K)$ bytes),

³In the event of sensor mobility, a system that uses CLCP can gracefully degrade to the status quo for those sensors that are mobile (*i.e.*, mobile nodes use standard channel estimation).

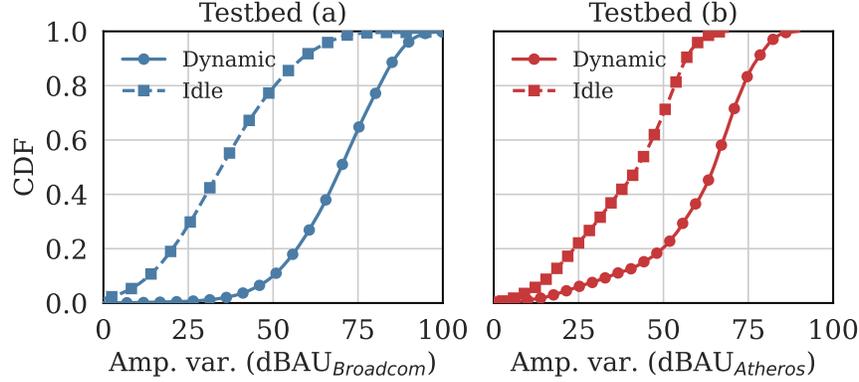


Figure 6.13: The variance of channel amplitude over time with and without human mobility for testbed (a) and (b). The channel amplitude is measured in AU as reported by the tool. Broadcom chipset’s maximum AU is 2000, while Atheros chipset’s maximum AU is 500.

where K denotes the number of users. SIFS takes $10\mu s$. We note that BSRs and CSIs are delivered to the AP via OFDMA transmission to minimize the overhead.

Algorithms. We compare CLCP to the following algorithms which collectively represent the state-of-the-art in channel prediction: (1) R2F2 [142] infers downlink CSI for a certain LTE mobile-base station link based on the path parameters of uplink CSI readings for that *same* link, using an optimization method in lieu of an ML-based approach; (2) OptML [14] leverages a neural network to estimate path parameters, which, in turn, are used to predict across frequency bands. Both algorithms predict downlink based on uplink channels in LTE scenarios, where the frequencies for downlink and uplink are different. To adopt these algorithms into our OFDMA scheduling problem, the AP triggers all users to transmit pilot signals simultaneously in the 242-subcarrier RUs, and these algorithms predict a full-band channel (2048 subcarriers) of each user based on the received RU. We use a maximum likelihood approach for fast path parameter estimation.

Channel variability. It is important to take into account environmental dynamics in the training dataset since training in a static environment does not generalize well. Fig. 6.13 demonstrates the channel variance of idle channels without human mobility and our channel environment affected by moving reflectors for testbed (a) and (b). Specifically, we calculate

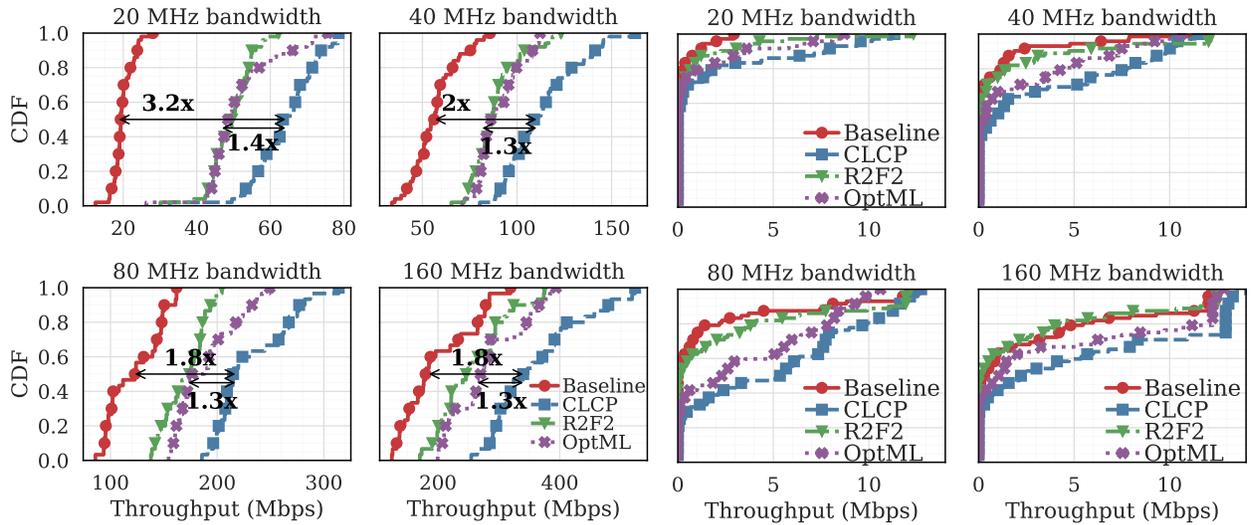
the variance of channels ($20 \log_{10}[\text{var}(|H|)]$) over time. Each point indicates different subcarriers and links. Here, the channel amplitude is in arbitrary units as reported by the tool. In Fig. 6.13, the variance of our channel data is 20-30 dB higher than that of idle channel data, suggesting that our links are not idle and there is environment variability caused by moving reflectors.

6.5.2 End-to-End Performance

We evaluate CLCP’s end-to-end throughput against R2F2 and OptML, and compare its power consumption performance

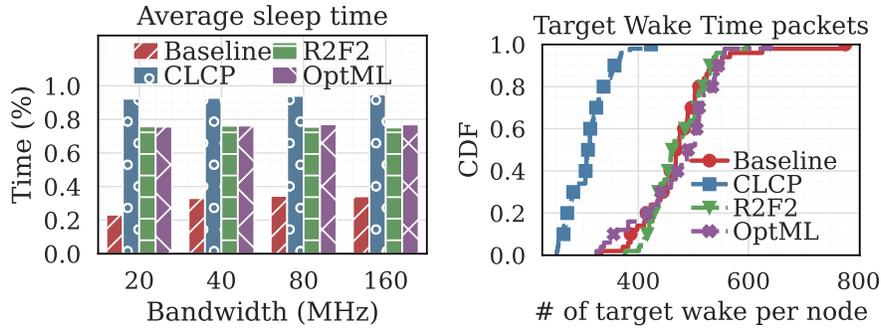
Significant throughput improvement. The end-to-end throughput performance of CLCP is evaluated and compared to the baseline, R2F2, and OptML across time and users. Fig. 6.14(a) summarizes the results for 20, 40, 80, and 160 MHz bandwidth channels. Here, each data point of the curves indicates an aggregated uplink throughput within 500 ms duration. With 20 MHz bandwidth, CLCP shows a 3.2x improvement in throughput compared to the baseline and a 1.4x improvement over R2F2 and OptML. Similarly, CLCP provides a 1.9x to 2x throughput improvement over the baseline for 40, 80, and 160 MHz channels and a 1.3x improvement over R2F2 and OptML. Particularly, CLCP improves spectral efficiency for smaller bandwidths. While existing cross-band prediction algorithms need a dedicated pilot signal for channel sounding from all users, CLCP makes use of the channel estimates obtained from existing transmissions, completely eliminating the need for additional signal transmissions and corresponding control beacons.

In Fig. 6.14(b), we present the end-to-end throughput performance across users. Here, each data indicates a throughput of one user within 10 second duration of uplink traffics. It is worth noting that as the bandwidth increases, more users get an opportunity to send their data. For 20 MHz bandwidth, only 20% to 40% of users can send their data while for 160 MHz bandwidth, more than 50% to 70% of users communicate with the AP. More



(a) Throughput across time.

(b) Throughput across users.



(c) Power consumption.

Figure 6.14: End-to-end performance on throughput and power consumption: (a) aggregated throughput across time for every 500 ms, (b) throughput across users for 20, 40, 80, and 160 MHz bandwidth, and (c) device sleep time over the entire transmission duration and the total number of Target Wake Time (TWT) triggered on every user.

importantly, we observe that for all bandwidths, CLCP enables 15% to 20% more users to delivery their data within 10 second duration.

Increasing device sleep time. TWT reduces the power consumption by letting users sleep and only waking up when they need to transmit their packets. Fig. 6.14(c) (*upper*) shows the average sleep time of all users over the entire measurement duration. While the users sleep slightly over 25% of the time using the baseline, CLCP enables 90% of users to remain in sleep state, which is roughly 65% and 15% longer than the baseline and cross-band prediction algorithms, respectively. It is worth noting that the sleep time of the cross-band algorithms is longer than the baseline because each user sends at least 260 byte of its channel information for the baseline while users simply transmit a pilot signal for R2F2 and OptML. However, CLCP does not need bulky feedbacks nor pilot signals, minimizing contention between users and increasing the duration of power save mode. Fig. 6.14(c) (*lower*) shows how many TWT packets are exchanged when 150 MB data are delivered to the AP. (*i.e.*, how frequently each user wakes up to participate in channel sounding and data transmission). Here, CLCP has significantly less TWT counts because the users stay idle during channel acquisition. In contrast, all users are forced to transmit a signal with the baseline, R2F2, and OptML. Given that the average wake power and transmit power is 600 uW and 135 mW, respectively, we can infer that CLCP’s power consumption is significantly less.

6.5.3 Microbenchmark

We present a microbenchmark on the prediction accuracy, channel capacity, packet error rate, and PHY-layer bit rates. Microbenchmark results are obtained with 8 devices under the settings in Fig. 6.12(a) for a cashierless store testbed and 19 devices under the settings in Fig. 6.12(b) for a smart warehouse testbed.

Prediction accuracy. We use an *error vector magnitude* (EVM) to represent how far a predicted channel H deviates from a ground-truth H_{gt} : $\text{EVM} = 10 \log_{10} (|H - H_{\text{gt}}|^2 / |H_{\text{gt}}|^2)$. According to 802.11 spec. [71, 72], BPSK modulation requires an EVM between -5 to -10

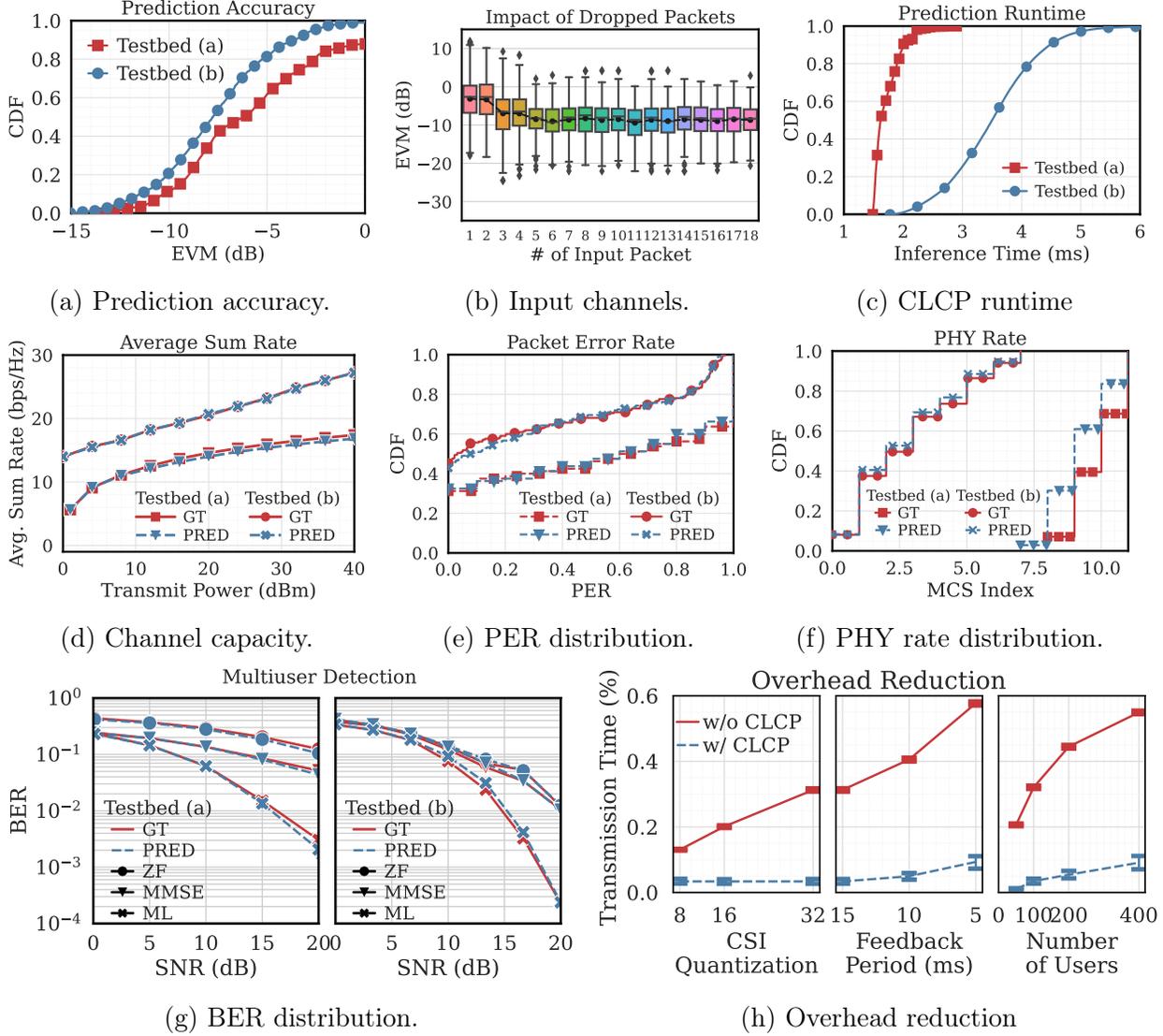


Figure 6.15: Microbenchmark on the prediction accuracy, runtime, overhead reduction, and scheduling performance under two scenarios: testbed (a) for cashierless store with 8 devices and testbed (b) for smart warehouse with 19 devices.

dB, and QPSK needs the EVM from -10 to -13 dB. In Fig. 6.15(a), CLCP provides an EVM of approximately -6 dB and -8 dB for testbed (a) and (b), respectively.

Impact of the number of observed channels. We evaluate the prediction accuracy with varying number of observed channels. As shown in Fig. 6.15(b), the prediction performance significantly improves when there are more than two input users. However, further increasing the number of input channels does not greatly enhance CLCP’s prediction accuracy.

This result suggests that CLCP can predict channels correctly even when there are many unobserved channels

Overhead reduction. Fig. 6.15(c) shows that CLCP with 19 users achieves only about 4 ms of inference time, which comply with the runtime of other VAE-based models [80]. The inference time reduces to 2 ms with 8 users due to less number of encoders in CLCP. Next, we present the overhead reduction with varying parameters in Fig. 6.15(h). We define the overhead as the percentage of CSI transmission time over the total traffic time. In the absence of CLCP, a short feedback period, an increase in the number of users, and a greater number of sub-carriers result in a larger CSI overhead. In a densely deployed scenario, our CLCP notably reduces the overhead. Fig. 6.15(h) (*right*) shows that with 400 users, CLCP can free up more than 40% overhead.

Channel capacity. Fig. 6.15(d) evaluates the channel capacity of OFDMA packets that are scheduled using predicted channels. Channel capacity is defined as the sum of achieved rates at each subcarrier s , that is $R_{\text{capacity}}(RU_i) = \sum_{s \in RU_i} R_{\text{capacity}}(s)$ where RU_i is RU at i -th location. Then, we define the capacity of a final user schedule g as:

$$\sum_j R_{\text{capacity}}(p_j, u_j) = \sum_j \sum_{s \in p_j} \sum_{u \in u_j} \log_2(1 + P_{u,s}) \quad (6.13)$$

where $P_{u,s}$ denotes a transmit power for user u and subcarrier s . We obtain the average sum rate by dividing $\sum_j R_{\text{capacity}}$ by the total number of subcarriers. Fig. 6.15(d) shows that the channel capacity of packets scheduled based on predicted CSIs is nearly identical to that of ground-truth CSIs. These results indicate that our predicted channels are accurate enough for OFDMA scheduling.

PER distribution. Fig. 6.15(e) shows PER distributions of packets scheduled based on groundtruth CSIs and predicted CSIs under SNR below 10 dB. Even if PER is high, packets scheduled with ground-truth CSIs and predicted CSIs share similar PER distributions.

We conclude that even if the channel condition is bad, CLCP still provides good channel prediction.

PHY rate distribution. 802.11ax enables each RU to have its own MCS index, which is calculated based on its channel condition. Therefore, accurate channel estimates are essential for rate adaptation. In Fig. 6.15(f), we present the PHY rate distributions calculated using an effective SNR (ESNR)-based rate adaptation [58]. The algorithm chooses the MCS index from 0 to 11 for testbed (a) while MCS index between 0 and 7 for testbed (b), as defined by 802.11n and 802.11ax standards. This algorithm leverages channel estimates to find a proper MCS index. The results show that the PHY rate distributions of both ground-truth and predicted channel are similar.

Multiuser detection. 802.11ax uses multi-user detection algorithms to separate uplink streams from multiple users. For uplink MU-MIMO, it is crucial to select a subset of users with low spatial channel correlation *and* determine an appropriate decoding precedence. Here, we deploy several multiuser detection algorithms, such as zero-forcing (ZF) and minimum mean squared error (MMSE), that are integrated with a successive interference cancellation (SIC) and the optimal maximum-likelihood (ML) decoder. Fig. 6.15(g) shows a BER of packets that are scheduled with ground-truth CSIs and predicted CSIs. We decode these packets using ZF-SIC, MMSE-SIC, or ML techniques across different SNR values. We observe that the BER of packets from predicted CSIs is slightly higher than that from ground-truth CSIs for ML decoder when SNR ranges from 10 to 16 dB. In contrast, BER with ZF- and MMSE-SIC decoder shows no difference, indicating that CLCP's prediction is accurate for ZF-SIC and MMSE-SIC decoders.

Joint transmission mode. 802.11ax supports joint MU-MIMO and OFDMA for RUs that contain more than 106 subcarriers. Since MU-MIMO is suitable for users with high-data rate, we configure 12 users to continuously deliver standard quality video of 1080p using UDP protocol while the rest of 83 users to send low-rate sensory data. We use a 20 MHz

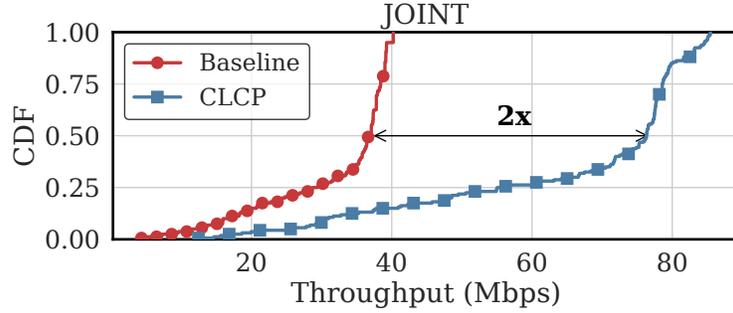


Figure 6.16: Aggregated throughput of joint OFDMA and MU-MIMO.

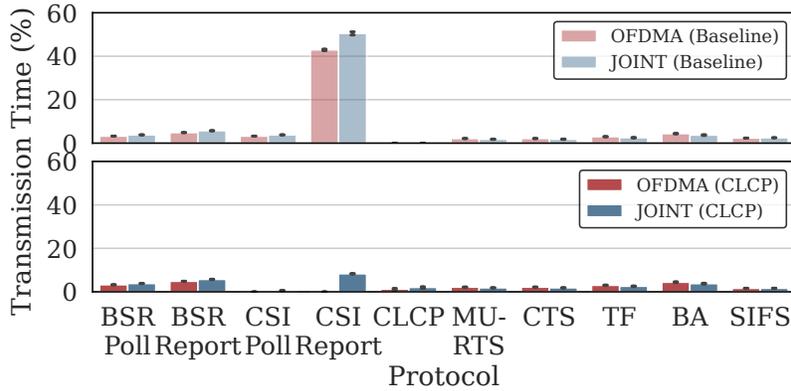


Figure 6.17: Overhead comparison across different protocols.

channel bandwidth and run the experiment multiple times with different transmit powers ranging from 10 to 30 dBm.

Figure 6.16 summarizes the aggregated throughput of the baseline and CLCP in the joint mode. CLCP shows a 2x improvement in throughput compared to the baseline. Fig. 6.17 shows the overhead comparison across different protocols with the baseline and CLCP for the OFDMA only mode and joint mode, separately. The overheads are averaged across different experimental runs. Without CLCP, the average overhead from CSI exchange takes up approximately 43% and 51% of the total transmission time for the OFDMA and joint mode, which is significantly higher than the overhead from other protocols, such as BSR and TF. In the OFDMA mode, CLCP induces no CSI overhead and therefore completely frees up the overhead. In the joint mode, CLCP reduces the CSI overhead by approximately 40%. This is because CLCP explicitly acquires CSIs of users scheduled for uplink MU-MIMO.

6.6 Discussion and Limitations

Environmental changes. During the deployment, the model’s predictive performance may degrade over time due to changes in the environment that violate the model’s assumptions. To mitigate this model drift, we can use continuous learning. In our OFDMA scheduling process, once UEs transmit data in response to the trigger frame, the AP can leverage the multiple sub-channels in received OFDMA transmissions for dual purposes: predicting other users’ channels as described in Section 6.3 and updating model weights. The AP computes the loss between previously predicted channels and received channel estimates, accumulating this loss in batches to update model weights accordingly. If the post-update loss exceeds a threshold for a specified time hysteresis, the AP can initiate re-training for the affected model, indicating a significant environmental change. This adaptive approach helps maintain the model’s accuracy in dynamic environments.

Limitations on the training data. To ensure sufficient channel diversity, our training data was collected over several weeks. However, to accelerate the training process, we can leverage synthetic datasets based on physical models. Recent studies [170, 56] have demonstrated the effectiveness of realistic synthetic data in improving wireless channel prediction accuracy, especially when dealing with limited datasets. Likewise, we may generate and use synthetic data to overcome the limitation. Specifically, we can integrate Sionna’s ray tracing (RT) module [68], a differentiable ray tracer designed for simulating radio wave propagation. By inputting environmental information from our acquired dataset and a visual map of the indoor scene into Sionna RT, we can compute gradients of the channel response. This process allows us to simulate and generate a synthetic dataset featuring environment-specific and physically accurate channel realizations. This method not only enhances the diversity of our

training data but also ensures that the synthetic channels closely reflect the characteristics of the real-world environment we are modeling.

Channel uncertainty in scheduling model. User scheduling is more robust to CSI errors compared to complex MIMO operations like beamforming and spatial multiplexing. For scheduling, CSI primarily serves to compare users' relative channel conditions, aiming to select those with strong channels. Here, the precise CSI values are less critical; the relative ordering of users' channel qualities is more important. In contrast, beamforming and spatial multiplexing demand much more accurate CSI. Beamforming heavily relies on precise phase information for constructive signal alignment, where even minor phase errors can significantly degrade performance. Similarly, spatial multiplexing is highly sensitive to inter-stream interference, necessitating accurate CSI for effective precoding and stream separation. Thus, while beamforming and spatial multiplexing require highly accurate CSI to exploit MIMO capabilities, scheduling can tolerate some level of channel uncertainty.

6.7 Supporting Applications

We describe a couple of supporting cases for CLCP. In particular, we show how CLCP can be used to significantly improve different types of wireless sensor networks.

Smart Retailers. The retail industry is going through a rapid transformation, with IoT systems taking a central part. Recently, Amazon launched Amazon Go, a cashierless grocery store equipped with a massive number of IoT devices. Similarly, Alibaba announced its own cashierless store, Hema, to automate the shopping experience. Smart stores like Amazon Go and Hema exploit massive IoT sensor networks to automate checkouts, manage inventory in real time, and analyze customer behavior. For automatic checkout and inventory management, smart stores use a combination of ceiling-mounted cameras and smart shelf sensors to automatically track customers and items. For example, data from smart shelves are fused with the images captured from many cameras to assist in determining the identity of items

picked and/or placed in inventory locations. Smart stores also analyze mall traffic using a massive number of IoT devices to understand customers' shopping patterns. Specifically, they exploit cameras and Wi-Fi traffic monitoring to see if customers dwell over a product area. Then, in real-time, they analyze that information to adjust store layouts for more efficient customer visits.

To increase the reliability and accuracy of the system, smart stores need to obtain multiple data at various locations from a massive number of IoT sensors. However, massive IoT systems lead to large amounts of aggregate uplink traffic. By minimizing communication overhead from a large number of IoT sensors, CLCP increases the overall spectral efficiency of massive IoT networks in smart stores. CLCP can observe channels from ceiling-mount cameras to predict links from low-power sensors like smart shelves and smart checkout. This way, the AP does not wake up these low-power smart sensors for channel sounding thereby further reducing the power consumption.

Smart Cities. In smart cities, IoT networks provide the ability to remotely monitor, manage, and control cities from massive streams of real-time data. For example, IoT sensors in smart cities improve safety and reduce traffic congestion. Specifically, devices like traffic sensors, cameras, and smart parking sensors collect real-time data that can be shared with both drivers and local authorities. Then they stream collected data to the base station, such that a centralized system can alert when there are traffic issues and provide helpful information regarding parking availability and the best routes to take at any given time. Smart waste management is another use-case of IoT sensors in smart cities. Data-orchestrated IoT systems like smart meters help in gathering and monitoring critical infrastructure data to prevent system failures and use predictive analytics to optimize resource use. Also, smart lighting systems in smart cities help conserve energy and improve safety through usage monitoring, adaptive lighting settings, and real-time data analysis. Overall, CLCP can provide a solution to efficiently manage massive streams of real-time data from various smart city systems.

6.8 Related Works

Over the last few years, ML techniques have been widely applied to wireless communications for several different tasks [39]. To our knowledge, using ML techniques for cross-link channel prediction is a novel contribution of this work and represents a field of research worth investigating more in detail. Still, it is possible to compare CLCP and other works from some of the research topics touched by this work.

Localization and Sensing. Typically, ML is used in *sensing* applications, including localization, human activity recognition, and motion detection [91, 133]. It could be argued that in some way CLCP also concerns *sensing* activity, even if in this case it is performed with the final goal of predicting the wireless link conditions of a different user rather than measuring some property of the surroundings. Since the seminal paper that proposed to derive environment information from Wi-Fi signals [10], one of the key challenges of sensing applications is the capability to work sufficiently well in new environments. Indeed, while traditional localization algorithms can work in different environments because they rely on physical properties such as AoA or ToF [111, 82, 9], this assessment becomes more complicated when ML techniques are used. Several works in literature have already tackled this problem by proposing learning frameworks based on environment- and user-independent features. For example, several works [98, 106] propose to process the CSI to extract the micro-Doppler features that filter out human movements from the overall effect of the surroundings. Preliminary results show very high accuracy in different rooms without the need for re-training of the model. However, all these techniques still require extensive validation by using heterogeneous datasets, collected in different scenarios. While most of the datasets available today are still using legacy IEEE 802.11n signals (e.g., [172]) collected with the discontinued

Intel 5300 chipset [59], newer datasets using high-dimensional IEEE 802.11ax signals have appeared [28, 97].

Deep Probabilistic Networks for Wireless Signals. Meanwhile, other learning frameworks have emerged that are intrinsically more robust to channel variations. For instance, EI [75] leverages adversarial networks to classify motion information embedded in the wireless signal. It uses a probabilistic learning model to extract general features (*i.e.*, not dependent on the environment or specific users) of the wireless signals when someone is moving. On the other hand, RF-EATS [56] leverages a probabilistic learning framework that adopts variational inference networks [113], in this case to sense food and liquids in closed containers with the back-scattered RF signals as an input. Both EI and RF-EATS build a model capable of generalizing to new, unseen environments. Similarly, CLCP’s goal is to learn the underlying dynamics of the wireless channel shared by different links by means of a deep probabilistic model. CLCP does not only decompose the observed wireless channel into a representation conveying environment-specific information, but it also integrates the representation with user-specific information to predict the wireless channels of a target (still unobserved) user.

Learning-based Channel Prediction. A growing body of work leverages various ML techniques for the broader goal of radio resource management. CSpy [123] uses a Support Vector Machine (SVM) to predict, on a single link, the channel with the strongest average magnitude, but does not venture into cross-link prediction at a subcarrier-level granularity, which modern wireless networks require in order to perform efficient OFDMA channel allocation for a group of users. Also, to manifest compression-based channel sounding for uplink-dominant massive-IoT networks, it requires extremely regular and frequent traffic patterns for every users, which is impractical. R2F2 [142] infers downlink CSI for a certain LTE mobile-base station link based on the path parameters of uplink CSI readings for that *same* link, using an optimization method in lieu of an ML-based approach. Similarly, [14] leverages a neural network to estimate path parameters, which, in turn, are used to predict across frequency bands. However, in 802.11ax, there is instead a different need and opportu-

nity: to predict *different* links' channels as recent traffic has used, in order to reduce channel estimation overhead, the opportunity CLCP targets.

Resource Allocation Algorithms. OFDMA scheduling has been extensively researched for LTE systems [20, 49, 154, 19]. In LTE, the channel is divided into equal-sized resource blocks, with independent scheduling on each subchannel, and base stations can allocate any combination of these blocks to a single UE. While this approach might seem applicable to 802.11ax, 802.11ax has stricter constraints on resource allocation: a user can only be assigned a single RU and the size of an RU is variable. Furthermore, unlike the flexible resource block locations in LTE, 802.11ax RUs are confined to specific positions. Hence, we need to develop a new scheduling solution that specifically addresses the unique characteristics of 802.11ax OFDMA.

Various schedulers have been developed to address the unique OFDMA structure of 802.11ax, which features varying RU sizes. For instance, researchers have adapted classic schedulers like proportional fair and shortest remaining processing time (SRPT) to meet 802.11ax specifications [15]. Another example is MMRU-ALLOC [37], a resource scheduler designed for 802.11ax that aims to minimize transmission duration. However, these approaches ignore that different RUs undergo varying levels of attenuation due to frequency-selective fading. An effective scheduler should account for varying signal quality across the channel, assigning RUs based on channel measurements rather than assuming uniform conditions. Wang *et al.* [148] proposed a max sum rate-based scheduler using greedy and recursive algorithms. While this approach considers frequency-selective fading, it does not account for the varying amounts of data users need to transmit, potentially leading to suboptimal scheduling. Tutelian *et al.* [139] developed a scheduler that jointly optimizes channel conditions and fairness. For each MCS, it prioritizes users and allocates the widest available RU to each user in order of priority. This process is repeated for all MCS options, with the best overall allocation selected. However, this method calculates channel conditions by aver-

aging quality across an entire 20 MHz band, which may not accurately represent frequency selectivity at a finer scale.

In contrast, our proposed algorithm considers both frequency-selective fading and the amount of data each user needs to transmit. It assigns users with smaller data amounts to smaller RUs and accounts for capacity at the subcarrier level, from the smallest 26-tone RUs to the largest 996-tone RU. Also, unlike [139], our algorithm allows for varying MCS among scheduled users, as permitted by 802.11ax, potentially leading to more optimal results.

6.9 Conclusion

This chapter presents in detail the first study to explore cross-link channel prediction to improve scheduling and resource allocation in WLAN networks using IEEE 802.11ax. To our knowledge, the CLCP technique presented in this paper is the first that employs a deep learning-based model to predict the wireless channel's properties across different links. By considering two large-scale indoor testbeds with up to 144 Wi-Fi channels and four different channel bandwidths, from 20 MHz to 160 MHz, we show that the proposed CLCP technique provides a 2x throughput gain with respect to baseline implementations and 30% throughput increment over existing prediction algorithms. Future works may assess the performance of CLCP in scenarios where the target environment exhibits significant changes after the deployment.

Chapter 7

Concluding Remarks

This thesis presents the design, implementation, and experimental evaluation of a smart and programmable network architecture that optimizes and learns wireless networks in real-time, tackling key challenges in NextG networks. Our work spans both the physical and link layers, providing comprehensive solutions to enhance network performance and reliability. At the physical layer, we have designed programmable smart surface systems and deployed them on buildings and vehicles to manipulate wireless environments. These surfaces relay high-frequency signals around obstacles, enabling ultra-reliable, high-bandwidth connections required for emerging real-time applications such as augmented and virtual reality, video conferencing, and volumetric video holograms. Also, we've developed surfaces that enhance the connectivity of satellite networks, allowing reliable and low-latency Internet service to rural areas that terrestrial networks can't reach. At the link layer, I developed a machine learning system that manages networking resources among massively deployed devices, which is crucial for applications, such as smart retailer, smart city and smart factory. The contributions of this thesis encompass the development of new high-frequency surfaces, rigorous theoretical analysis, software-hardware integration, extensive experimentation using the advanced wireless testbed, and novel cross-layer network design and protocols. Specifically, we make the following contributions:

1. **mmWall** is the first electronically near-360° steerable meta-surface that brings outdoor high-frequency 5G signals (mmWave), indoors. mmWave signals offers immense bandwidth but are highly directional and easily obstructed by walls, forcing mmWave traffic to offload onto low frequencies. With over 4,000 meta-materials, each under a millimeter in diameter, mmWall steers signals *through* the surface or reflects signals to bypass obstacles, significantly improving mmWave coverage. It conducts a rapid beam search by splitting and sweeping signals in multi-armed beams as a user moves. A mmWall deployment boosts signals throughout the room by up to 30 dB, even in half of the previously dead spots. Future work may implement a time-modulated control on the mmWall prototype for beam manipulation at multiple frequencies.
2. **Wall-Street** tackles the challenges of 5G mmWave roadside networks, such as frequent handovers, complex beam alignment, and signal attenuation caused by vehicle bodies. By independently controlling two beams, Wall-Street enables simultaneous serving cell data transfer and handover cell measurement, allowing seamless handovers without service interruption. Furthermore, it combines beams from both the old and new cells during handover, realizing a make-before-break mechanism at mmWave for the first time. Integrated into the COSMOS testbed and evaluated with multiple mobile clients inside a surface-equipped vehicle driving on a nearby road, Wall-Street demonstrates a 2× improvement in TCP throughput. We believe Wall-Street is the first step in scaling out the capacity of mmWave networks and paves the way for wider adoption of mmWave communications.
3. **Wall-E** addresses the challenges of satellite networks, where uplink and downlink transmission occur simultaneously but over different frequency bands. This electronically tunable surface programs dual-frequency radio at 10 and 15 GHz, refracting the satellite link directly into homes and to satellites. My design ensures angular reciprocity,

allowing simultaneous steering of downlink and uplink beams at the same angle, thus accelerating the beam alignment process.

4. **Monolith** proposes a holistic OAM-based ambient backscatter design for rank-deficient wireless environments. We theoretically analyze the OAM-MIMO multiplexing and provide a preliminary study on generating the OAM beam with different states. Future works may address a solution to demultiplex the diverged beams by capturing only a subsection of the received beam and to reduce the size of metasurface at the receiver side.
5. **CLCP** uses a ML technique to predict wireless conditions and accordingly allocates resources. A key insight is that modeling the radio environment in a shared wireless medium is feasible with sufficient background data. Inspired by how computer vision builds 3D models from photos taken at different angles, CLCP treats each channel observation from prior transmissions as a snapshot from a specific viewpoint, merging these views into a cohesive representation of the environment. This representation is then used to predict unobserved wireless links. The experimental evaluation of CLCP in a 144-node deployment demonstrates that CLCP doubles the throughput gain and increases battery life by up to 65%. Future works may assess the performance of CLCP in scenarios where the target environment exhibits significant changes after the deployment.

This thesis thus establishes the foundation for a programmable smart radio environment, yielding a paradigm shift in the design of NextG wireless networks. We envision future wireless systems harnessing low-cost, tunable, controllable surfaces integrated into our everyday surroundings - embedded in building interiors and exteriors, windows, and vehicles. They will actively shape the propagation of ambient RF signals, enabling endpoints such as base stations and handheld devices to become simpler and smaller, and realizing the full potential of NextG wireless networks.

Bibliography

- [1] 3GPP. Study on new radio access technology physical layer aspects. Technical Report (TR) 38.802, 3rd Generation Partnership Project (3GPP), 2017.
- [2] 3GPP. Study on new radio (nr) to support non-terrestrial networks. Technical Report (TR) 38.811, 3rd Generation Partnership Project (3GPP), 2017.
- [3] 3GPP. Nr; base station (bs) radio transmission and reception. Technical Report (TR) 38.104, 3rd Generation Partnership Project (3GPP), 2019.
- [4] 3GPP. Nr; user equipment (ue) radio transmission and reception; part 2: Range 2 standalone. Technical Report (TR) 38.101-2, 3rd Generation Partnership Project (3GPP), 2019.
- [5] 3GPP. Study on narrow-band internet of things (nb-iot) / enhanced machine type communication (emtc) support for non-terrestrial networks (ntn). Technical Report (TR) 36.763, 3rd Generation Partnership Project (3GPP), 2020.
- [6] 3GPP. 5G; NR; NR and NG-RAN Overall description; Stage-2. Technical Report (TR) 38.300, 3rd Generation Partnership Project (3GPP), January 2021.
- [7] Omid Abari, Dinesh Bharadia, Austin Duffield, and Dina Katabi. Enabling High-Quality Untethered Virtual Reality. In *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pages 531–544, 2017.
- [8] Omid Abari, Haitham Hassanieh, Michael Rodriguez, and Dina Katabi. Millimeter wave communications: From point-to-point links to agile network connections. In *ACM HotNets Workshop*, page 169–175, 2016.
- [9] Fadel Adib, Zachary Kabelac, and Dina Katabi. Multi-person localization via {RF} body reflections. In *12th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 15)*, pages 279–292, 2015.
- [10] Fadel Adib and Dina Katabi. See through walls with wifi! In *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*, pages 75–86, 2013.
- [11] M Albulet. Spacex non-geostationary satellite system. *FCC Application SAT-LOA2016111500118*, 2016.

- [12] Yaniv Azar, George N. Wong, Kevin Wang, Rimma Mayzus, Jocelyn K. Schulz, Hang Zhao, Felix Gutierrez, DuckDong Hwang, and Theodore S. Rappaport. 28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in New York city. In *IEEE International Conference on Communications (ICC)*, pages 5143–5147, June 2013. ISSN: 1938-1883.
- [13] Hareesh Bahuleyan, Lili Mou, Olga Vechtomova, and Pascal Poupart. Variational attention for sequence-to-sequence models. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1672–1682. Association for Computational Linguistics, 8 2018.
- [14] Arjun Bakshi, Yifan Mao, Kannan Srinivasan, and Srinivasan Parthasarathy. Fast and efficient cross band channel prediction using machine learning. In *MobiCom*, 2019.
- [15] Dmitry Bankov, Andre Didenko, Evgeny Khorov, and Andrey Lyakhov. Ofdma uplink scheduling in ieee 802.11 ax networks. In *IEEE Intl. Conf. on Comms. (ICC)*, pages 1–6. IEEE, 2018.
- [16] Federico Boccardi, Robert W. Heath, Angel Lozano, Thomas L. Marzetta, and Petar Popovski. Five disruptive technology directions for 5g. *IEEE Communications Magazine*, 52(2):74–80, 2014.
- [17] Shkelzen Cakaj. The parameters comparison of the “starlink” leo satellites constellation for different orbital shells. *Frontiers in Communications and Networks*, 2:643095, 2021.
- [18] Hailin Cao, Wang Zhu, Wenjuan Feng, and Jin Fan. Robust beamforming based on graph attention networks for irs-assisted satellite iot communications. *Entropy*, 24(3):326, 2022.
- [19] Eduardo Castaneda, Adao Silva, Atilio Gameiro, and Marios Kountouris. An overview on resource allocation techniques for multi-user mimo systems. *IEEE Communications Surveys & Tutorials*, 19(1):239–284, 2016.
- [20] Peter WC Chan and Roger S Cheng. Capacity maximization for zero-forcing mimo-ofdma downlink systems with multiuser diversity. *IEEE Transactions on Wireless Communications*, 6(5):1880–1889, 2007.
- [21] Ke Chen, Yijun Feng, Francesco Monticone, Junming Zhao, Bo Zhu, Tian Jiang, Lei Zhang, Yongjune Kim, Xumin Ding, Shuang Zhang, et al. A reconfigurable active Huygens metalens. *Advanced materials*, 29(17):1606422, 2017.
- [22] Kun Woo Cho, Marco Cominelli, Francesco Gringoli, Joerg Widmer, and Kyle Jamieson. Scalable multi-modal learning for cross-link channel prediction in massive iot networks. In *Proceedings of the Twenty-Fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, MobiHoc ’23, page 221–229, New York, NY, USA, 2023. Association for Computing Machinery.

- [23] Kun Woo Cho, Yasaman Ghasempour, and Kyle Jamieson. Towards dual-band reconfigurable metasurfaces for satellite networking. In *Proceedings of the 21st ACM Workshop on Hot Topics in Networks, HotNets '22*, page 17–23, New York, NY, USA, 2022. Association for Computing Machinery.
- [24] Kun Woo Cho, Srikar Kasi, and Kyle Jamieson. A low-power oam metasurface for rank-deficient wireless environments. In *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, pages 5695–5700, 2023.
- [25] Kun Woo Cho, Prasanthi Maddala, Ivan Seskar, and Kyle Jamieson. Wall-street: Smart surface-enabled 5g mmwave for roadside networking, 2024.
- [26] Kun Woo Cho, Mohammad H. Mazaheri, Jeremy Gummesson, Omid Abari, and Kyle Jamieson. Mmwall: A reconfigurable metamaterial surface for mmwave networks. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications, HotMobile '21*, page 119–125, New York, NY, USA, 2021. Association for Computing Machinery.
- [27] Kun Woo Cho, Mohammad H. Mazaheri, Jeremy Gummesson, Omid Abari, and Kyle Jamieson. mmWall: A Steerable, Transflective Metamaterial Surface for NextG mmWave Networks. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pages 1647–1665, 2023.
- [28] Marco Cominelli, Francesco Gringoli, and Francesco Restuccia. Exposing the csi: A systematic investigation of csi-based wi-fi sensing capabilities and limitations. In *2023 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 81–90. IEEE, 2023.
- [29] Erik Dahlman, Gunnar Mildh, Stefan Parkvall, Janne Peisa, Joachim Sachs, Yngve Selén, and Johan Sköld. 5g wireless access: requirements and realization. *IEEE Communications Magazine*, 52(12):42–47, 2014.
- [30] Linglong Dai, Bichai Wang, Min Wang, Xue Yang, Jingbo Tan, Shuangkaisheng Bi, Shenheng Xu, Fan Yang, Zhi Chen, Marco Di Renzo, Chan-Byoung Chae, and Lajos Hanzo. Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results. *IEEE Access*, 8:45913–45923, 2020.
- [31] Spyridon-Nektarios Daskalakis, John Kimionis, Ana Collado, Manos M Tentzeris, and Apostolos Georgiadis. Ambient fm backscattering for smart agricultural monitoring. In *2017 IEEE MTT-S International Microwave Symposium (IMS)*, pages 1339–1341. IEEE, 2017.
- [32] Philipp del Hougne, Mohammadreza F Imani, Timothy Sleasman, Jonah N Gollub, Mathias Fink, Geoffroy Lerosey, and David R Smith. Dynamic metasurface aperture as smart around-the-corner motion detector. *Scientific reports*, 8(1):6536, 2018.
- [33] Inigo Del Portillo, Bruce G Cameron, and Edward F Crawley. A technical comparison of three low earth orbit satellite constellation systems to provide global broadband. *Acta astronautica*, 159:123–135, 2019.

- [34] Haotian Deng, Chunyi Peng, Ans Fida, Jiayi Meng, and Y. Charlie Hu. Mobility Support in Cellular Networks: A Measurement Study on Its Configurations and Implications. In *Proceedings of the Internet Measurement Conference 2018*, pages 147–160, Boston MA USA, October 2018. ACM.
- [35] Xumin Ding, Zhuochao Wang, Guangwei Hu, Jian Liu, Kuang Zhang, Haoyu Li, Badreddine Ratni, Shah Nawaz Burokur, Qun Wu, Jiubin Tan, et al. Metasurface holographic image projection based on mathematical properties of Fourier transform. *Photonix*, 1(1):1–12, 2020.
- [36] Kate Duffy. Starlink has hit more than 1 million users despite a drop in download speeds. here’s what you need to know about the service. *Business Insider*, 2022.
- [37] Avik Dutta, Naman Gupta, Syamantak Das, and Mukulika Maity. Mmru-alloc: An optimal resource allocation framework for ofdma in ieee 802.11 ax. In *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–6. IEEE, 2020.
- [38] Ove Edfors and Anders J Johansson. Is orbital angular momentum (oam) based radio communication an unexploited area? *IEEE Transactions on Antennas and Propagation*, 60(2):1126–1131, 2011.
- [39] Yonina C Eldar, Andrea Goldsmith, Deniz Gündüz, and H Vincent Poor. *Machine learning and wireless communications*. Cambridge University Press, 2022.
- [40] Jose A Encinar, Manuel Arrebola, F Luis, and Giovanni Toso. A transmit-receive reflectarray antenna for direct broadcast satellite applications. *IEEE Transactions on Antennas and Propagation*, 59(9):3255–3264, 2011.
- [41] Jose A Encinar, L Sh Datashvili, J Agustín Zornoza, Manuel Arrebola, Manuel Sierra-Castañer, Jose Luis Besada-Sanmartin, Horst Baier, and Herve Legay. Dual-polarization dual-coverage reflectarray for space applications. *IEEE Transactions on Antennas and Propagation*, 54(10):2827–2837, 2006.
- [42] Ariel Epstein and George V. Eleftheriades. Passive lossless Huygens metasurfaces for conversion of arbitrary source field to directive radiation. *IEEE Transactions on Antennas and Propagation*, 62(11):5680–5695, 2014.
- [43] Ariel Epstein and George V Eleftheriades. Huygens’ metasurfaces via the equivalence principle: design and applications. *JOSA B*, 33(2):A31–A50, 2016.
- [44] BAF Esmail, HA Majid, ZZ Abidin, SH Dahlan, MKA Rahim, O Ayop, et al. New metamaterial structure with reconfigurable refractive index at 5G candidate band. *J. Optoelectron Adv M*, 21(1-2):101–107, 2019.
- [45] Huilong Fan, Zhan Yang, Shimin Wu, Xi Zhang, Jun Long, and Limin Liu. An efficient satellite resource cooperative scheduling method on spatial information networks. *Mathematics*, 9(24):3293, 2021.

- [46] FEDERAL COMMUNICATIONS COMMISSION. Ericsson ab pre-nr and nr base radio akrd901059-1.
- [47] FEDERAL COMMUNICATIONS COMMISSION. Sfg-aa100dc 5g access unit test report 1 samsung electronics.
- [48] FEDERAL COMMUNICATIONS COMMISSION. Title 47, code for federal regulations.
- [49] Guillem Femenias and Felip Riera-Palou. Scheduling and resource allocation in downlink multiuser mimo-ofdma systems. *IEEE Transactions on Communications*, 64(5):2019–2034, 2016.
- [50] H.T. Friis. A note on a simple transmission formula. *Proc. of the IRE*, 34(5):254–256, 1946.
- [51] Francesco Gringoli, Marco Cominelli, Alejandro Blanco, and Joerg Widmer. AX-CSI: Enabling CSI Extraction on Commercial 802.11ax Wi-Fi Platforms. In *Proceedings of the 15th ACM Workshop on Wireless Network Testbeds, Experimental Evaluation & CHaracterization*, WiNTECH’21, page 46–53. Association for Computing Machinery, 2021.
- [52] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, et al. Recent advances in convolutional neural networks. *Pattern recognition*, 77:354–377, 2018.
- [53] Bei Guo, Chenhao Sun, and Meixia Tao. Two-way passive beamforming design for ris-aided fdd communication systems. In *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2021.
- [54] Fengxian Guo, F. Richard Yu, Heli Zhang, Xi Li, Hong Ji, and Victor C. M. Leung. Enabling massive iot toward 6g: A comprehensive survey. *IEEE Internet of Things Journal*, 8(15):11891–11915, 2021.
- [55] Kai Guo, Qun Zheng, Zhiping Yin, and Zhongyi Guo. Generation of mode-reconfigurable and frequency-adjustable oam beams using dynamic reflective metasurface. *IEEE Access*, 8:75523–75529, 2020.
- [56] Unsoo Ha, Junshan Leng, Alaa Khaddaj, and Fadel Adib. Food and liquid sensing in practical environments using RFIDs. In *NSDI*, 2020.
- [57] Muhammad Kumail Haider, Yasaman Ghasempour, Dimitrios Koutsonikolas, and Edward W. Knightly. LiSteer: mmWave Beam Acquisition and Steering by Tracking Indicator LEDs on Wireless APs. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 273–288, New Delhi India, October 2018. ACM.

- [58] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. Predictable 802.11 packet delivery from wireless channel measurements. *ACM SIGCOMM CCR*, 40(4):159–170, 2010.
- [59] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM computer communication review*, 41(1):53–53, 2011.
- [60] Jiaqi Han, Long Li, Hao Yi, and Weimin Xue. Versatile orbital angular momentum vortex beam generator based on reconfigurable reflective metasurface. *Japanese journal of applied physics*, 57(12):120303, 2018.
- [61] Juzheng Han and Rushan Chen. Dual-band metasurface for broadband asymmetric transmission with high efficiency. *Journal of Applied Physics*, 130(3):034503, 2021.
- [62] Mark Handley. Delay is not an option: Low latency routing in space. In *Proceedings of the 17th ACM Workshop on Hot Topics in Networks, HotNets '18*, page 85–91, New York, NY, USA, 2018. Association for Computing Machinery.
- [63] Ahmad Hassan, Arvind Narayanan, Anlan Zhang, Wei Ye, Ruiyang Zhu, Shuowei Jin, Jason Carpenter, Z. Morley Mao, Feng Qian, and Zhi-Li Zhang. Vivisecting mobility management in 5G cellular networks. In *Proceedings of the ACM SIGCOMM 2022 Conference, SIGCOMM '22*, pages 86–100, New York, NY, USA, August 2022. Association for Computing Machinery.
- [64] Haitham Hassanieh, Omid Abari, Michael Rodriguez, Mohammed Abdelghany, Dina Katabi, and Piotr Indyk. Fast millimeter wave beam alignment. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, pages 432–445, Budapest Hungary, August 2018. ACM.
- [65] Ibrahim A. Hemadeh, Katla Satyanarayana, Mohammed El-Hajjar, and Lajos Hanzo. Millimeter-Wave Communications: Physical Channel Models, Design Considerations, Antenna Constructions, and Link-Budget. *IEEE Communications Surveys & Tutorials*, 20(2):870–913, 2018.
- [66] Geoffrey E Hinton. Training products of experts by minimizing contrastive divergence. *Neural computation*, 14(8):1771–1800, 2002.
- [67] Christopher L. Holloway, Edward F. Kuester, Joshua A. Gordon, John O’Hara, Jim Booth, and David R. Smith. An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials. *IEEE Antennas and Propagation Magazine*, 54(2):10–35, 2012.
- [68] Jakob Hoydis, Fayçal Aït Aoudia, Sebastian Cammerer, Merlin Nimier-David, Nikolaus Binder, Guillermo Marcus, and Alexander Keller. Sionna rt: Differentiable ray tracing for radio propagation modeling. *arXiv preprint arXiv:2303.11103*, 2023.

- [69] Sheng Hua, Yong Zhou, Kai Yang, Yuanming Shi, and Kunlun Wang. Reconfigurable intelligent surface for green edge inference. *IEEE Transactions on Green Communications and Networking*, 5(2):964–979, 2021.
- [70] Cheng Huang, Bo Sun, Wenbo Pan, Jianhua Cui, Xiaoyu Wu, and Xiangang Luo. Dynamical beam manipulation based on 2-bit digitally-controlled coding metasurface. *Scientific reports*, 7:42302, 2017.
- [71] IEEE Standard Part 11: P802.11ax/D4.0, February 2019.
- [72] IEEE Standard Part 11: 802.11p, Amendment 6: Wireless Access in Vehicular Environments, 2010.
- [73] Borja Imaz-Lueje, Daniel R Prado, Manuel Arrebola, and Marcos R Pino. Reflectarray antennas: A smart solution for new generation satellite mega-constellations in space communications. *Scientific Reports*, 10(1):1–13, 2020.
- [74] Ish Kumar Jain, Raghav Subbaraman, and Dinesh Bharadia. Two beams are better than one: Towards reliable and high throughput mmwave links. In *ACM SIGCOMM Conf.*, page 488–502, New York, NY, USA, 2021.
- [75] Wenjun Jiang, Chenglin Miao, Fenglong Ma, Shuochoao Yao, Yaqing Wang, Ye Yuan, Hongfei Xue, Chen Song, Xin Ma, Dimitrios Koutsonikolas, et al. Towards environment independent device free human activity recognition. In *MobiCom*, 2018.
- [76] Zhi Hao Jiang, Lei Kang, Wei Hong, and Douglas H Werner. Highly efficient broadband multiplexed millimeter-wave vortices from metasurface-enabled transmit-arrays of subwavelength thickness. *Physical Review Applied*, 9(6):064009, 2018.
- [77] Suraj Jog, Jiaming Wang, Junfeng Guan, Thomas Moon, Haitham Hassanieh, and Romit Roy Choudhury. Many-to-Many Beam Alignment in Millimeter Wave Networks. In *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pages 783–800, Boston, MA, 2019.
- [78] Mohammad Ali Khalighi and Murat Uysal. Survey on free space optical communication: A communication theory perspective. *IEEE Communications Surveys*, 16(4):2231–2258, 2014.
- [79] Minseok Kim, Namjo Ahn, and Song Min Kim. NR-Surface: NextG-ready μ W-reconfigurable mmWave Metasurface. In *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2024.
- [80] Minyoung Kim and Vladimir Pavlovic. Recursive inference for variational autoencoders. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 19632–19641. Curran Associates, Inc., 2020.
- [81] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. In *Proceedings of the 2nd International Conference on Learning Representations, ICLR*, 2014.

- [82] M Kotaru, K Joshi, D Bhardia, and S Katti. SpotFi: Decimeter level localization using Wi-Fi. In *SIGCOMM*, 2015.
- [83] Yuanjie Li, Qianru Li, Zhehui Zhang, Ghufan Baig, Lili Qiu, and Songwu Lu. Beyond 5G: Reliable Extreme Mobility Management. In *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication*, pages 344–358, Virtual Event USA, July 2020. ACM.
- [84] Zhuqi Li, Yuanchao Shu, Ganesh Ananthanarayanan, Longfei Shangguan, Kyle Jamieson, and Paramvir Bahl. Spider: A multi-hop millimeter-wave network for live video analytics. In *2021 IEEE/ACM Symp. on Edge Computing (SEC)*, pages 178–191, 2021.
- [85] Baiyang Liu, Sirong Li, Yejun He, Yin Li, and Sai-Wai Wong. Generation of an orbital-angular-momentum-mode-reconfigurable beam by a broadband 1-bit electronically reconfigurable transmitarray. *Physical Review Applied*, 15(4):044035, 2021.
- [86] Lixiang Liu, Xueqian Zhang, Mitchell Kenney, Xiaoqiang Su, Ningning Xu, Chunmei Ouyang, Yunlong Shi, Jianguang Han, Weili Zhang, and Shuang Zhang. Broadband metasurfaces with simultaneous control of phase and amplitude. *Advanced materials*, 26(29):5031–5036, 2014.
- [87] Mingkai Liu, David A Powell, Yair Zarate, and Ilya V Shadrivov. Huygens’ metadevices for parametric waves. *Physical Review X*, 8(3):031077, 2018.
- [88] Ruiqi Liu, Meng Hua, Ke Guan, Xiping Wang, Leyi Zhang, Tianqi Mao, Di Zhang, Qingqing Wu, and Abbas Jamalipour. 6g enabled advanced transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–17, 2024.
- [89] Vincent Liu, Aaron Parks, Vamsi Talla, Shyamnath Gollakota, David Wetherall, and Joshua R Smith. Ambient backscatter: Wireless communication out of thin air. *ACM SIGCOMM Computer Communication Review*, 43(4):39–50, 2013.
- [90] Yang Liu, Jun Zhao, Ming Li, and Qingqing Wu. Intelligent reflecting surface aided miso uplink communication network: Feasibility and power minimization for perfect and imperfect csi. *IEEE Transactions on Communications*, 69(3):1975–1989, 2020.
- [91] Yongsun Ma, Gang Zhou, and Shuangquan Wang. Wifi sensing with channel state information: A survey. *ACM Computing Surveys (CSUR)*, 52(3):1–36, 2019.
- [92] George R. MacCartney, Theodore S. Rappaport, and Sundeep Rangan. Rapid Fading Due to Human Blockage in Pedestrian Crowds at 5G Millimeter-Wave Frequencies. In *IEEE Global Communications Conference, GLOBECOM*, pages 1–7, Singapore, December 2017. IEEE.
- [93] George R. MacCartney, Theodore S. Rappaport, Shu Sun, and Sijia Deng. Indoor office wideband millimeter-wave propagation measurements and channel models at 28 and 73 GHz for ultra-dense 5G wireless networks. *IEEE Access*, 3:2388–2424, 2015.

- [94] Fuchun Mao, Ming Huang, Jingjing Yang, Chengfu Yang, Tinghua Li, and Jialin Zhang. Capacity performance of wireless oam-based massive mimo system. *Progress In Electromagnetics Research M*, 82:149–156, 2019.
- [95] Daniel Martinez-de Rioja, Eduardo Martinez-de Rioja, Yolanda Rodriguez-Vaqueiro, Jose A Encinar, and Antonio Pino. Multibeam reflectarrays in ka-band for efficient antenna farms onboard broadband communication satellites. *Sensors*, 21(1):207, 2020.
- [96] Bho Matthiesen, Emil Björnson, Elisabeth De Carvalho, and Petar Popovski. Intelligent reflecting surface operation under predictable receiver mobility: A continuous time propagation model. *IEEE Wireless Communications Letters*, 10(2):216–220, 2020.
- [97] Francesca Meneghello, Cheng Chen, Carlos Cordeiro, and Francesco Restuccia. Ieee 802.11ax csi dataset for human activity recognition, 2023.
- [98] Francesca Meneghello, Domenico Garlisi, Nicolò Dal Fabbro, Ilenia Tinnirello, and Michele Rossi. Sharp: Environment and person independent activity recognition with commodity ieee 802.11 access points. *IEEE Transactions on Mobile Computing*, 2022.
- [99] Marco Mezzavilla, Menglei Zhang, Michele Polese, Russell Ford, Sourjya Dutta, Sundeeep Rangan, and Michele Zorzi. End-to-End Simulation of 5G mmWave Networks. *IEEE Communications Surveys & Tutorials*, 20(3):2237–2263, 2018.
- [100] Joanna Najder and Krzysztof Sońnica. Quality of orbit predictions for satellites tracked by slr stations. *Remote Sensing*, 13(7):1377, 2021.
- [101] Arvind Narayanan, Eman Ramadan, Jason Carpenter, Qingxu Liu, Yu Liu, Feng Qian, and Zhi-Li Zhang. A First Look at Commercial 5G Performance on Smartphones. In *Proceedings of The Web Conference (WWW)*, pages 894–905, Taipei Taiwan, April 2020. ACM.
- [102] Arvind Narayanan, Xumiao Zhang, Ruiyang Zhu, Ahmad Hassan, Shuwei Jin, Xiao Zhu, Xiaoxuan Zhang, Denis Rybkin, Zhengxuan Yang, Zhuoqing Morley Mao, Feng Qian, and Zhi-Li Zhang. A variegated look at 5G in the wild: performance, power, and QoE implications. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, pages 610–625, New York, NY, USA, August 2021. ACM.
- [103] Mahyar Nemati, Behrouz Maham, Shiva Raj Pokhrel, and Jinho Choi. Modeling RIS empowered outdoor-to-indoor communication in mmWave cellular networks. *IEEE Transactions on Communications*, 69(11):7837–7850, 2021.
- [104] Yunzhe Ni, Feng Qian, Taide Liu, Yihua Cheng, Zhiyao Ma, Jing Wang, Zhongfeng Wang, Gang Huang, Xuanzhe Liu, and Chenren Xu. POLYCORN: Data-driven Cross-layer Multipath Networking for High-speed Railway through Composable Schedulerlets. In *20th USENIX Symposium on Networked Systems Design and Implementation (USENIX NSDI)*, pages 1325–1340, 2023.

- [105] Kentaro Nishimori, Koshiro Kitao, and Tetsuro Imai. Interference-based decode and forward scheme using relay nodes in heterogeneous networks. *International Journal of Antennas and Propagation*, 2012, 2012.
- [106] Kai Niu, Fusang Zhang, Xuanzhi Wang, Qin Lv, Haitong Luo, and Daqing Zhang. Understanding wifi signal frequency features for position-independent gesture sensing. *IEEE Transactions on Mobile Computing*, 21(11):4156–4171, 2022.
- [107] Anton Patyuchenko, Carolina Tienda, Marwan Younis, Sebastian Bertl, Paco Lopez-Dekker, and Gerhard Krieger. Digital beamforming sar interferometer based on a multi-beam reflectarray antenna. In *EUSAR 2014; 10th European Conference on Synthetic Aperture Radar*, pages 1–4. IEEE, 2014.
- [108] Hao Peng and Xiaoli Bai. Improving orbit prediction accuracy through supervised machine learning. *Advances in Space Research*, 61(10):2628–2646, 2018.
- [109] Yao Peng, Longfei Shangguan, Yue Hu, Yujie Qian, Xianshang Lin, Xiaojiang Chen, Dingyi Fang, and Kyle Jamieson. Plora: A passive long-range data network from ambient lora transmissions. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pages 147–160, 2018.
- [110] Carl Pfeiffer and Anthony Grbic. Metamaterial Huygens’ surfaces: Tailoring wave fronts with reflectionless sheets. *Phys. Rev. Lett.*, 110:197401, May 2013.
- [111] Alejandro Blanco Pizarro, Joan Palacios Beltrán, Marco Cominelli, Francesco Gringoli, and Joerg Widmer. Accurate ubiquitous localization with off-the-shelf ieee 802.11ac devices. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys ’21*, page 241–254. Association for Computing Machinery, 2021.
- [112] Daniel Rodríguez Prado, Adrián Campa, Manuel Arrebola, Marcos R Pino, José A Encinar, and Fernando Las-Heras. Design, manufacture, and measurement of a low-cost reflectarray for global earth coverage. *IEEE antennas and wireless propagation Letters*, 15:1418–1421, 2015.
- [113] Yunchen Pu, Zhe Gan, Ricardo Henao, Xin Yuan, Chunyuan Li, Andrew Stevens, and Lawrence Carin. Variational autoencoder for deep learning of images, labels and captions. In *Advances in neural information processing systems*, pages 2352–2360, 2016.
- [114] Kun Qian, Lulu Yao, Xinyu Zhang, and Tse Nga Ng. MilliMirror: 3D printed reflecting surface for millimeter-wave coverage expansion. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking (MobiCom)*, pages 15–28, Sydney NSW Australia, October 2022. ACM.
- [115] Sundeeep Rangan, Theodore S. Rappaport, and Elza Erkip. Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges. *Proceedings of the IEEE*, 102(3):366–385, March 2014.

- [116] Maryam Eslami Rasekh, Zhinus Marzi, Yanzi Zhu, Upamanyu Madhow, and Haitao Zheng. Noncoherent mmWave path tracking. In *ACM HotMobile Workshop*, page 13–18, 2017.
- [117] Dipankar Raychaudhuri, Ivan Seskar, Gil Zussman, Thanasis Korakis, Dan Kilper, Tingjun Chen, Jakub Kolodziejcki, Michael Sherman, Zoran Kostic, Xiaoxiong Gu, Harish Krishnaswamy, Sumit Maheshwari, Panagiotis Skrimponis, and Craig Gutterman. Challenge: COSMOS: A city-scale programmable testbed for experimentation with advanced wireless. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking, MobiCom '20*, pages 1–13, New York, NY, USA, April 2020. Association for Computing Machinery.
- [118] David Rotshild and Amir Abramovich. Ultra-wideband reconfigurable x-band and ku-band metasurface beam-steerable reflector for satellite communications. *Electronics*, 10(17):2165, 2021.
- [119] Sudhir K. Routray, Richa Tengshe, Abhishek Javali, Sutapa Sarkar, Laxmi Sharma, and Aritri D. Ghosh. Satellite based iot for mission critical applications. In *2019 International Conference on Data Science and Communication (IconDSC)*, pages 1–6, 2019.
- [120] Yasir Saifullah, Qinzhuo Chen, Guo-Min Yang, Abu Bakar Waqas, and Feng Xu. Dual-band multi-bit programmable reflective metasurface unit cell: design and experiment. *Optics Express*, 29(2):2658–2668, 2021.
- [121] Edgar Schonfeld, Sayna Ebrahimi, Samarth Sinha, Trevor Darrell, and Zeynep Akata. Generalized zero-and few-shot learning via aligned variational autoencoders. In *CVPR*, pages 8247–8255, 2019.
- [122] David Schurig, JJ Mock, BJ Justice, Steven A Cummer, John B Pendry, AF Starr, and David R Smith. Metamaterial electromagnetic cloak at microwave frequencies. *Science*, 314(5801):977–980, 2006.
- [123] S Sen, B Radunović, J Lee, and K-H Kim. CSpy: Finding the best quality channel without probing. In *MobiCom*, 2013.
- [124] Wei-Liang Shen, Kate Ching-Ju Lin, Ming-Syan Chen, and Kun Tan. SIEVE: Scalable user grouping for large MU-MIMO systems. In *Infocom*, pages 1975–1983. IEEE, 2015.
- [125] Timothy Slesman, Michael Boyarsky, Mohammadreza F Imani, Thomas Fromenteze, Jonah N Gollub, and David R Smith. Single-frequency microwave imaging with dynamic metasurface apertures. *JOSA B*, 34(8):1713–1726, 2017.
- [126] Zhenyu Song, Longfei Shangguan, and Kyle Jamieson. Wi-Fi Goes to Town: Rapid Picocell Switching for Wireless Transit Networks. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, pages 322–334, New York, NY, USA, August 2017. ACM.

- [127] Adrian Spurr, Jie Song, Seonwook Park, and Otmar Hilliges. Cross-modal deep variational hand pose estimation. In *CVPR*, pages 89–98, 2018.
- [128] Shao-Hua Sun, Minyoung Huh, Yuan-Hong Liao, Ning Zhang, and Joseph J Lim. Multi-view to novel view: Synthesizing novel views with self-learned confidence. In *ECCV*, pages 155–171, 2018.
- [129] Sanjib Sur, Ioannis Pefkianakis, Xinyu Zhang, and Kyu-Han Kim. WiFi-assisted 60 GHz wireless networks. In *ACM MobiCom Conf.*, page 28–41, New York, NY, USA, 2017.
- [130] Michele Tamagnone, Christophe Craeye, and Julien Perruisseau-Carrier. Comment on ‘encoding many channels on the same frequency through radio vorticity: first experimental test’. *New Journal of Physics*, 14(11):118001, 2012.
- [131] Fabrizio Tamburini, Elettra Mari, Anna Sponselli, Bo Thidé, Antonio Bianchini, and Filippo Romanato. Encoding many channels on the same frequency through radio vorticity: first experimental test. *New Journal of Physics*, 14(3):033001, 2012.
- [132] Fabrizio Tamburini, Elettra Mari, Bo Thidé, Cesare Barbieri, and Filippo Romanato. Experimental verification of photon angular momentum and vorticity with radio techniques. *Applied Physics Letters*, 99(20):204102, 2011.
- [133] Sheng Tan, Yili Ren, Jie Yang, and Yingying Chen. Commodity wifi sensing in ten years: Status, challenges, and opportunities. *IEEE Internet of Things Journal*, 9(18):17832–17843, 2022.
- [134] Xin Tan, Zhi Sun, Dimitrios Koutsonikolas, and Josep M. Jornet. Enabling Indoor Mobile Millimeter-wave Networks Based on Smart Reflect-arrays. In *IEEE Conference on Computer Communications (INFOCOM)*, pages 270–278, Honolulu, HI, April 2018. IEEE.
- [135] Wankai Tang, Ming Zheng Chen, Xiangyu Chen, Jun Yan Dai, Yu Han, Marco Di Renzo, Yong Zeng, Shi Jin, Qiang Cheng, and Tie Jun Cui. Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement. *IEEE Trans. on Wireless Comms.*, 20(1):421–439, 2021.
- [136] Wankai Tang, Xiangyu Chen, Ming Zheng Chen, Jun Yan Dai, Yu Han, Marco Di Renzo, Shi Jin, Qiang Cheng, and Tie Jun Cui. Path loss modeling and measurements for reconfigurable intelligent surfaces in the millimeter-wave frequency band. *IEEE Trans. on Comms. (Early Access)*, 2022.
- [137] Andrea Tassi, Malcolm Egan, Robert J. Piechocki, and Andrew Nix. Modeling and Design of Millimeter-Wave Networks for Highway Vehicular Communication. *IEEE Transactions on Vehicular Technology*, 66(12):10676–10691, December 2017.
- [138] D Tse and P Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

- [139] Sergei Tutelian, Dmitry Bankov, Dmitri Shmelkin, and Evgeny Khorov. Ieee 802.11 ax ofdma resource allocation with frequency-selective fading. *Sensors*, 21(18):6099, 2021.
- [140] Andrea Vallecchi, Ekaterina Shamonina, and Christopher J Stevens. Analytical model of the fundamental mode of 3D square split ring resonators. *J. of Applied Physics*, 125(1):014901, 2019.
- [141] Greg Van Houdt, Carlos Mosquera, and Gonzalo Nápoles. A review on the long short-term memory model. *Artificial Intelligence Review*, 53(8):5929–5955, 2020.
- [142] Deepak Vasisht, Swarun Kumar, Hariharan Rahul, and Dina Katabi. Eliminating Channel Feedback in Next-Generation Cellular Networks. In *SIGCOMM*, 2016.
- [143] Deepak Vasisht, Jayanth Shenoy, and Ranveer Chandra. L2d2: Low latency distributed downlink for leo satellites. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*, SIGCOMM '21, page 151–164, New York, NY, USA, 2021. Association for Computing Machinery.
- [144] Miguel Ángel Vázquez, MR Bhavani Shankar, Charilaos I Kourogiorgas, Pantelis-Daniel Arapoglou, Vincenzo Icolari, Symeon Chatzinotas, Athanasios D Panagopoulos, and Ana I Pérez-Neira. Precoding, scheduling, and link adaptation in mobile interactive multibeam satellite systems. *IEEE Journal on Selected Areas in Communications*, 36(5):971–980, 2018.
- [145] Suresh Venkatesh, Xuyang Lu, Hooman Saeidi, and Kaushik Sengupta. A high-speed programmable and scalable terahertz holographic metasurface based on tiled CMOS chips. *Nature Electronics*, 3(12):785–793, December 2020.
- [146] E.J. Violette, R.H. Espeland, R.O. DeBolt, and F.K. Schwing. Millimeter-wave propagation at street level in an urban environment. *IEEE Trans. on Geoscience and Remote Sensing*, 26(3):368–380, 1988.
- [147] J Wang, J Tang, Z Xu, Y Wang, G Xue, X Zhang, and D Yang. Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach. In *Infocom*, 2017.
- [148] Kaidong Wang and Konstantinos Psounis. Efficient scheduling and resource allocation in 802.11 ax multi-user transmissions. *Computer Communications*, 152:171–186, 2020.
- [149] Xuan Wang, Pei-Yuan Qin, Anh Tuyen Le, Hao Zhang, Ronghong Jin, and Y Jay Guo. Beam scanning transmitarray employing reconfigurable dual-layer Huygens element. *IEEE Transactions on Antennas and Propagation*, 70(9):7491–7500, 2022.
- [150] Teng Wei, Anfu Zhou, and Xinyu Zhang. Facilitating robust 60 GHz network deployment by sensing ambient reflectors. In *USENIX NSDI Symp.*, pages 213–226, March 2017.

- [151] Alan E Willner, Zhe Zhao, Cong Liu, Runzhou Zhang, Haoqian Song, Kai Pang, Karapet Manukyan, Hao Song, Xinzhou Su, Guodong Xie, et al. Perspectives on advances in high-capacity, free-space communications using multiplexing of orbital-angular-momentum beams. *APL Photonics*, 6(3):030901, 2021.
- [152] Timothy Woodford, Kun Qian, and Xinyu Zhang. Metasight: High-resolution nlos radar with efficient metasurface encoding. In *Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems, SenSys '23*, page 308–321, New York, NY, USA, 2024. Association for Computing Machinery.
- [153] Mike Wu and Noah Goodman. Multimodal generative models for scalable weakly-supervised learning. In *Advances in Neural Information Processing Systems*, pages 5575–5585, 2018.
- [154] Xiao Xiao, Xiaoming Tao, and Jianhua Lu. Energy-efficient resource allocation in lte-based mimo-ofdma systems with user rate constraints. *IEEE Transactions on Vehicular Technology*, 64(1):185–197, 2014.
- [155] Guodong Xie, Zhe Zhao, Yan Yan, Long Li, Yongxiong Ren, Nisar Ahmed, Yinwen Cao, Asher J Willner, Changjing Bao, Zhe Wang, et al. Demonstration of tunable steering and multiplexing of two 28 ghz data carrying orbital angular momentum beams using antenna array. *Scientific reports*, 6(1):1–9, 2016.
- [156] Xiufeng Xie, Xinyu Zhang, and Karthikeyan Sundaresan. Adaptive feedback compression for mimo networks. In *MobiCom*, pages 477–488, 2013.
- [157] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. mD-Track: Leveraging Multi-Dimensionality for Passive Indoor Wi-Fi Tracking. In *MobiCom*, 2019.
- [158] Yunchou Xing, Frederick Vook, Eugene Visotsky, Mark Cudak, and Amitava Ghosh. Raytracing-based system performance of intelligent reflecting surfaces at 28 GHz. In *IEEE Intl. Conf. on Comms.*, pages 498–503, 2022.
- [159] Dongzhu Xu, Anfu Zhou, Xinyu Zhang, Guixian Wang, Xi Liu, Congkai An, Yiming Shi, Liang Liu, and Huadong Ma. Understanding Operational 5G: A First Measurement Study on Its Coverage, Performance and Energy Consumption. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, pages 479–494, New York, NY, USA, July 2020. ACM.
- [160] Chunhua Xue, Qun Lou, and Zhi Ning Chen. Broadband double-layered huygens’ metasurface lens antenna for 5g millimeter-wave systems. *IEEE Transactions on Antennas and Propagation*, 68(3):1468–1476, 2019.
- [161] Tian Yang, Xiaoming Liu, Chen Wang, Zetan Liu, Jingbo Sun, and Ji Zhou. Polarization conversion in terahertz planar metamaterial composed of split-ring resonators. *Optics Communications*, 472:125897, 2020.
- [162] Yan Yang and Hao Wang. Multi-view clustering: A survey. *Big Data Mining and Analytics*, 1(2):83–107, 2018.

- [163] Alison M Yao and Miles J Padgett. Orbital angular momentum: origins, behavior and applications. *Advances in optics and photonics*, 3(2):161–204, 2011.
- [164] Keke Ying, Zhen Gao, Shanxiang Lyu, Yongpeng Wu, Hua Wang, and Mohamed-Slim Alouini. GMD-based hybrid beamforming for large reconfigurable intelligent surface assisted millimeter-wave massive MIMO. *IEEE Access*, 8:19530–19539, 2020.
- [165] Xianghao Yu, Dongfang Xu, and Robert Schober. Miso wireless communication systems via intelligent reflecting surfaces. In *2019 IEEE/CIC International Conference on Communications in China (ICCC)*, pages 735–740, China, 2019. IEEE.
- [166] R. Ivan Zelaya, William Sussman, Jeremy Gummesson, Kyle Jamieson, and Wenjun Hu. LAVA: fine-grained 3D indoor wireless coverage for small IoT devices. In *ACM SIGCOMM Conf.*, page 123–136, 2021.
- [167] Lei Zhang, Xiao Qing Chen, Shuo Liu, Qian Zhang, Jie Zhao, Jun Yan Dai, Guo Dong Bai, Xiang Wan, Qiang Cheng, Giuseppe Castaldi, et al. Space-time-coding digital metasurfaces. *Nature Communications*, 9(1):1–11, 2018.
- [168] Xueqian Zhang, Zhen Tian, Weisheng Yue, Jianqiang Gu, Shuang Zhang, Jianguang Han, and Weili Zhang. Broadband terahertz wave deflection based on c-shape complex metamaterials with phase discontinuities. *Advanced Materials (Deerfield Beach, Fla.)*, 25(33):4567–4572, 2013.
- [169] Hang Zhao, Rimma Mayzus, Shu Sun, Mathew Samimi, Jocelyn K. Schulz, Yaniv Azar, Kevin Wang, George N. Wong, Felix Gutierrez, and Theodore S. Rappaport. 28 GHz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in New York city. In *IEEE International Conference on Communications (ICC)*, pages 5163–5167, June 2013. ISSN: 1938-1883.
- [170] Xiaopeng Zhao, Zhenlin An, Qingrui Pan, and Lei Yang. Nerf2: Neural radio-frequency radiance fields. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, pages 1–15, 2023.
- [171] Beixiong Zheng, Shaoe Lin, and Rui Zhang. Intelligent reflecting surface-aided leo satellite communication: Cooperative passive beamforming and distributed channel estimation. *arXiv preprint arXiv:2201.02913*, 2022.
- [172] Yue Zheng, Yi Zhang, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. Zero-effort cross-domain gesture recognition with wi-fi. In *Proceedings of the 17th annual international conference on mobile systems, applications, and services*, pages 313–325, 2019.