RICE UNIVERSITY

MU-MIMO WLANs in Diverse Bands and Environments

by

Narendra Anand

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
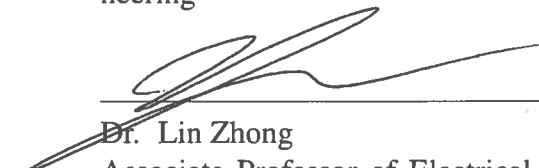REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy
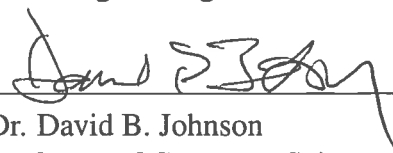
APPROVED, THESIS COMMITTEE:

Dr. Edward W. Knightly, *Chair*
Professor of Electrical and Computer Engineering and Computer Science

Dr. Ashutosh Sabharwal
Professor of Electrical and Computer Engineering

Dr. Lin Zhong
Associate Professor of Electrical and Computer Engineering

Dr. David B. Johnson
Professor of Computer Science

HOUSTON, TEXAS
APRIL 2015

ABSTRACT

MU-MIMO WLANs in Diverse Bands and Environments

by

Narendra Anand

Multi-user MIMO (MU-MIMO) is a precoding technique that allows for
an Access Point (AP) to transmit data to multiple receivers in parallel resulting
in an overall capacity increase. However, achieving these gains requires signif-
icant overhead in first choosing the subset of users to serve and measuring the
Channel State Information (CSI) between that user group and the AP. Man-
agement of this overhead and other Media Access Control (MAC) decisions
are key to the performance of MU-MIMO transmissions. Previous solutions
attempt all-purpose approaches which attempt to work well regardless of de-
ployment environment or frequency band.

However, through thorough analysis of MU-MIMO transmissions with re-
spect to band and environment, we show that an all-purpose solution will not
allow for MU-MIMO transmissions to reach their full potential. In fact, lever-
aging the differences in MU-MIMO transmission characteristics for different
bands and environments to develop separate protocols will allow for increased
overall system performance.

To address transmission scenarios with high channel variability, we present
Pre-sounding **U**ser and **M**ode selection **A**lgorithm (PUMA), a transmission
mode and user selection protocol that leverages the characteristics of the highly
variable 2.4/5.8 GHz transmission scenarios for efficient MU-MIMO overhead

amortization. While PUMA results in significant MU-MIMO capacity gains for 2.4/5.8 GHz indoor environments, it will not allow UHF-band indoor and outdoor MU-MIMO transmissions to reach their full potential. To that end, we design and implement **F**eedback **R**emoval with **O**pportunistic **Z**ero-overhead channel **E**stimatio**N** (FROZEN), a protocol that harnesses the channel stability of the UHF band to eliminate sounding overhead by relying on CSI measurements from previously received uplink packets.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor, Dr. Edward Knightly, whose guidance has been invaluable in producing this thesis. Additionally, I would like to thank my thesis committee members for their comments, evaluation, and their time spent examining my work.

I am also grateful for my colleagues and collaborators (and now my dear friends) in the Rice Network Group without whom I would not have made it through graduate school. I would specifically like to thank Ryan Guerra, Oscar Bejarano, and Adriana Flores for their friendship, technical input, and willingness to take time out of their day to join me on the perch for conversations frivolous and meaningful. Also, I would like to thank Oscar for making sure that I didn't miss a single paperwork deadline set by Rice's Registrar Office or the ECE department. If it weren't for him, I'd still be resubmitting my 599 report.

I owe my success to my parents N.K. and Veena Anand, and my brother, Shankara Anand. Without them, I would have never thought any of this possible. Their continual love, support, and occasional goading has made me into who I am today.

Finally, I would like to thank Elaine Savoy. Your love has kept me going through this process. Thank you for being my co-pilot. This work marks the end of the first stage of my life and I look forward to beginning the next with you.

CONTENTS

LIST OF ACRONYMS

LIST OF SYMBOLS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Multi-user MIMO (MU-MIMO) achieves substantial capacity gains by using precoding to support multiple, concurrent data streams to a group of clients. Precoding comprises of computing the transmitter's antenna gains and phases from the Channel State Information (CSI), i.e., the channel matrix in which each element represents the magnitude and phase offset for each transmitter-receiver antenna path. In this way, each receiver can simultaneously decode its streams [1]. Moreover, the recent IEEE 802.11ac amendment promises multi-Gb/s Wireless LANs (WLANs) via down-link MU-MIMO with up to 8 transmit antennas at the AP [2, 3].

However, the potential gains of MU-MIMO transmissions come at the cost of channel sounding overhead. The aforementioned CSI requires an overhead intensive measurement process that must occur before a multi-stream transmission can occur.

For MU-MIMO transmitters to achieve their full potential, the transmitter must execute novel Media Access Control (MAC) protocols that carry out efficient transmissions, i.e., the significant channel sounding overhead must be amortized with respect to the transmitted data so that the overall system throughput approaches the theoretical capabilities of this multi-antenna technology.

Existing protocols that attempt to increase MU-MIMO throughput do so by attempting

to control the channel sounding overhead using methods such as reducing the rate of channel sounding (e.g., [4]) or reducing the amount of data fed back from the sounding process (e.g., [5]).

However, these methods fail to consider or otherwise exploit the different characteristics of MU-MIMO transmissions in different frequency bands or environments. While all-purpose protocols seem ideal, they do not allow for MU-MIMO to reach its full potential in these different scenarios.

To that end, this work focuses on the development of separate MU-MIMO protocols for different band and environment combinations. These protocols leverage the beneficial characteristics of these scenarios to avoid the detrimental characteristics of these scenarios and increase overall system throughput.

The contributions of this thesis are as follows:

First, we conduct a thorough measurement study of MU-MIMO transmissions with respect to band and environment to fully understand the differences between each scenario. This analysis drives our protocol development in that it provides a unique understanding of the coupled effect of these scenarios on key MU-MIMO performance determining characteristics[1].

Second, we present **P**re-sounding **U**ser and **M**ode selection **A**lgorithm (PUMA), a transmission mode and user selection protocol that leverages the characteristics of 2.4/5.8 GHz indoor transmission scenarios for efficient MU-MIMO overhead amortization.

Finally, we present **F**eedback **R**emoval with **O**pportunistic **Z**ero-overhead channel **E**stimatio**N** (FROZEN), a protocol that harnesses the channel stability of the indoor or outdoor UHF band to eliminate sounding overhead by relying on CSI measurements from previously received uplink packets.

The remainder of this thesis is organized as follows: Chapter 2 provides a background

---

[1]Which are User Separability and Temporal Correlation, defined and discussed in Sec. 3.2.

of MU-MIMO transmissions, Chapter 3 discusses our MU-MIMO measurement study, Chapter 4 and Chapter 5 present PUMA and FROZEN respectively, Chapter 6 discusses related work, and finally Chapter 7 concludes the thesis.

MU-MIMO BACKGROUND

In this chapter we briefly discuss MU-MIMO transmission techniques and the key performance determining factors for MU-MIMO.

2.1 Concurrent Multi-user Transmissions

Multi-user Multiple-Input, Multiple-Output (MIMO) is a multi-antenna transmission technique that allows a transmitter to spatially reuse a wireless channel by transmitting multiple concurrent streams. This is achieved in two steps: First, each data stream is multiplied by a length M vector of complex steering weights (where M is the number of transmit antennas) resulting in M phase twisted copies of each data stream. Each user's steering weight vector ($\vec{w}$) is represented in the columns of the steering weight matrix W Second, each receiver's set of M copies are summed together at each antenna to construct K parallel data streams (where K is the number of concurrent receivers) emanating from M antennas as shown in Fig. 2.1.

The aforementioned steering weights are chosen such that the interference between the parallel streams is minimal (ideally zero). To compute these weights, the transmitter must first measure the CSI matrix (H) where each element corresponds to the magnitude and
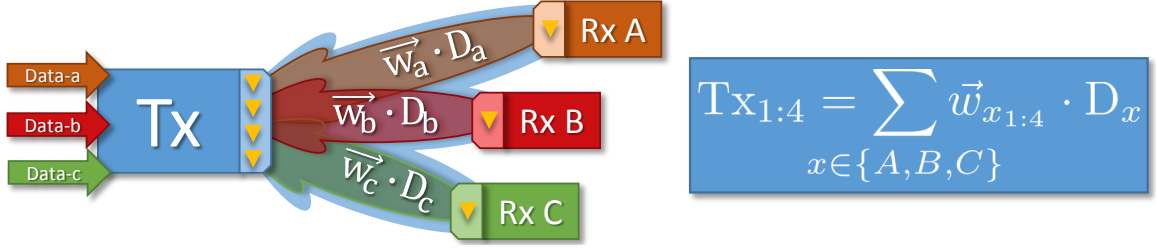
Figure 2.1: Depiction of MU-MIMO precoding.

phase difference between each transmit and receive antenna. The optimal method of constructing the steering matrix is Dirty Paper Coding (DPC) [6]; however, its complexity makes it unfeasible to implement. Instead, a method known as Zero-forcing Beamforming (ZFBF) [1] is shown to approach the optimal performance of DPC while employing a computationally feasible weight matrix calculation method, the pseudo-inverse (represented as $(\cdot)^{\dagger}$) given by

$$W = H^{\dagger} = H^* \cdot (H \cdot H^*). \tag{2.1}$$

A key element of ZFBF is the zero-interference condition which is a direct result of the pseudo-inverse. Because $W = H^{\dagger}$, $\vec{h}_i \vec{w}_j = 0$ for $i \neq j$ meaning that the interference from user i's stream on user j is nulled and vice versa. ZFBF precodes the transmitted data streams such that the combined wireless channel between the transmitter and the receivers (H) is separated. If ZFBF works perfectly, we can express the precoded transmission $(W \cdot \text{Tx})$ as:

$$W \cdot \text{Tx} \xrightarrow[H]{\text{Transmit}} H \cdot (W \cdot \text{Tx})$$
$$= \cancel{H} \cdot (\cancel{W} \cdot \text{Tx}) = \text{Tx} \tag{2.2}$$

In this work, we focus on the zero-forcing beamforming technique for MU-MIMO.

2.2   MU-MIMO Performance Limitations

The key to the success of this precoding operation is that $H \cdot W$ is the identity matrix so the transmitted streams are received separately at each receiver. We focus on two characteristics of H that can degrade the performance of this precoding operation: an ill-conditioned H [7] or an out-dated H [8].

An ill-conditioned H matrix renders matrix inversion inaccurate [9] and thus $H \cdot W$ is far less likely to equal the identity matrix I. This results in inter-stream interference degrading the received signal strength of a data stream to its intended receiver [8]. Ill-conditioned H matrices are a result of receiver channel correlation, an environment and frequency dependent characteristic that will be discussed in Sec. 3.2.1.

Out-dated H matrices are a direct result of the latency between the measurement of the H matrix and the transmission of the W precoded data streams. Increased time between the measurement of H and the transmission of $W \cdot \mathrm{Tx}$, results in a higher probability of incorrect transmit precoding. Essentially, the transmitter measures H_t and then calculates $W_t = H_t^\dagger$. However, the subsequent precoded transmission is $H_{(t+\Delta t)} \cdot W_t$, which may not equal I. Whether or not $H_t = H_{(t+\Delta t)}$ is based on environmental variability and user mobility; and, like channel conditioning, is also an environment and frequency dependent characteristic that will be discussed in the following sections.

While a large number of studies in Sec. 6.2 have characterized the indoor and outdoor propagation environment for the purpose of network planning and algorithm design, few are applicable to evaluating MU-MIMO performance and most have focused on a single frequency band. This renders measurement studies of different frequencies and radio technologies difficult to compare.

Additionally, the complexity required for real-time implementations of multi-carrier MU-MIMO is prohibitive for today's Software-Defined Radio (SDR) platforms [10], thus providing a challenge to empirical measurement of MU-MIMO performance.

CHARACTERIZATION OF MU-MIMO TRANSMISSIONS

In this chapter, we discuss our MU-MIMO measurement study. The purpose of our study is to understand the MU-MIMO propagation differences between the UHF and 2.4/5.8 GHz band along with the characteristics of performing MU-MIMO transmissions in indoor or outdoor environments.

In Sec. 3.1, we discuss the implementation of our custom built MU-MIMO test hardware (the Wideband UHF Radio Card (WURC) Array), in Sec. 3.2 we analyze existing channel models to gain additional insight into MU-MIMO performance in the aforementioned band and environment combinations, and finally in Sec. 3.3 we discuss the over-the-air, experimental results of our measurement study.

3.1 WURC Array Implementation

In this section, we describe the hardware platforms and software frameworks that we designed and deployed in order to enable an experimental evaluation of UHF MU-MIMO. First, we design and implement a new SDR analog front-end designed for high-power, wideband Single-Input, Single-Output (SISO) UHF operation. Using the newly-developed radio and the WARPv3 SDR hardware platform, we then develop an integrated frequency-

diverse MU-MIMO system. Finally, we describe our extensive modifications to existing experimental software frameworks that allow us to rapidly gather channel state information and perform over-the-air MU-MIMO evaluations.

3.1.1 Wideband UHF Radio Card (WURC)

WURC is a new SDR analog front-end designed to enable high-power, long-range experiments and hardware prototypes in the UHF frequency band. It is designed for modularity and compactness, with the goal of enabling prototyping of new MAC and PHY enhancements for UHF and Industrial, Scientific, and Medical (ISM)-band applications [11].

It connects to the host FPGA board via an HSMC or FMC (with custom adapter board) daughtercard slot, and provides a 12-bit digital baseband quadrature interface to the host system, while permitting in-field reconfiguration of RF analog parameters such as channel bandwidth and center frequency between 300-3800 MHz, though it is currently optimized and calibrated for transmissions between 470-798 MHz, and 2400-2500 MHz.

3.1.1.1 WURC High-Level System Design

MU-MIMO systems generally require a large number of independent transmit and receive RF chains on the base station to generate multiple spatial streams. In addition, a large number of distributed client nodes are required to serve as the mobile user stations. In order to simplify the manufacturing and management of a large number of radios, WURC is designed to be modular, with calibration/control libraries and board-dependent calibration tables stored locally on each daughter-card on a micro-controller. This eases the requirements for integration with a host platform and makes the radios completely interchangeable.
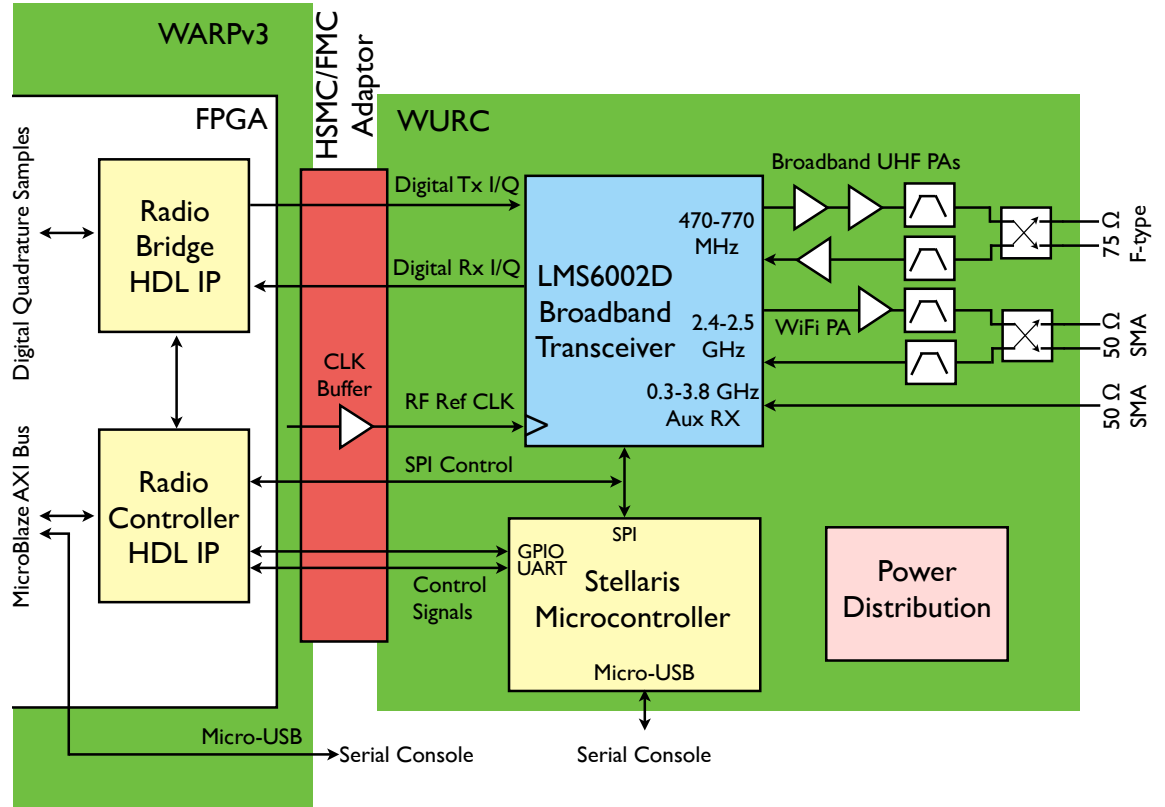
Figure 3.1: Block diagram of WURC module on a host WARPv3 board.

3.1.1.2 Power Amplifier Design and Verification

In order to operate as an opportunistic transmitter in the UHF band and adapt to various channel bandwidths, spectrum availability, and regulatory domains across the world, WURC is designed to operate at arbitrary channel bandwidths from 1.5 to 28 MHz and carrier frequencies ranging from 300-3800 MHz. This presents a challenge for a high-power RF design since power amplifiers and their associated impedance matching networks are generally optimized for a narrow frequency band.

A common technique for designing high-power analog front-ends is to build multiple switched amplification and filtering chains, each optimized for a narrow band. However, when the system operating frequency range spans multiple octaves, space and cost constraints require that each chain support a wide range of frequencies. In the design

of WURC, we target two optimized transmit and receive chains for 470-698 MHz and 2400-2500 MHz, chosen because these two bands allow unlicensed operation and are invaluable for research and testing. In addition, a wide-band balun transformer enables a 300-3800 MHz receiver port that can serve as a wideband spectrum sensor, if required.

Since the bandwidth of an RF chain is generally proportional to $\Delta f/f$, common techniques for designing and implementing discrete power transfer networks (such as multi-section Chebyshev transformers [12]) either cannot meet design requirements for passband flatness or result in non-realizable circuits when applied to bandpass designs spanning a large frequency range like 470-698 MHz.

In order to address this problem, we implemented a wideband linear power-transfer network utilizing real-frequency techniques [13] for the UHF front end. We target a design goal of transmit powers up to 30 dBm from 300 to 750 MHz, the maximum power currently allowed by the Federal Communications Commission (FCC) in the United States for unlicensed operation [14]. The 2.4 GHz ISM transmit chain provides up to 27 dBm between 2400-2500 MHz. The RF chain of WURC provides up to 30 dB of dynamic transmit gain, and up to 61 dB of dynamic receive gain, which when combined with its on-board Low-Noise Amplifier (LNA) can provide up to 83 dB of receive gain for improved sensitivity, although noise figure considerations generally limit this application to 72 dBm.

While this design was confirmed at the early design stage with SPICE simulation models, early prototypes demonstrated that package parasitics in the lumped-element broadband power transfer chain were not modeled by the ideal SPICE simulator. These parasitics severely impaired the implemented high-frequency gain response and required a more advanced model and simulation technique to correctly predict their effect. Re-modeling the RF chain in the SpectreRF circuit simulator utilizing empirical S-parameter models resulted in a more accurate simulation, allowing package parasitics to be compensated for in the lumped-element design.

In order to verify the correctness of the implemented design and understand how man-ufacturing process variation might effect the output frequency response of multiple RF chains in a MU-MIMO system, we built a Python-based batch interface to the WURC's serial UART in order to sweep transmit frequencies while simultaneously controlling a bench-top vector signal analyzer to measure the output power. We implemented a digital frequency synthesizer within the digital baseband reference design in order to generate a constant-power complex sinusoid for ease of measurement.
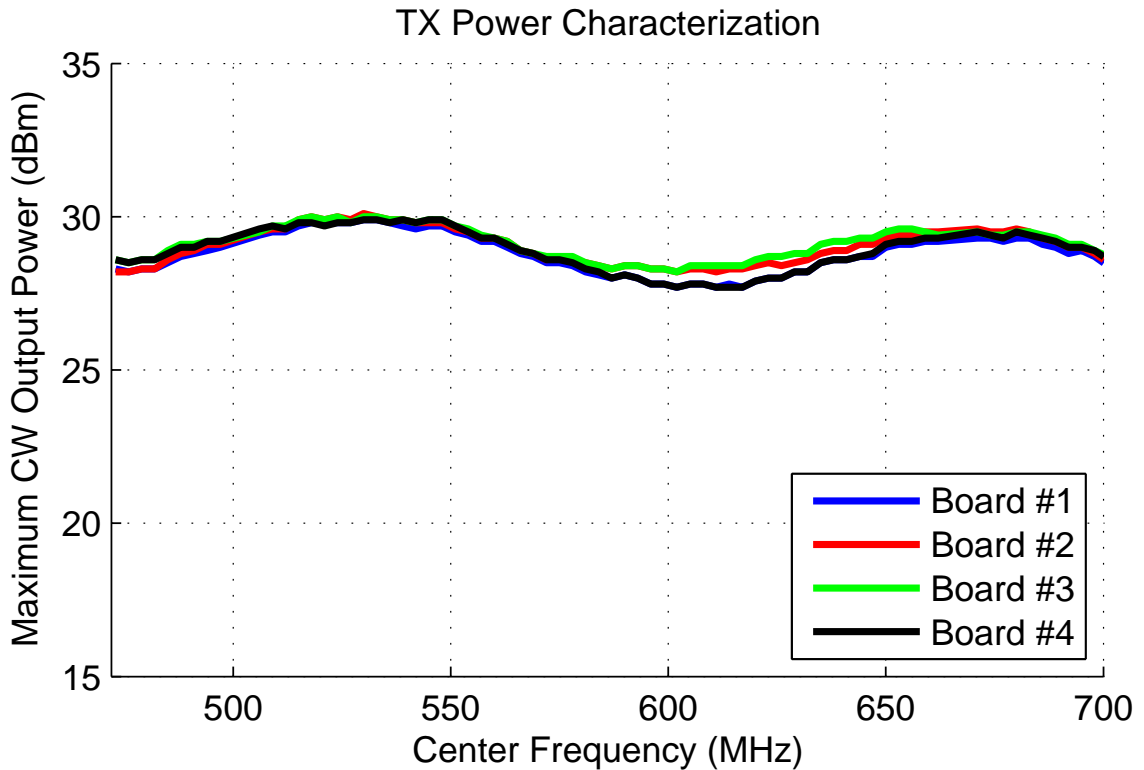
Figure 3.2: Process variation across multiple fabricated WURC boards has a small effect on output power.

The process variation plot in Fig. 3.2 was generated by increasing the output transmit gain of each WURC at each center frequency until its output PA began to saturate. This is the delivered output power of WURC near the 1dB compression point of the RF chain. Notably, the process variation across different boards is less than 1 dB, with passband ripple on the order of 2.5 dB. This means that multiple RF chains will maintain similar

output power across the entire UHF frequency range.

3.1.1.3 Radio Architecture

WURC uses a *direct-conversion* quadrature transceiver architecture based around the Lime Microsystems LMS6002D "field-programmable" transceiver IC in order to minimize size, implementation complexity, and energy-consumption [15]. Because of this, we are able to power the high-power RF chain from the FMC/HSMC-compliant daughtercard slot, further decreasing size and complexity.

All 12-bit DACs/ADCs, programmable analog anti-aliasing channel filters, frequency synthesizers, and direct-conversion mixers are integrated on a single chip while the rest of the board contains power amplifiers and filters, antenna diversity DPDT switch, power distribution, and a microcontroller (Fig. 3.1). We designed and tested fast-switching control circuits on the discrete amplification stages that allow the system to operate as a TDD transceiver with a switching time of less than 7 μs, or an FDD system with independent transmit and receive fractional-N frequency synthesizers.

Clocking. Since the transmit and receive chains in a MU-MIMO base station require precise phase synchronization, WURC was designed to draw RF reference clocks from the host digital baseband board as in Fig. 3.3. We placed an additional RF reference and sampling clock buffer on the FMC/HSMC adaptor rather than on the daughtercard itself so that the designed system can scale up to four WURCs driven from a single host FPGA with synchronized clocks; however, we only implement a single-radio adapter at this time.

Control and Calibration. An on-board micro-controller provides a simple, scriptable, two-wire UART or USB UART interface to a host system for command and control of analog parameters such as center frequency, transmit power, and analog channel bandwidth, while providing full read/write configuration register access to the transceiver.

We designed embedded libraries complete with calibration macros that offload com-
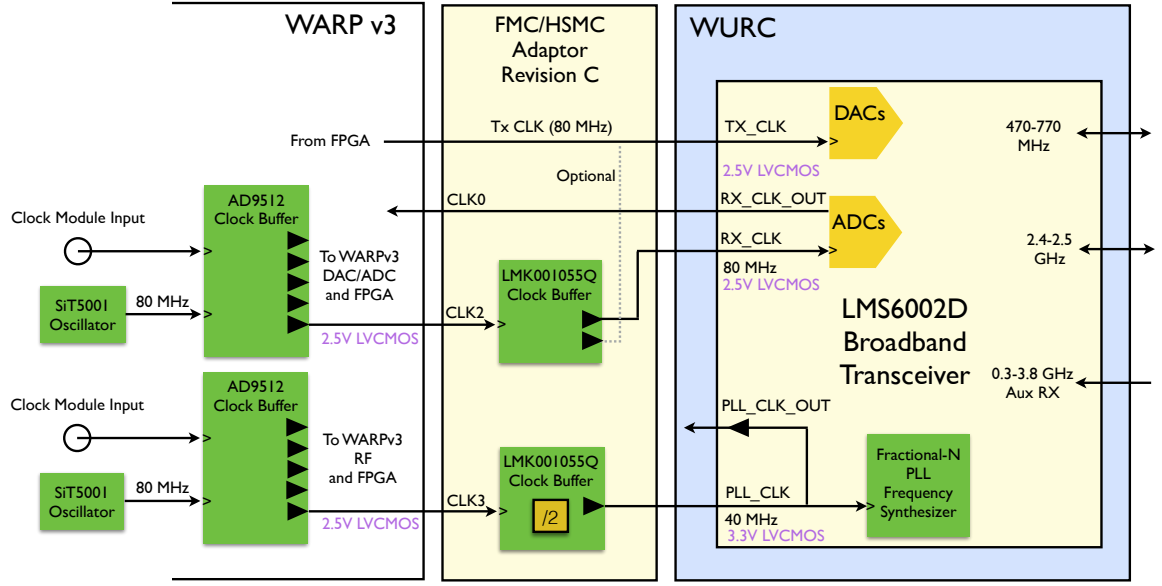
Figure 3.3: Source-synchronous sampling clocks and RF reference clocks are buffered in stages, permitting daisy-chaining and future fanout to multiple radios.

plex computation from the host system and handle the loading of stored factory calibration values for transmit and receive baseband IQ-imbalance and local oscillator feed-through compensation. In addition, we developed automated "factory" calibration procedures that allow us to rapidly calibrate a large number of WURCs for field deployment with minimal setup time.

Each WURC is a highly-integrated SDR front-end module that provides unprecedented capabilities in a small form factor, enabling a wide range of experimental trials and system implementations with excellent RF flexibility.

3.1.2 WURC Array

In order to evaluate MU-MIMO transmissions at various carrier frequencies and node topologies, we integrate WURC and four WARPv3 modules into a coherent 4-radio array.

Clock Sharing. The MU-MIMO WURC array combines four WARPv3 boards and 4 WURC daughtercards into a single prototype base station providing combined sample and

RF-reference clock synchronization, power, and structural support. Synchronization of reference clocks for ADC sampling and RF frequency synthesizers is required for coherent beamforming and is accomplished by forwarding a daisy-chained reference clock from one master WARPv3 baseband board to the others in the array. All radios derive their sampling and RF reference clocks from this forwarded clock and thus remain phase-synchronized.

Antennas. Most studies of UHF propagation involve large, directional antennas intended for signal reception over many kilometers. This is because optimal signal reception and transmission requires antennas of at least 1/2 wavelength to generate a resonating standing wave. On the other hand, a WLAN deployment utilizing UHF frequencies may wish to keep the size of the base station somewhat limited, particularly for indoor deployments. For our experiments, we utilize off-the-shelf passive, omni-directional 3 dBi DTV antennas (August DTA240) that would provide the largest range of coverage with minimal dependance on direction. In our experimental platform (Fig. 3.4), it is actually the dual-band 2.4/5.8 GHz band antennas (L-com HG2458-5RD-RSP with 3 dBi and 5 dBi gain, respectively) that are larger in size.

This type of omni-directional antenna array is ideal for indoor MU-MIMO as it provides many opportunities for multipath reflections [10]. In order to guarantee the required channel diversity, each antenna was spaced at least 1/2 wavelength for its respective transmit frequency.

3.1.3 Software Framework

In addition to the development of custom hardware to meet our design requirements, we build upon or modify a number of existing applications in order to develop an experimental framework for the WURC MU-MIMO array.

(a) WURC.   (b) WURC Array Internals.   (c) WURC Array Antennas.
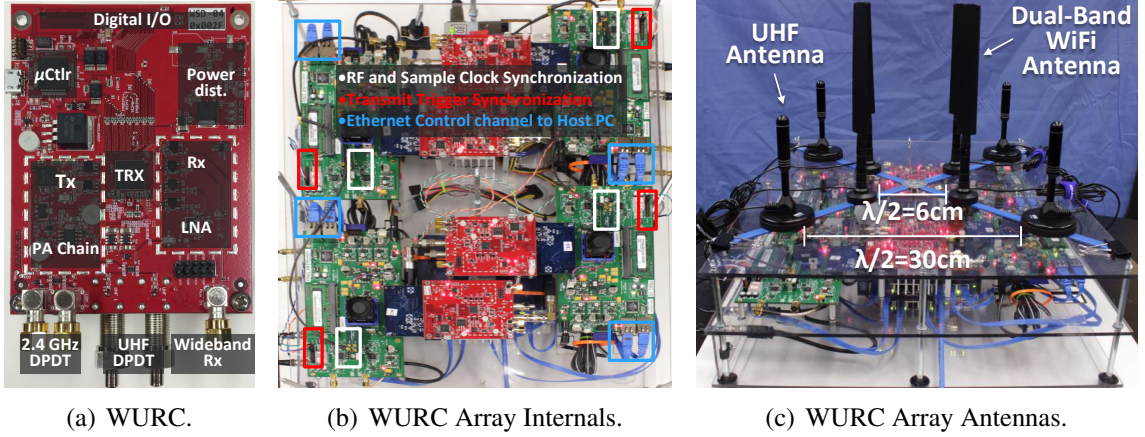
Figure 3.4: Wideband UHF Radio Card (WURC) hardware platform.

3.1.3.1   WARPLab

The WARPLab 7 framework for WARP hardware provides a means to pre-compute base-band signals in MATLAB, load transmit sample buffers into an array of WARP boards, and then trigger a simultaneous RF transmission of all buffered signals via a back-end ethernet network or a GPIO trigger [16]. Similarly, an arbitrary number of radios can be configured to perform Automatic Gain Control (AGC) and store their received RF samples in buffers for off-line retrieval and processing.

We extend WARPLab's object-oriented framework with additional classes and methods to support the WURC's interfaces. This system provides a powerful workflow for UHF PHY prototyping and measurement studies for multi-antenna systems.

Measurement Speed. WARPLab 7 contains a number of transport improvements that result in the ability to perform near-real-time experiments by rapidly performing cycles of: precompute, load, transmit/receive, fetch, and process on the order of 2.5 ms. A fast central coordinator using jumbo ethernet frames for transporting IQ buffers and a compiled MATLAB-mex transport layer can operate at per-packet time intervals. We observed that extra switches between WARPLab nodes produce measurable switching delay and recommend the use of long ethernet cables and minimization of the number of ethernet hops for

backhaul.

While powerful, the primary drawbacks of WARPLab is that it requires a central coordinator connected via gigabit ethernet switches, and real-time protocol implementations generally require processing at sub-packet timescales. These two factors hinder long-distance or mobile experiments.

SINR Measurement Technique. In order to overcome these limitations yet still accurately measure the MU-MIMO channel, we employ a WARPLab-based MU-MIMO transmission framework that is based on measuring received SINR and then computing the Shannon capacity to estimate the achievable rate of a transmission system. This is accomplished by a measurement technique adapted from [10] and shown in Fig. 3.5. Here, the transmitter beamforms sections of the transmission packet independently to accurately measure the SINR.
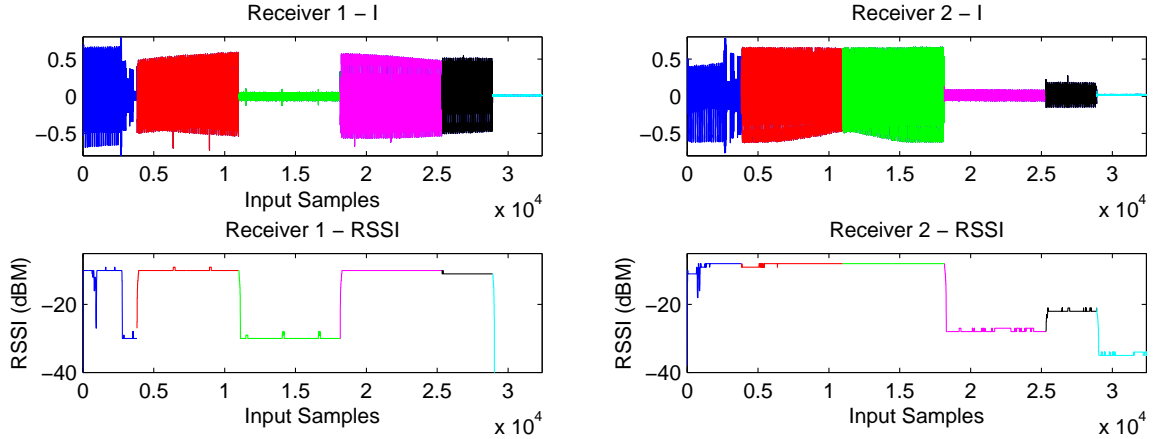
Figure 3.5: Example RSSI Measurement used in achievable capacity calculation.

In the depicted 4x2 transmission example, the transmitter first sends an LTS preamble for timing synchronization (blue) and then performs a MU-MIMO transmission to both users (red). In the following two sections (green, purple), the transmitter sequentially zeroes out the steering vector to each receiver in order to measure noise and interference at each receiver during a MU-MIMO transmission. In the example 4x2 case, this becomes

a single-user beamformed transmission, however in the 4x3 or 4x4 case, two or three receivers would be beamformed to during this measurement.

Thus, the difference between the full MU-MIMO transmission containing both signal, noise, and interference at each receiver (red) and the transmission containing just interference and noise at the zeroed-out receiver (green or purple) is each transmitter's SINR. From there, we can compute aggregate Shannon Capacity as:

$$C = log_2(1 + \text{SINR}) \tag{3.1}$$

3.1.3.2 Real-Time 802.11a/g-Like Reference Design

We realize a real-time 802.11a/g-interoperable design utilizing the WARPv3 802.11 Reference Design and WURC to transmit over UHF frequencies, with modifications to provide 10 and 5 MHz channels [17]. We develop custom HDL for the radio interface, AGC, and digital filtering necessary for a real-time broadband system, and integrate the hardware and software design with the WARPv3 802.11 Reference Design. This system implements a real-time layer-2 wireless bridge utilizing an 802.11a/g AP and STA design with a completely open network stack.

In particular, the real-time capabilities of the 802.11 reference design are leveraged to provide fine-grained continuous channel estimates from multiple transmitting antennas in order to directly measure the MU-MIMO channel capacity instantaneously and over a long period of time.

3.1.3.3 Framework Enhancements

In order to enable long-range MIMO channel sounding by a large number of mobile nodes, we make the following enhancements and modifications to the WARP frameworks de-

scribed in Sec. 3.1.3.2 and 3.1.3.1.

Hardware Integration. We adapt both the WARPLab and 802.11 reference design to work seamlessly with the WURC hardware in place of normal WARP daughter cards. From the perspective of the digital baseband, the analog front-end is transparent, which allows interchangeable analog PHYs to be used with the same digital PHY/MAC for fair comparison. This is especially useful for the MU-MIMO comparison study as it controls for a large number of variables in the radio MAC and PHY chain.

Channel Bandwidth. The 802.11 reference design operates in a 20 MHz channel bandwidth. In order to enable a UHF transmission to fit within one or two contiguous UHF channels of 6 MHz, we modify the 802.11 reference design to operate at 10 and 5 MHz channel bandwidths in compliance with the 802.11 standard. This is accomplished by halving the data sampling rate with added programmable decimation filters and adjusting MAC parameters and receiver DSP blocks to match.

Automatic Gain Control. The range of received power in realistic deployments is sufficiently large that an AGC subsystem is required to guarantee the robust and accurate reception of wideband channel sounding packets, particular when wireless nodes are mobile. Reference designs from the WARP project rely on external power detectors and autocorrelation to detect incoming packets and estimate a target receive gain setting, whereas an external power detector would require additional external circuitry on WURC.

Instead, we design a custom real-time digital loop in hardware to provide AGC convergence within 5.6 μs as required by the 20 MHz 802.11 PLCP and utilizing only the ADC output for packet detection and power estimation. AGC is an enabling technique required to ensure that channel measurement samples have the proper resolution. This is guaranteed when the received signal strength at the ADC input falls within the dynamic range of the ADC (the ADC's Effective Number of Bits (ENOB) is 10 [15]).

Fig. 3.6 depicts the performance of the implemented power estimator and AGC sub-
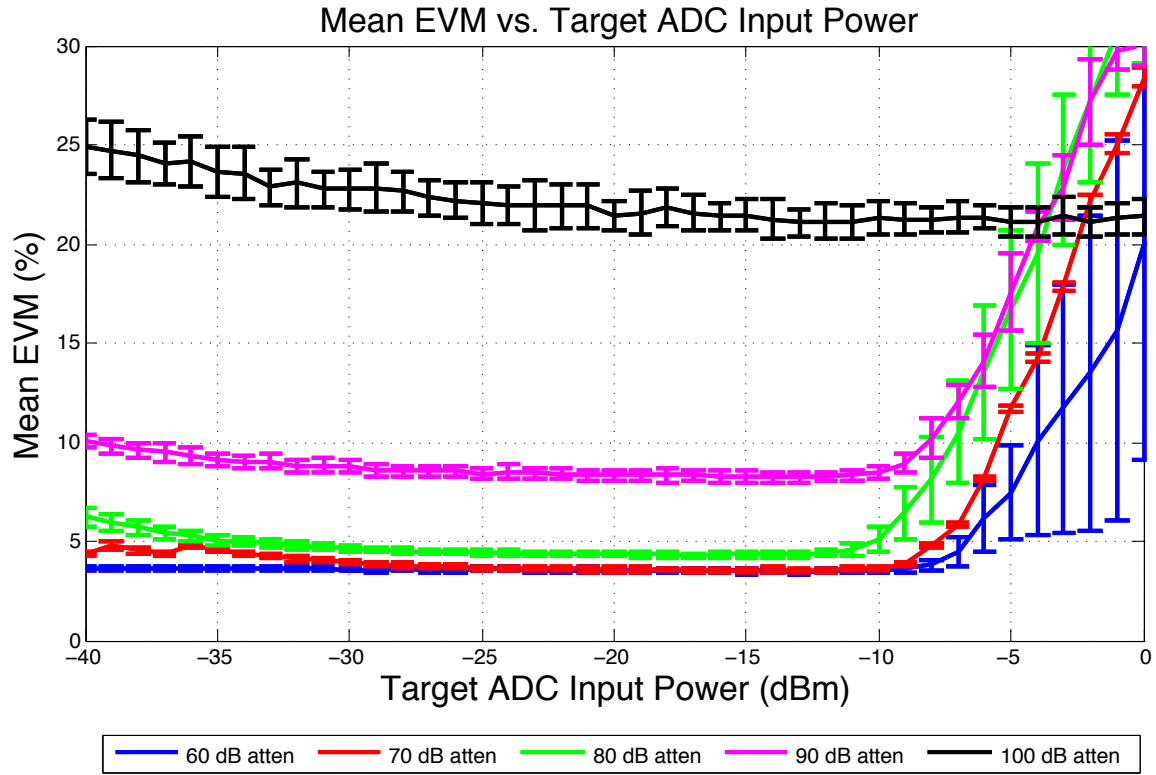
Figure 3.6: Verification of the implemented received power estimator and AGC operation. Measured with Tx gain at 25 dB, with 16-QAM OFDM modulation.

system design by reporting a series of experiments over a cable between two WURC nodes with a variable attenuator. The transmit gain is fixed to 25 dB and a 802.11g-like packet with random data payload and 16-QAM OFDM modulation is constructed in MATLAB and transmitted over the cable using the WARPLab framework developed in Sec. 3.1.3.1. The received packet without Forward Error Correction (FEC) is decoded and its received EVM is calculated as the mean across subcarriers and OFDM symbols of the normalized distance between the received decoded symbol and the intended decoded symbol.

Both the RF path attenuation and target ADC input power are varied under these conditions, resulting in the plot shown in Fig. 3.6, where the error bars represent one standard deviation across 50 trials. The top plot shows that the ADC operating target of -26 to -13 dBm is optimal for received EVM under a wide range of input powers. For each attenuation value, the bowl of the EVM curve represents the lower bound on the system's

16-QAM receive EVM, with the right-most bound of this range limited by saturation at the ADC and the left-most bound is determined by the system noise floor and quantization error. As expected, high signal attenuation of 100 dB results in a decrease in SINR and thus, minimum achievable EVM. We therefore fix our target ADC input power to -18 dBm in order to ensure that channel measurement packets are detected and received without quantization error and with maximum precision.

Channel Sounding. While the legacy 802.11 design calculates and stores channel state information as required by its OFDM channel equalizers, this information is generally discarded after packet reception. The channel estimation extracted from each received 802.11 PLCP header [3] provides a complete CSI estimation matrix that can be used as a single-antenna sounding event. We modify the physical layer of the 802.11 reference design to treat each of a series of transmitted PLCP headers as separate "packets" for the purpose of CSI measurement from multiple transmitting antennas.

Our custom sounding "packet" is a brief 802.11g-like signal containing PLCP header for packet detection, AGC convergence, and symbol timing extraction. The payload is just long enough to provide error detection bits and identifying information about the transmitter so that the transmitting antenna can be identified. Due to the small size of this sounding packet, it is not compliant with the requirement that 802.11 packets contain an 802.11 and link-layer header. Therefore, we modify the MAC software to pass all packets regardless of valid header or fields to the Ethernet interface for processing.

We construct this special sounding packet in MATLAB and preconfigure the WURC array, running our WARPLab modification, to transmit these packets continuously staggered in time as shown in Figure 3.7. Tests show that WARPLab continuous-transmit mode remains synchronous over long periods of time if the boards are clock synchronized. We provide sufficient spacing between sounding packets to allow the 802.11 PHY to process the previous packet and reset, and we find that the WARPLab buffer size of 32768 samples

over 819.2 μs is sufficient to capture channel variation even at higher frequencies.

We combine this structure with a set of multiple listening nodes that process these channel sounding packets and can then store them for later retrieval. A ten-minute packet trace for a single antenna can run over 1 GB in size, so substantial buffering and disk I/O speed is required for the recording nodes.
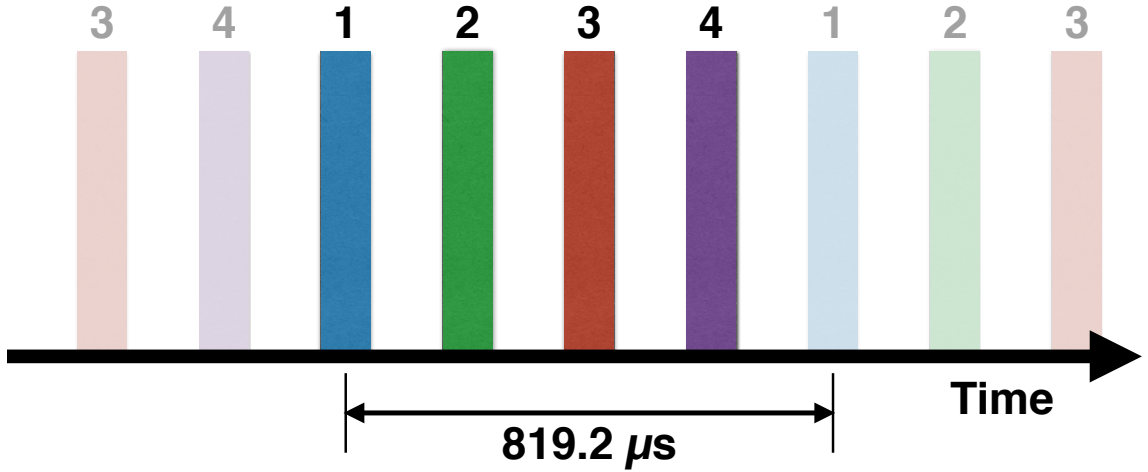
Figure 3.7: Short timing packets are sent from each of the WURC array antennas in rapid succession consisting of an 802.11 PLCP preamble and a short, 14-byte payload.

3.2 Model-Driven Evaluation

In order to understand how different environments and operational frequencies will effect the performance of a MU-MIMO system, we first turn to modern statistical MIMO channel models [18]. Since this statistical model requires tuning for different environments and frequencies, we compare the results using two published parametrizations for 300 MHz [19] and 5.8 GHz [20]. These results provide the theoretical motivation for over-the-air experiments to explore common application scenarios for UHF and 2.4/5.8 GHz WiFi.

3.2.1   UHF vs. 2.4/5 GHz: Channel Models

Spectrum differences between 2.4/5 GHz WiFi and sub-gigahertz frequencies are essentially due to the different manifestations of Doppler effects given each band's wavelength. Doppler effects are a result of transmitter, receiver, and client movement with respect to a transmission's wavelength. Because sub-gigahertz wavelengths are 2-4 times longer than 2.4/5 GHz, environmental variation will affect sub-gigahertz transmissions 2-4 times less (without considering multi-path effects).

Fig. 3.8 shows the theoretical, freespace 50% coherence time for various sub-gigahertz and 2.4/5 GHz frequencies [21]. The 50% coherence time is expected length of time that the channel characteristics will vary at most 50% given some velocity (effectively channel variation).

The coherence time difference between 2.4/5 GHz WiFi and sub-gigahertz frequencies is between 1-2 *orders of magnitude*. This channel characterization does not consider many real world effects such as multi-path or fading but provides a coarse characterization of the key differences in the two bands.

For a more realistic characterization of the spectrum differences, we employ the COST 2100 MIMO channel model, a flexible channel model that is well suited for MU-MIMO scenarios [18]. This channel model is tuned with parameters that are extracted from empirical measurements and thus does consider real-world channel effects such as fading, multi-path, and non-line-of-sight (NLOS) transmissions. Parametrized realizations of the COST 2100 model have been created for 300 MHz [19] and 5 GHz [20] bands. Using these models, we generate 15,000 channel snapshots at a simulated rate of 100 snapshots per second to characterize the variation of channel state over time and the separability of individual users. Specifically we explore the temporal correlation and receiver separability (shown in Fig. 3.9) of the generated matrices.

Temporal correlation is the average autocorrelation between channel snapshots at vary-
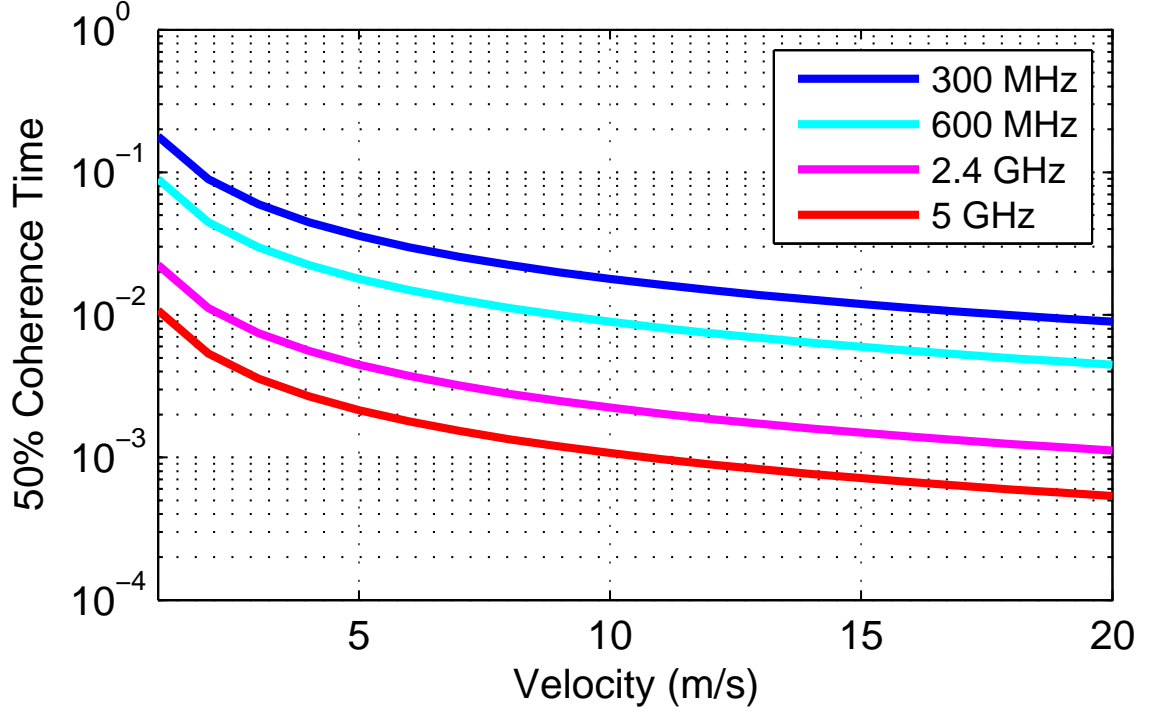
Figure 3.8: 50% coherence time for various sub-gigahertz and 2.4/5 GHz WiFi frequencies.

ing intervals of time calculated as described in [22]. The correlation coefficient ρ at time interval (Δt) is defined as:

$$\rho_{(\Delta t)} = \frac{\mathbb{E}\big[H_{mk}[t]H^*_{mk}[t + \Delta t]\big]}{\mathbb{E}\big[H_{mk}[t]H^*_{mk}[t]\big]} \tag{3.2}$$

where expectation is calculated for all combinations of transmit antenna m, receive antenna k and starting time sample t.

We show the magnitude of the temporal correlation coefficient in Fig. 3.9(a) for our generated channels. Lower temporal correlation results in less robust MU-MIMO transmissions because the measured channel state has a high probability of being stale. As seen in Fig. 3.9(a), the temporal correlation of 5 GHz WiFi almost immediately drops to below 0.9 ($T_{0.9}$) a point when when re-sounding the channel is strongly suggested [23].

According to the channel models, the approximate re-sounding time for 5 GHz is 50 ms
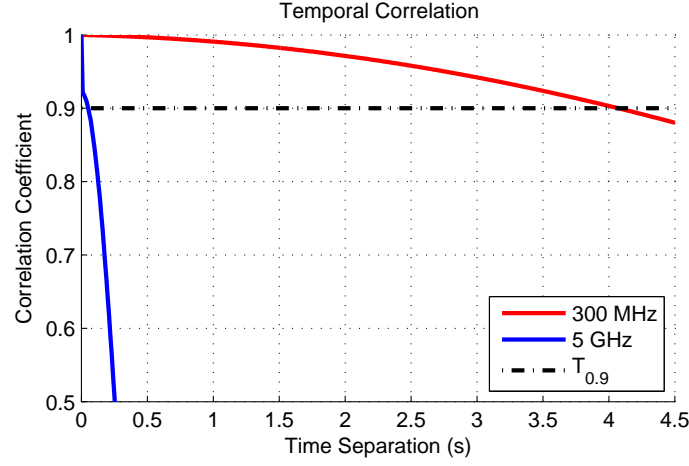
(a) Temporal correlation between channel snapshots from 0 to 10 seconds apart. Higher time correlation allows for more robust MU-MIMO performance. $T_{0.9}$ is 50 ms and 4 s for 5 GHz and 300 MHz respectively.
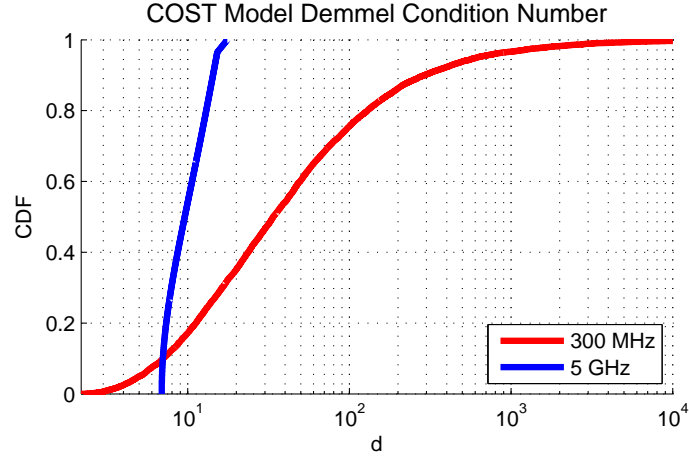
(b) CDF of model-generated Demmel condition number. Left is better for MU-MIMO

Figure 3.9: Temporal correlation and channel condition of 300 MHz and 5 GHz 2x2 MU-MIMO channels generated by COST 2100 MIMO channel model.

and 300 MHz is approximately 4.5 s (almost two orders of magnitude longer). This result is similar to what we expect from Doppler effects of the different frequency bands (Fig. 3.8) and is similar to our indoor temporal characterization in Sec. 3.3.1.

User separability refers to how well a multi-antenna transmitter can serve a set of users in parallel. The Demmel condition number is a modified matrix condition number that

directly predicts the efficacy of an adaptive MIMO or MU-MIMO transmission for a particular channel realization [7].

The Demmel condition number is computed using the eigenvalues λ_k of $HH^\dagger$ as:

$$d \triangleq \frac{\sum_{k=1}^{n} \lambda_k}{\lambda_n} \tag{3.3}$$

where $\lambda_1 > \lambda_2 > \ldots > \lambda_n$. This ratio represents how well a matrix can be inverted, a key component of many adaptive MU-MIMO techniques such as Zero-Forcing Beamforming [1] and MMSE [24]. Specifically, the higher the condition number, the more numerically unstable the inverse and thus the more inter-user interference during MU-MIMO transmissions reducing received SINR. The condition number ranges from 1 to infinity for well to ill-conditioned matrices, respectively.

This method of calculating the condition number is less forgiving than the traditional singular value ratio. The singular values (σ_k) of H are the square root of the eigenvalues of $HH^\dagger$. Thus, instead of σ_k/σ_n, the Demmel condition number is equivalent to $\sum \sigma^2/\sigma_n^2$ meaning that channel matrices with low singular values (resulting in inaccurate inversion) are even further "penalized." This modification to the condition number better predicts MU-MIMO performance, in fact, it is consistent and accurate enough to be used for determining parameters such as supported modulation rate and user selection [7].

The COST channel models show a significant difference between the 5 GHz WiFi and UHF bands. The CDF shown in Fig. 3.9(b) depicts how almost all of the generated 5 GHz channel matrices have a Demmel condition number less than 10 while UHF's channel condition varies far more and is significantly worse. This results in an increased ability for a MU-MIMO transmitter to invert the channel matrix and send orthogonal streams to each intended user.

Thus, existing MIMO channel models show that while the UHF channel is more temporally stable over time, its ill-conditioned channel matrices can result in lower served SINR

due to inter-user interference. However, the available parametrizations of the COST model are for indoor 5 GHz and outdoor UHF scenarios. We show in Sec. 3.3 how restricting these bands to these transmission environments does not tell the full story.

3.3   Experiment-Driven Evaluation

The models analyzed in Sec. 3.2 are parametrized for particular environments, frequency bands, and topologies. While they suggest that the performance of MU-MIMO beamforming in UHF bands may be advantageous, it it difficult to directly predict or simulate UHF performance using these models as they were not validated for application scenarios such as indoor or urban outdoor, nor the UHF frequency band.

In order to address uncertainty in these models for our target application (indoor and outdoor WLAN), we perform a set of experiments utilizing our custom SDR radio platform that allows us to measure the performance of a MU-MIMO transmission over a diverse set of carrier frequencies and characterize the wireless MU-MIMO channel for important temporal and spatial correlation properties.

We perform over-the-air beamforming transmissions in a densely packed, office scenario with multiple subscriber nodes and demonstrate not only the ability to simultaneously beamform to distinct users in relatively close proximity, but also the relative improvement that shifting to UHF frequencies provides.

Finally, we perform two sets of experiments with a customized MAC and PHY designed to gather dense, wideband, over-the-air channel estimates in realistic indoor and outdoor WLAN scenarios with multiple subscriber nodes. Using this data, we then demonstrate that the spatial correlation for outdoor users remains similar to that of 2.4 GHz WiFi, thus incurring no beamforming "penalty" for utilizing a frequency band with superior propagation and temporal correlation.

3.3.1   Indoor MU-MIMO Transmissions

Experimental Setup. First, we evaluate the performance of UHF MU-MIMO in an indoor, NLOS, office environment. Experiments were conducted during the work day with people walking through the halls in the environment depicted in Fig. 3.10.
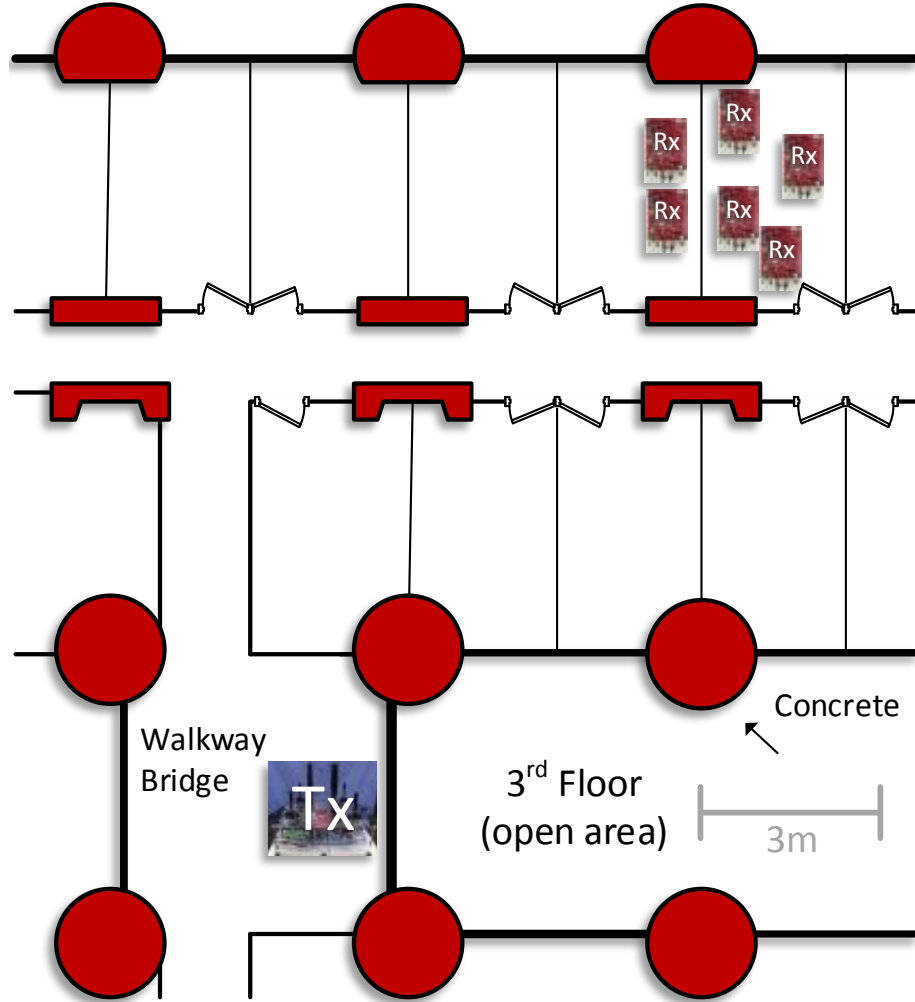
Figure 3.10: Indoor Experimental Test Setup.

The transmitting array was placed on a third floor walkway bridge and 6 separate receivers in two adjacent offices within the adjoining hallway. Note that the to-scale depiction in Fig. 3.10 shows the relative co-location of all receiving nodes with respect to the distance from the transmitter to simulate a densely packed office environment. This rep-

resents a realistic, challenging case for indoor stationary MU-MIMO transmissions due to the co-located receivers.

To encompass a wide range of user grouping conditions, every possible combination of transmit and receive antennas are considered. Sixty transmissions are performed for each topology. The center frequencies for each frequency band (*i.e.* channel) were chosen so that transmissions did not encounter interference from other equipment. Specifically, the UHF channel was first directly scanned for existing DTV or microphone transmissions and an experimental license was obtained to operate equipment on that channel. The channels selected for 2.4 and 5.8 GHz are not currently supported by the regulatory domain where these experiments were performed, thus ensuring minimal ISM-band interference. Using the measurement technique specified in Sec. 3.1.3.1, every possible topology's MU-MIMO capacity is measured for each frequency band and shown in Fig. 3.11.
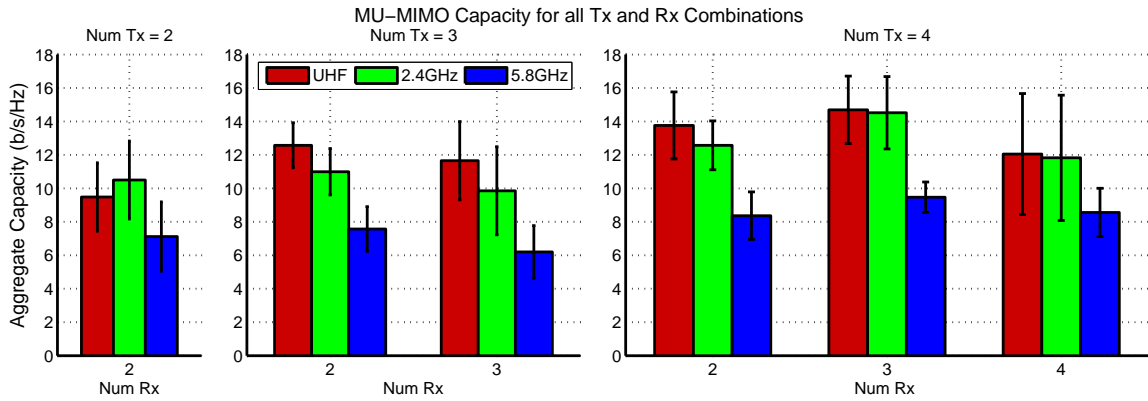
Figure 3.11: Received MU-MIMO Capacity.

MU-MIMO Achievable Sum-Rate Capacity. Based on the channel models and accompanying analysis presented in Sec. 3.2.1, we expect that the increased spatial correlation of UHF channels will not allow for MU-MIMO transmissions to accurately separate nearby users. However, we find that UHF MU-MIMO transmissions can actually achieve a sum capacity similar to that of 2.4 GHz WiFi transmissions (always between 1-2 b/s/Hz above of below the 2.4 GHz band).

In fact, we find that majority of the intuition and channel models surrounding UHF MU-MIMO are not specific to the frequency band itself but rather generalized characteristics of MU-MIMO transmissions. For example, the available MU-MIMO channel models characterize *indoor* WiFi and *outdoor* UHF channel environments where, regardless of frequency band, we expect increased difficulty in user separability in outdoor environments. Note the channel condition of the different transmission bands in the NLOS environment in Fig. 3.12 are similar in contrast to Fig. 3.9(b). Even though the wavelength of UHF is longer resulting in better propagation through materials, the UHF-band transmission still experiences enough multi-path to successfully beamform to multiple users in parallel.
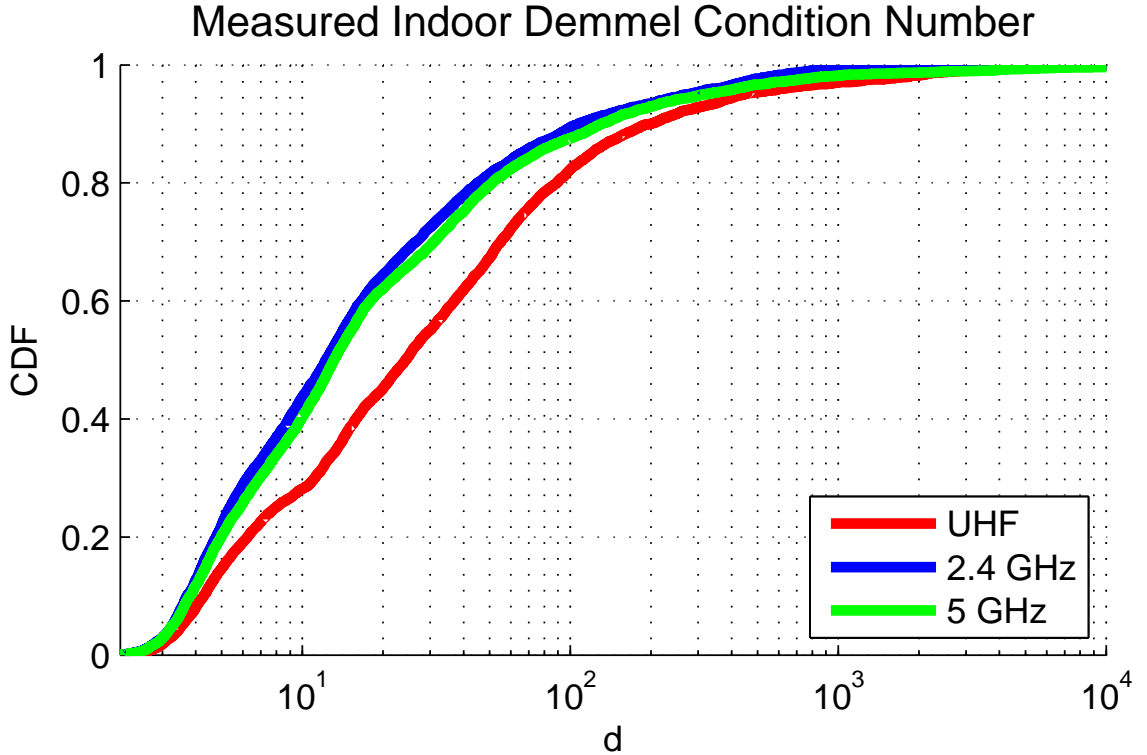
Figure 3.12: Demmel condition number measured for the indoor environment. Left is better for MU-MIMO.

Additionally, the results shown in Fig. 3.11 show a known trend of achievable capacity for MU-MIMO transmissions where the MU-MIMO gain plateaus as the available Degrees

of Freedom (DoF)[1] are reached. The consistently worse performance of 5.8 GHz is explained by the high attenuation experienced by that frequency band in NLOS conditions combined with its sensitivity to environmental variation.

Note that UHF MU-MIMO consistently outperforms 2.4 GHz transmissions except for in the 2x2 transmission scenario. Because the sum transmit power emanating from the array is held constant regardless of the number of transmit antennas in use, the performance differential is solely a result of channel state, specifically it is an indicator of temporal channel correlation due to the WARPLab measurement platform.

As discussed on Sec. 3.1.3.1, the latency in the WARPLab platform is due to the rate at which the host PC can download and upload samples to each of the WARP boards over Ethernet. In our system, we benchmark a read/write rate of approximately 2.5 ms per buffer and the closed loop beamforming method employed requires between 10 to 20 ms to complete depending on the number of transmit and receive antennas (the difference between a 2x2 and 4x4 transmission scenario).

Measured Temporal Correlation. To gain additional insight into the measured capacity results and to infer real world performance from our MU-MIMO transmissions, we also consider the channel correlation measured during each experiment.

For each topology, we consider each of the 60 MU-MIMO transmissions and their channel matrices. We calculate channel correlation between varying times during the experiment to measure the rate of change of the channel information with respect to time. These calculations are an average over all topologies (all combinations of transmit and receive antennas).

Fig. 3.13 shows how the channels decorrelate over the course of one measured second in time. This is effectively an indicator of how long a transmitter has after measuring the channel matrix and before actually transmitting parallel streams using that measurement.

---

[1]DoF here refers to how many more transmit antennas there are than receive antennas in a MIMO transmission.

A coherence time of $T_{0.9}$ represents when the probability of the channel being too stale to successfully beamform over is high.
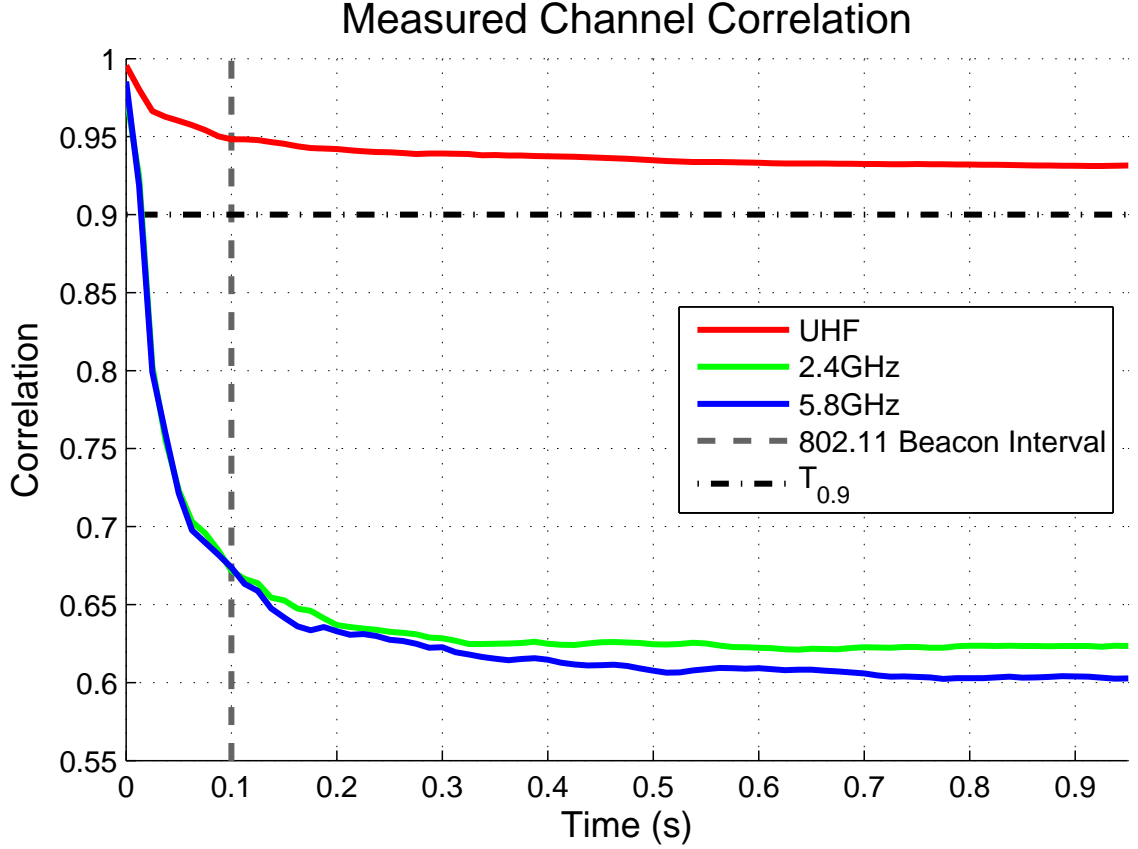
Figure 3.13: Measured Temporal Channel Correlation, depicting Beacon Interval, and $T_{0.9}$. WARPLab latency is 10-20 ms depending number of transmit and receive antennas.

First, note that the WARPLab latency range of 10 to 20 ms is approximately $T_{0.9}$ for the the two WiFi frequency bands. This indicates why only the 2x2 transmission scenario has the 2.4 GHz transmitter outperform UHF MU-MIMO; the latency between the sounding and transmission phase was the lowest and just at the $T_{0.9}$ limit.

While the 2.4/5 GHz frequencies both drop significantly within 100 ms, UHF remains above the $T_{0.9}$ threshold for the maximum one measured second difference between channel matrices. While these correlation values are not asymptotic and will eventually degrade, the performance of 2.4 and 5.8 GHz is suffcently low for stationary devices [23].

Also note that the 802.11 beacon packet rate (100 ms) is greater than the interval that 2.4/5 GHz MU-MIMO channels decorrelate. However, the stability of the UHF channel implies that a UHF MU-MIMO system could use periodic protocol packets for exchanging channel state information.

Finally, the channel correlation result shown in Fig. 3.13 effectively scales the MU-MIMO achievable rate shown in Fig. 3.10. The rate at which the 2.4/5 GHz channel decorrelates necessitates channel sounding on a per packet basis adding considerable overhead to MU-MIMO transmissions. However, the temporal stability of the UHF MU-MIMO channel allows a transmitter to significantly reduce this overhead intensive sounding process and thus significantly increase the potential MU-MIMO gains.

3.3.2   Outdoor Channel Characterization
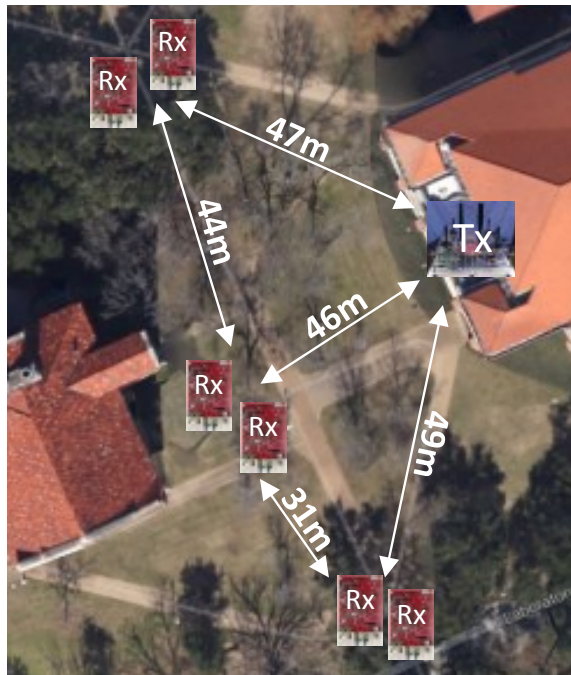
Figure 3.14: Experimental setup for outdoor channel sounding experiments. Distances between transmitter (on third floor balcony) and receivers shown. Note building and tree locations.

Finally, using the experimental framework developed in Sec. 3.1, we perform outdoor channel sounding experiments to directly compare the performance and stability of UHF MU-MIMO channels. To that end, we setup an experimental network of a collection of nodes located outdoors being served by our array from a third floor balcony. Although the UHF transmitter is capable of transmitting much further distances, we limited the scale of the topology as shown in Fig. 3.14 to ensure a fair comparison between UHF and 2.4/5 GHz bands. The locations of the nodes were chosen such that the transmissions from the UHF and 2.4/5 GHz bands would reach the receivers (the UHF band transmitters can easily transmit further than 50 m). However, even by reducing the receiver distance to what is shown in Fig. 3.14, the 5 GHz band transmissions did not reliably reach the receiving nodes severely limiting the number of measured channel matrices. Thus, we restrict our outdoor comparison to the UHF and 2.4 GHz bands.
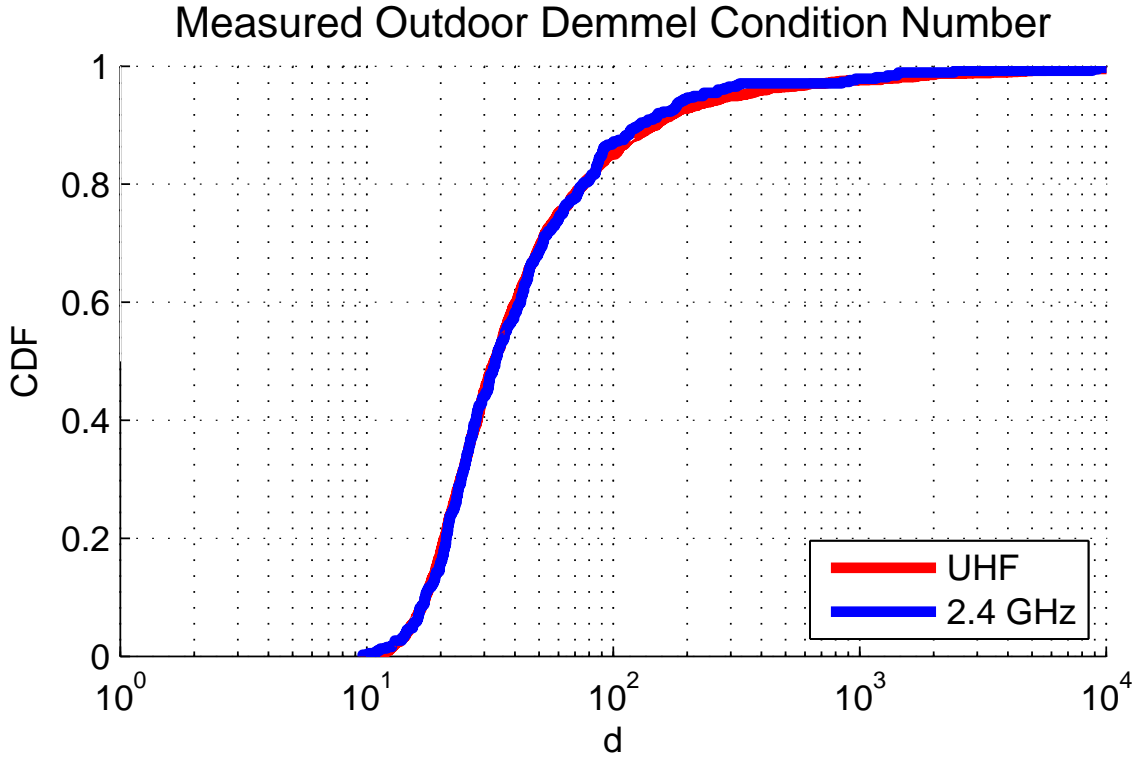
Figure 3.15: Measured Demmel Condition Number of the outdoor MU-MIMO channel.

Just as we evaluated temporal correlation in the multi-path rich, indoor transmission environment, we seek to similarly characterize the most detrimental aspect of the outdoor MU-MIMO channel: receiver separability. Ill-conditioned channel matrices, as discussed in Sec. 3.2.1, have a detrimental effect on an MU-MIMO enabled transmitter's ability to separate multiple users.

In the previous section, we found that while temporal stability of UHF was greater than that of 2.4/5 GHz, spatial correlation did not suffer as the UHF MU-MIMO transmissions were able to separate the co-located receivers. However, in an open, outdoor line-of-sight (LOS) environment, we find that both the UHF and 2.4 GHz bands exhibit the same Demmel condition number. Additionally, the CDF of the Demmel condition number closely matches the COST UHF channel condition shown in Fig. 3.9(b). This suggests that the comparison shown in Fig. 3.9(b) is not a result of the frequency band itself, but rather the wholly different channel environments in which the model was parametrized.

MU-MIMO PROTOCOL FOR VARIABLE CHANNELS

To increase system throughput of MU-MIMO transmissions in variable environments (e.g., the 2.4/5.8 GHz band), we present **P**re-sounding **U**ser and **M**ode selection **A**lgorithm (PUMA), a method for efficiently making MU-MIMO MAC-layer decisions before the users' CSI is collected.

PUMA allows for an MU-MIMO Access Point (AP) to make the key MAC-layer decisions of (number of transmit antennas and collective number of receive antennas) and the group (set of receivers) prior to transmission and thus without requiring measured CSI.

To accomplish this, PUMA exploits theoretical properties of MU-MIMO system scaling with respect to mode, characterizes the relative cost and overhead of realizing each potential mode, and estimates the per-stream transmission rate and aggregate throughput in each mode for a potential set of users. We show that on average, PUMA selects the mode and group that achieves a joint transmission rate that is within several percent of what would have been achieved by sounding all users. Yet, PUMA makes its selection prior to sounding any users and therefore does not require any additional sounding overhead for user and mode selection. Moreover, we show that PUMA obtains 30% higher aggregate throughput in comparison to even the best fixed-mode policy (e.g., a policy such as always using the maximum number of transmit antennas and collective receive antennas).

The remainder of this chapter is organized as follows: Sec. 4.1 provides an overview of PUMA, Sec. 4.2 describes how to integrate PUMA with an existing MU-MIMO transmission protocol (specifically 802.11ac), and finally, Sec. 4.3 provides an experimental evaluation of PUMA.

4.1   PUMA Overview

The PUMA algorithm is executed before the start of any MU-MIMO transmission with only *a priori* information; i.e., without any information garnered from previous multi-stream communication or channel sounding. The necessity of using only immediately available information for mode selection stems from the highly volatile nature of an indoor WLAN environment.

Although this multi-path rich, fading environment results in well conditioned channels, indoor WLAN environments can have an unpredictable coherence time [25, 10, 26]. The measured CSI can easily become stale between packets, making previous transmissions unhelpful for predicting future environments. Additionally, because of the high overhead incurred from measuring CSI, it would be costly for the transmission mode selection to rely on channel sounding. To alleviate the effects of variable and costly CSI measurement, PUMA selects the best mode and user group without CSI.

The following sections detail how, using only the information available before channel sounding (Sec. 4.1.1), PUMA predicts the per-user MU-MIMO datarate (Sec. 4.1.2), computes a potential group's throughput (Sec. 4.1.3), and finally selects the appropriate user/group combination (Sec. 4.1.4).

4.1.1   Available Pre-sounding Information

Before initiating an MU-MIMO transmission, the AP has the following information: *system state*, *queue state*, and *link state*. By leveraging this *a priori* information, PUMA enables an AP to select the best mode. The system state and queue state are used directly in protocol overhead calculation detailed in Sec. 4.1.3. Link state is leveraged in per-user data rate prediction detailed in Sec. 4.1.2. Combining these components, PUMA estimates the throughput of any possible MU-MIMO transmission an AP can execute.

System State. Before any transmission, the AP knows the hardware configurations of itself and its clients. This includes the available (maximum) number of transmit antennas M_{max} and the available number of associated users' receive antennas K_{max}. PUMA leverages this system state for overhead computation. While a greater number of overall antennas results in increased data transmissions, it also significantly increases sounding overhead.

Queue State. The AP is also aware of each receiver's backlog or queue size. The amount of available data directly affects how much sounding overhead is amortized. If the amount of available data for a particular user is relatively small compared to sounding overhead, the potential gains of a MU-MIMO transmission to that user are severely diminished. We express the available data in terms of available packets b.

Link State. The AP is aware of each user's link state or omnidirectional Signal-to-Noise Ratio (SNR). The AP automatically gathers this information from periodic beacon messages and updates this information after each received packet. Unlike CSI, this metric need not be instantaneous since received signal strength stays coherent longer than multiple packet transmissions (approximately 90 ms at 0.9 kph [27]). PUMA leverages each client's link state to estimate the achievable data rate.

4.1.2 Predicting User-specific MU-MIMO Datarate

The first key technique of PUMA is the estimation of per-user datarate. PUMA accomplishes this by computing the expected Signal-to-Interference-and-Noise Ratio (SINR) of an MU-MIMO transmission using only pre-sounding information based on theoretical MU-MIMO system scaling. PUMA then estimates the achievable rate using a protocol specific minimum SINR table (such as Table 4.2).

4.1.2.1 Post-Sounding Rate Estimation

Many works provide expressions for the expected received SINR or aggregate capacity of a MU-MIMO transmission (e.g. [1, 28]) such as the following (where C is in b/s/hz):

$$C = \max_{\vec{w}_k, P_k} \sum_{k=1}^{K} \log_2 \left(\frac{1 + \sum_{j=1}^{K} P_j |\vec{h}_k \vec{w}_j|^2}{1 + \sum_{j=1, j \neq k}^{K} P_j |\vec{h}_k \vec{w}_j|^2} \right). \tag{4.1}$$

However, such methods are not suitable for our purposes because they require information only available after sounding: the measured channel matrix H. Instead, we seek to estimate the expected performance of a MU-MIMO transmission *before* the channel is sounded.

Given the significant overhead of channel sounding (which we discuss specifically for 802.11ac in Sec. 4.2), a transmitter must serve whatever user it sounds to maximize performance (to be described in Eq. (4.4)). Additionally, the channel state is highly variable for the frequencies used in WLANs [10] and thus channel sounding must occur before every packet transmission (i.e., previously measured channel matrices cannot be used reliably for future transmissions).

4.1.2.2 Pre-Sounding Rate Estimation

PUMA's pre-sounding rate estimation method is based on theoretical MU-MIMO system scaling. PUMA exploits this scaling by estimating the received SINR for a particular client

and converting it into an expected achievable datarate using a standard specific minimum SNR table such as Table 4.2.

The basis of Eq. (4.1) is the computation of SIR from the multiplication of the $\vec{h}$ and $\vec{w}$ vectors. While the H matrix represents the measured CSI, the W matrix is what the AP computes from H to actually construct parallel streams.

A commonly employed MU-MIMO precoding technique, ZFBF [1], requires that the W matrix be computed as the inverse of the H matrix. Beamforming itself is the application of the W steering matrix through the channel H or $H \cdot W$. While a matrix times its inverse should result in the identity matrix, the actual value of H and W may not precisely meet this criterion due to per-user power allocation or an ill-conditioned H.

Eq. (4.1) is based on this matrix multiplication $H \cdot W$. The additional computations are simply to convert SINR into Shannon Capacity.[1] Thus, the result of the matrix multiplication is the diagonal matrix L and is a representation of the received SINR. Each diagonal element l_i corresponds to each of the K receiving antennas and its magnitude encompasses the beamforming gain (or loss) with respect to the received omnidirectional signal strength P/N_o [29].

Thus, assuming equal per-user power allocation (and normalized per-antenna power allocation), the actual SINR for a beamforming transmission based on this measured H is:

$$\text{SINR} = 10 * \log_{10}\left(\frac{P/N_o}{M} |l_i|^2 \right). \tag{4.2}$$

This expression still leaves us in the same position as with Eq. (4.1). However, instead of attempting to calculate $|l_i|^2$, we consider its distribution.

The distribution of the SINR determining $|l_i|^2$ factor can be shown to be Erlang for Rayleigh matrices [30]. The distribution is dependent upon the dimensions of H (M and K) and has mean $(M - K + 1)/K$. Therefore, combining with Eq. (4.2), we estimate

---

[1]Recall that the Shannon Capacity is computed as $C = \log_2(1 + \text{SINR})$.

per-user SINR as

$$\mathcal{E}\{\text{SINR}_{\text{BF}}\} = 10 \cdot \log_{10}\left(\frac{M - K + 1}{K}\frac{(P/N_o)}{M}\right). \tag{4.3}$$

While the resulting expected value computation of per-user SINR shown in Eq. (4.3) is inherently less precise than Eq. (4.1) because it does not use CSI, Eq. (4.3) exploits general system scaling properties of MU-MIMO transmissions to produce a sufficiently accurate result (verified in Sec. 4.3.2).

4.1.2.3 Model Rationale

This scaling is proportional to the available DoFs of a particular transmission mode. A mode's DoFs refer to how many more transmit antennas there are than receive antennas or $M - K + 1$. The larger this value, the easier it is to construct interference free parallel streams. An $[M, K]$ transmission requires an $M \times K$ channel matrix, which is easier to accurately invert or otherwise decompose when $M > K$ since it will be better conditioned [31].

However, absolute DoFs do not reveal the full solution for theoretical MU-MIMO received SINR scaling; instead we consider normalized DoFs. For example, although both $[M_{10}, K_9]$ and $[M_3, K_2]$ transmissions have equivalent absolute DoFs (2), the per-user SINR increase would be far more noticeable in the latter system because it has relatively more DoFs with respect to K. Thus, MU-MIMO SINR should scale relatively with $(M - K + 1)/K$.

4.1.2.4 Inferring Rate from SINR

PUMA's SINR estimation method, Eq. (4.3), only requires the M and K of a potential mode and a particular user's omnidirectional signal strength (P/N_o) periodically updated from beacon packets and previous transmissions. Like CSI, the omnidirectional SNR can

become stale after a period of time. However, omnidirectional SNR is far more robust to environmental variation than CSI. Channel matrices used for MU-MIMO transmissions are dependent on precise magnitude and phase offsets between each antenna path. Given the wavelengths of the frequencies used for WLANs and their physical interactions with obstacles, slight variations in the transmission environment can render a previously measured magnitude or phase useless. Instead, SNR is a coarse grained measurement that is an aggregate of all amplitudes and thus varies more slowly.

PUMA estimates the received SINR for each user in an $[M, K]$ system, which allows an AP to not only compare an $[M, K]$ system to an $[M', K']$ system, but also estimate an approximate Modulation and Coding Scheme (MCS) rate for each user using the SNR-MCS tables provided by the standard (for 802.11ac see Table 4.2).

4.1.3   Computing Expected Throughput

The second component of PUMA is the analysis of the selected MU-MIMO protocol specifically with respect to its aggregate throughput (R) using renewal arguments.

Other than the inferred expected rate (detailed in Sec. 4.1.2), the main components of expected throughput calculation are node backlog (the number of available packets to transmit per receiver) and net transmission overhead. By combining node backlog (client dependent), net overhead (mode dependent), and the expected rate (client and mode dependent), PUMA computes the expected goodput of any possible MU-MIMO transmission. Thus, PUMA provides a quantifiable metric for any possible transmission, allowing an AP to accurately compare potential mode and group selections. For simplicity, please refer to Table 4.1 for parameter notation:

The throughput R for any wireless transmission is generally represented as the amount

Parameter	Description	
R	Aggregate Goodput	(bits/s)
r_i	Per-user achievable rate	(bits/s)
b_i	Per-user aggregation rate	(packets)
L_D	Net Data Length	(bits)
L_p	Packet Length	(bits)
T_{OH}	Net Overhead Time	(s)
T_S	Channel Sounding Time	(s)
T_{CF}	CSI Feedback Time	(s)
T_{ACK}	Net Receiver Acknowledgment Time	(s)

Table 4.1: Expected Rate Notation

of data to transmit divided by the total transmit time (including overhead):

$$R = L_D/(T_D + T_{OH}). \tag{4.4}$$

The total amount of transmitted data across all streams L_D given the maximum packet length L_p is:

$$L_D = \sum_{i \in K} b_i \cdot L_p. \tag{4.5}$$

The overhead time T_{OH} is:

$$T_{OH} = T_S + T_{CF} + T_{ACK} \tag{4.6}$$

where T_S is the channel sounding time, T_{CF} is the channel feedback time, and T_{ACK} is the receiver acknowledgment time. Thus, T_D, the total data transmission time given the per-user rate r_i is:

$$T_D = \max_{i \in K} (b_i \cdot L_p)/r_i. \tag{4.7}$$

We express T_D as a maximum value in case the protocol (e.g., 802.11ac) supports

different per-user packet aggregation rates b_i or different per-user modulation rates r_i.

Through this formulation of aggregate goodput, we see that the value of T_{OH} limits the performance of a MU-MIMO transmission. A larger user set K served with more antennas M results in a larger amount of transmitted data L_D but it also results in a larger amount of overhead T_{OH} (which, like L_D, also scales with M and K). The appropriate mode is one that maximizes R by efficiently balancing L_D and T_{OH}.

While the basis of PUMA's overhead analysis is applicable to any standard, we focus on 802.11ac in Sec. 4.2.

4.1.4 Selecting Mode and User Group

PUMA seeks to jointly minimize the effects of T_{OH} and maximize the value of R from Eq. (4.4). Essentially, for a set of per-user rates r_i and per-user backlog b_i, PUMA computes:

$$\max_{M \in M_{max}, K \in K_R} R(M, K, b, r) \tag{4.8}$$

where M_{max} is the maximum possible number of transmit antennas and K_R is the subset of all potential associated receivers with packets in their queues.

One method of maximizing this expression is an exhaustive search. The value for b is set per-user and the value of r is dependent upon how many other concurrent users exist.

Thus, the overall search space is:

$$\sum_{M}^{M_{max}} \sum_{K=1}^{M} \binom{K_R}{K}. \tag{4.9}$$

This number of potential combinations can be exhaustively searched as long as K_R is not too large. Limiting K_R can be done in any number of ways such as truncating users with small b_i, fairness, or other QoS constraints.

Thus, the exhaustive search method is feasible. When an AP is ready to transmit, the algorithm is executed as follows:

1. Generate $K_R \in K_{max}$ (associated receivers with packets in their queues).

2. $\forall M \in M_{max}$ and sets of $K \in K_R$ (where $K \leq M$), compute the expected per-user SINR for each $[M, K]$ combination as shown in Eq. (4.3).

3. Using the standard's receiver sensitivity table for 90% packet reception, estimate the MCS for each K in every potential group to generate a list of all possible M, K, MCS combinations.

4. Using Eq. (4.4) combined with protocol specific values, calculate the expected aggregate throughput for each M, K, user group dependent MCS combination and choose the largest.

While multiple potential modes ($[M, K]$ combinations) could have equivalent expected throughputs if they have identical frame aggregation values or similar link qualities, the probability of this occurring is relatively low and the final selection can be chosen randomly.

4.2   PUMA with 802.11ac

While PUMA's basic mechanism is applicable to any random access MU-MIMO protocol, we demonstrate its functionality with 802.11ac. The two protocol-dependent components of PUMA are datarate inference from expected SINR and potential mode aggregate throughput calculation.

4.2.1   Datarate Inference from SINR

After computing the expected SINR from Eq. (4.3), PUMA employs protocol-specific minimum SNR tables to infer the expected per-user datarate. With this per-user datarate, PUMA computes the expected aggregate throughput of any possible mode and user group.

Table 4.2 is 802.11ac's minimum SNR table for ensuring a 90% packet reception rate. For each user and potential mode, PUMA selects the MCS index whose corresponding SNR is less than or equal to the expected value calculated using Eq. (4.3). The corresponding number of data bits per symbol (N_{DBPS}) is the per-user datarate.

MCS Index	Rate	N_{DBPS}*	SNR (dB)
0	BPSK $^1/_2$	117	1.1
1	QPSK $^1/_2$	234	4.1
2	QPSK $^3/_4$	351	6.7
3	16-QAM $^1/_2$	468	9.6
4	16-QAM $^3/_4$	702	12.8
5	64-QAM $^2/_3$	936	17.2
6	64-QAM $^3/_4$	1053	18.4
7	64-QAM $^5/_6$	1170	19.7
8	256-QAM $^3/_4$	1404	23.9
9	256-QAM $^5/_6$	1560	25.5

* N_{DBPS} (number of data bits per symbol) for each MCS when
channel bandwidth is 80 MHz.

Table 4.2: Required SNR (for 90% packet reception rate)

Note that the difference between MCS index's SNRs is as large as 4.4 db (MCS 4 vs. 5). Although PUMA's SINR estimation method is inherently less accurate than a post-sounding method since it does not utilize CSI, each MCS's large SNR ranges immensely reduce the resulting effect of this error (see Sec. 4.3.2).

(a) Three concurrent streams at MCS 2. $R = 145$ Mbps.

(b) Two concurrent streams at MCS 4. $R = 161$ Mbps.

LEGEND				
■	EBO	Expected Backoff	■ DIFS	DCF Interframe Spacing
■	NDPA	Null Data Packet Announcement	■ SIFS	Short Interframe Spacing
■	NDP	Null Data Packet	■ CBFR	Compressed Beamforming Report
■		Poll Time	■	Data
■	BA	Block Ack		
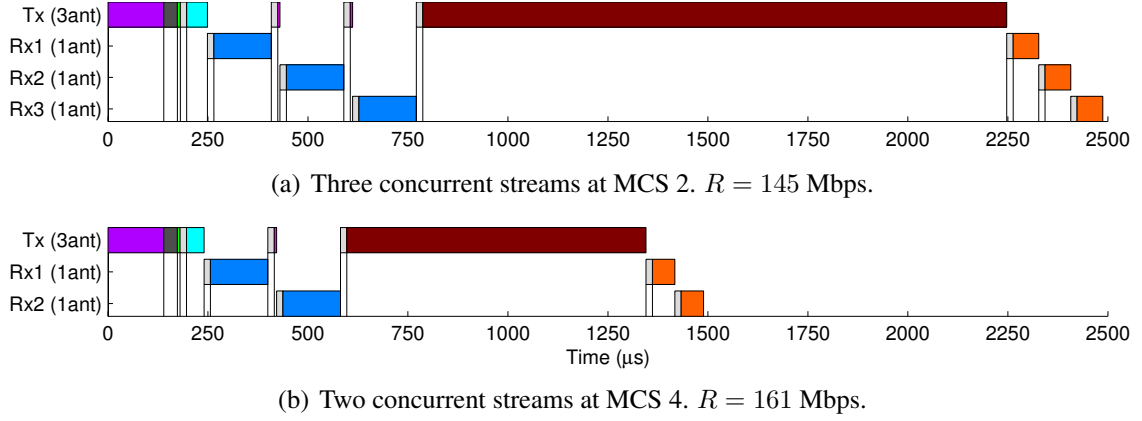
Figure 4.1: Example 802.11ac transmission timeline with 3 antenna transmitter sending multiple, 10 aggregated packet streams at 80 MHz (to scale).

4.2.2 Aggregate Throughput Calculation

Aggregate throughput calculation is dependent on protocol overhead in addition to per-user expected datarate and backlog. Analysis of the 802.11ac specification allows for PUMA to precisely compute the expected aggregate throughput for a potential mode and user group. We express each segment of an 802.11ac transmission as generally described in Eq. (4.4)–(4.7). An example of an 802.11ac MU-MIMO transmission is depicted as a timeline shown in Fig. 4.1. After the expected backoff duration EBO=139.5μs (or 15.5 slots at 9μs/slot) and DIFS=34μs, the AP begins transmitting.

Channel Sounding ($\mathrm{T_S}$). The transmitter first announces to all users in transmission range which specific subset is to expect the upcoming transmission with the Null Data Packet Announcement (NDPA=7.4μs) followed by the Null Data Packet (NDP). The NDP contains the sounding pilots used by the receivers to estimate the channel state between itself and each transmitting antenna (thus it scales with M).

Channel Feedback ($\mathbf{T_{CF}}$). Each receiver must sequentially reply with the Compressed Beamforming Report (CBFR) returning a compressed, per subcarrier version of the sounding pilots to the transmitter in the form of angle pairs (12 or 16 bits each for MU-MIMO, 10 or 12 bits for SU-MIMO). The number of angle pairs scales with the number of transmitting antennas M and the number of receiving antennas for that particular node (the 3×1 vector in the example shown in Fig. 4.1 requires two angle pairs per subcarrier).

The 80 MHz bandwidth has 234 usable subcarriers. One compressed angle set can be used to indicate groups of 1, 2, or 4 subcarriers. This results in a very high overhead due to channel feedback. The example timeline is scaled to 16 bits of feedback with a subcarrier grouping factor of 2. Each receiver has 3,744 bits to transmit back to the AP at the base rate (MCS 0, see Table 4.2) to ensure that the report is not lost.

While the $M=3$ antenna transmitter sent the NDP in 7.4μs, each receiver must spend 144μs responding with its CBFR.

Between each CBFR, the AP sends a short polling packet requesting the next user to send its report.

Data Transmission ($\mathbf{T_D}, \mathbf{L_D}$). Finally, the AP forms concurrent data streams at varying frame aggregation and MCS rates. If one stream finishes early, the remaining time until the longest stream completes is wasted. PUMA's throughput formulation shown in Eq. (4.4) accounts for the potential of unequal stream lengths. Instead of arbitrarily trying to avoid this scenario, PUMA searches for the option with the largest throughput. Because this wasted airtime does have a negative effect on throughput, the probability of such a transmission occurring is low. Nevertheless, PUMA can handle this scenario without explicitly considering it.

Block Acknowledgment ($\mathbf{T_{ACK}}$). Once the data transmission has completed, each receiver must sequentially reply with block acknowledgments (BA=6.4μs). Thus, this component of the overhead scales with K.

4.2.3 Example PUMA Transmission

Fig. 4.1 depicts the timeline of two separate transmissions from a 3-antenna transmitter to either three (Fig. 4.1(a)) or two single antenna receivers (Fig. 4.1(b)). Each stream consists of 10 aggregated full size (1,500 byte) packets. As previously discussed, the overhead (channel sounding, channel feedback, and acknowledgment) for both transmissions is sent at the base rate although the resulting data rates are different.

PUMA first calculates the expected datarate of each user. For this example, let each of the three potential users have 18 dB omnidirectional SNRs. Using Eq. (4.3), PUMA computes expected SINRs of 8.4 and 13.3 dB for users in the $[M_3, K_3]$ and $[M_3, K_2]$ modes respectively. By referencing Table 4.2, PUMA selects MCS 2 for $[M_3, K_3]$ and MCS 4 for $[M_3, K_2]$.

Once the per-user datarates are computed, PUMA computes the aggregate throughput of the potential modes given the expected datarate and node backlog as shown by the to-scale timelines in Fig. 4.1. Through this computation, PUMA identifies that the aggregate throughput of the $[M_3, K_3]$ transmission is 145 Mbps while the aggregate throughput for the $[M_3, K_2]$ transmission is 161 Mbps. Thus, PUMA selects the $[M_3, K_2]$ transmission mode.

This example highlights a common yet counterintuitive result that PUMA identifies. *A MU-MIMO transmission does not always benefit from using the most antennas.* Not only does the protocol overhead increase with additional antennas, but also the per-user MU-MIMO SINR decreases resulting in lower per-user datarates.

4.2.4 Numerical Analysis of Mode Selection

To observe the expected performance and gain an intuition into the effects of M and K, we present a numerical example of PUMA specific to 802.11ac. We consider four separate $[M_4, K_{1:4}]$ systems. To evaluate the best case performance for each $[M, K]$ system,
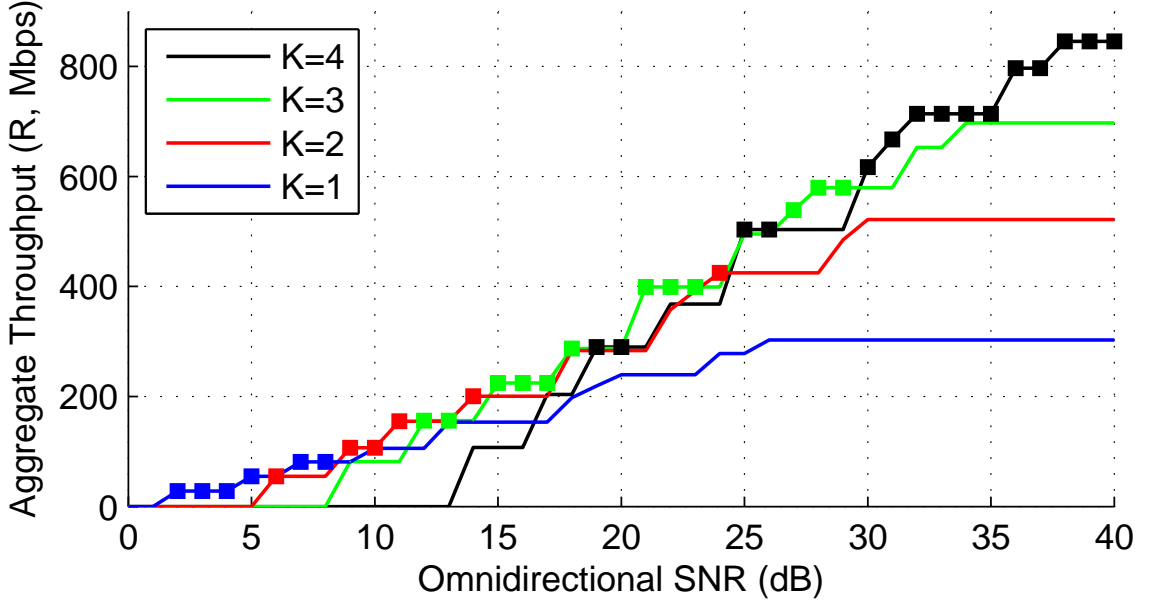
Figure 4.2: Theoretical expected throughput (Eq. (4.4)) of a M_4 antenna transmitter to $K_{1:4}$ parallel single antenna receivers with b=64. For each omnidirectional SNR, maximum expected aggregate throughput represented by square marker.

all receivers have equivalent omnidirectional SNRs and transmit at the maximum frame aggregation rate (b=64) for maximum overhead amortization.

For a range of omnidirectional SNRs, we infer the expected per-user datarate given M and K using Eq. (4.3) and Table 4.2. We then calculate the expected aggregate throughput using Eq. (4.4) considering 802.11ac-specific overhead and show the performance for each of the four [M_4, $K_{1:4}$] systems in Fig. 4.2. The mode that results in the highest aggregate throughput is marked for each omnidirectional SNR.

The key intuition gained from this numerical example is the effect of channel sounding overhead with respect to M and K. While increasing M slightly increases the size of an individual CBFR, additional K increases the number of CBFRs. Thus, to efficiently amortize the overhead induced from increased M and K, the channel state must support high per-user datarates.

The theoretical properties of MU-MIMO system scaling highlighted in Eq. (4.3) show

the effect of higher order modes on per-user SINR. For example, an $[M_4, K_4]$ system results in a per-user SINR approximately 12 dB less than the omnidirectional SNR. Thus, a per-user omnidirectional SNR much higher than 12 dB is required to perform a $[M_4, K_4]$ at a high datarate. The aggregate throughput for a $[M_4, K_4]$ system in Fig. 4.2 is non-zero at 14 dB, begins to contend with the other modes at 19 dB, and consistently outperforms all other modes starting at 30 dB.

The results for higher order modes in this numerical example further highlight the counterintuitive result shown in Fig. 4.1: *increasing the number of parallel streams is not always the most efficient transmission mode*. Additionally, because MCS defined datarates are discrete, the aggregate throughput curves for each mode are jagged and result in many intersection points. Thus, each mode does not have a clear SNR range where it exhibits the maximum throughput, necessitating PUMA's dynamic selection method.

PUMA's complete analysis of MU-MIMO system scaling and protocol specific overhead allows for the appropriate mode selection decision given the current state of the system. In fact, Fig. 4.2 graphically represents a PUMA enabled AP's decision engine and the markers represent the decisions themselves: the dynamically calculated, maximum throughput mode.

4.3   PUMA Evaluation

4.3.1   Experimental Methodology

We first characterize the performance of PUMA through OTA transmissions to verify PUMA's datarate inference method and generate realistic channel traces. We then utilize this realistic OTA data for channel-trace driven emulation.

4.3.1.1 OTA Experimentation

We conduct OTA experiments using the WARP software defined radio [16] utilizing a Zero-Forcing Beamforming framework developed in [10] and expanded in [32]. To perform our experiments, we modify WARPLab, a system that allows for baseband signals to be proccessed in MATLAB, downloaded to the board, and transmitted over-the-air. A full description of this transmission process is found in Sec. 3.1. Because the 802.11ac standard allows up to 8 spatial streams, we connect two 4 antenna WARP boards together to make an appropriate transmitter.

We place 8 receiving antennas in 8 different non-line of sight locations to emulate a typical indoor wireless LAN environment as shown in Fig. 4.3. We then serve every combination of $[M_{1:8}, K_{1:8}]$ to get a variety of different channel environments for each topology and measure the SINR.

The resulting variability of omnidirectional SNR measurements resulted in an overall mean of 18.3 dB and standard deviation of 5 dB. This allows us to verify the model using a wide range of P/N_o values.

4.3.1.2 Channel-Trace Driven Emulation

We construct a discrete time event emulator in MATLAB to evaluate the efficacy of our mode selection algorithm. For simplicity, we consider a topology wherein an M_4 antenna transmitter serves a subset ($K_{1:4}$) of 8 possible single antenna receivers. We consider $K_{1:4}$ because 802.11ac supports only up to four concurrent receivers (but up to 8 collective receive antennas).

Because we seek to isolate the effects of mode and user selection, we design an emulation engine that ignores collisions and retransmissions. Transmissions are executed as packets become available on a first-come first-serve basis. At the beginning of every transmission event, each mode and user group is selected based on how many available packets
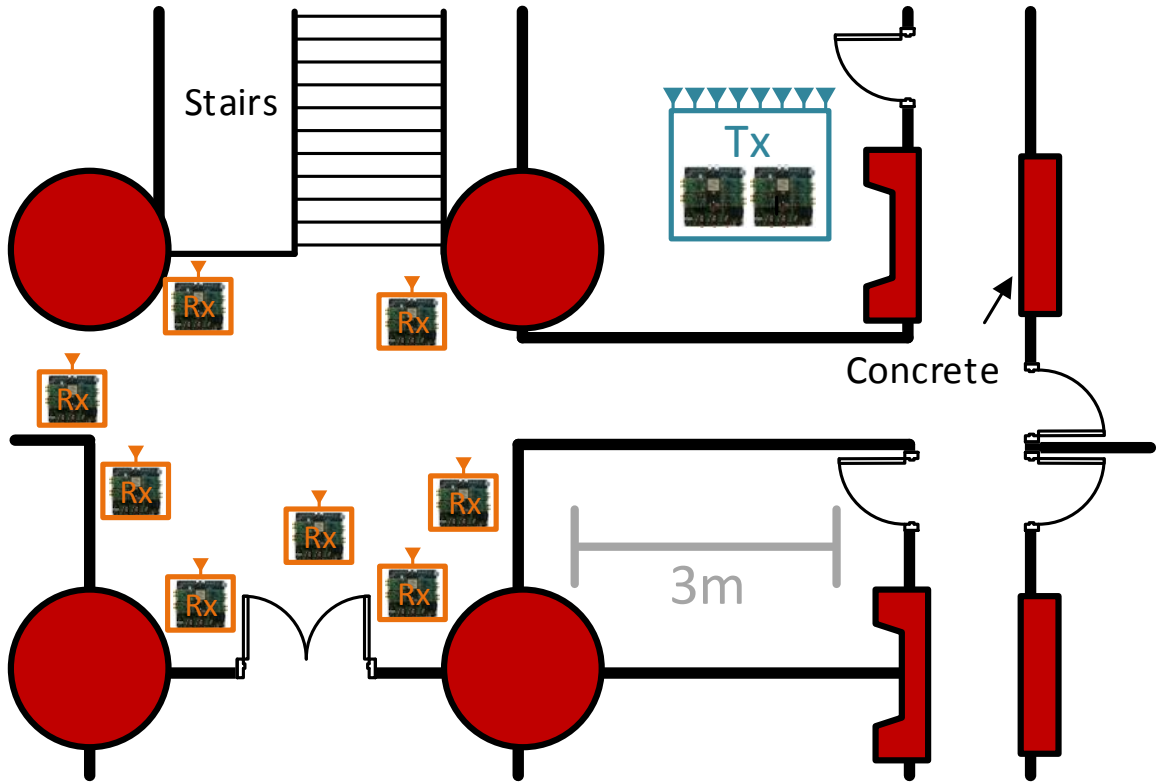
Figure 4.3: Experimental topology.

are in each receiver's queue (up to 802.11ac's 64 packet frame aggregation maximum). Packets are modeled to arrive as a Poisson process and the input traffic is defined as aggregate offered load (cumulative generated traffic) to all receivers.

We use our measured, OTA omnidirectional SNR values at each node as a channel trace to determine the expected per-user MCS. The variation in the measured values allows us to consider heterogeneous channels with different omnidirectional SNRs on the AP to client links

Since the 802.11ac standard supports unequal length parallel data streams and unequal per-user datarates, we allow the emulator to transmit parallel payloads with different MCS rates and frame aggregation sizes. Each emulation was run for 100 emulated seconds. As shown in 4.3, the emulation was conducted assuming an 80 MHz bandwidth and 4 μs symbol times. The CBFRs were quantized at 16 bits per subcarrier and a subcarrier grouping

of 2. The AP is allowed up to 4 transmit antennas and it serves a group of 8 single antenna receivers ($M_{max} = 4$ and $K_{max} = 8$).

Parameter	Value
Bandwidth	80 MHz
`cbfr` Bits per Subcarrier	16
`cbfr` Subcarrier Grouping	2
Symbol Time	4μs
M_{max}	4
K_{max}	8

Table 4.3: Simulation parameters.

4.3.2 Expected SINR Calculation Accuracy

We validate the accuracy of PUMA's SINR estimation method used for datarate inference. Our experiment consists of performing 8 OTA transmissions for all $[M_{1:8}, K_{1:8}]$ topologies as discussed in Sec. 4.3.1.1.

Using the measured omnidirectional SNR for each receiver, we use Eq. (4.3) to predict the per-user SINR for each receiver. Perfect SINR results are not expected as Eq. (4.3) is based on general MU-MIMO system scaling as opposed to CSI. However, the resulting error is almost zero mean and with a standard deviation of 2.43 dB as shown in Fig. 4.4. This error is tolerable because the standard deviation is approximately equivalent to SINR range for each modulation rate shown in Table 4.2. The use of the MCS table diminishes this error by effectively truncating it. We explore the effect of this error mitigation in Sec. 4.3.3.

Thus, the expected value equation for per-user SINR is accurate *even without considering measured channel matrices*. This holds true for indoor Wireless LANs because the channel vectors are relatively orthogonal, meaning that the channel matrices used for MU-MIMO transmissions are well conditioned (as experimentally verified in [33]).
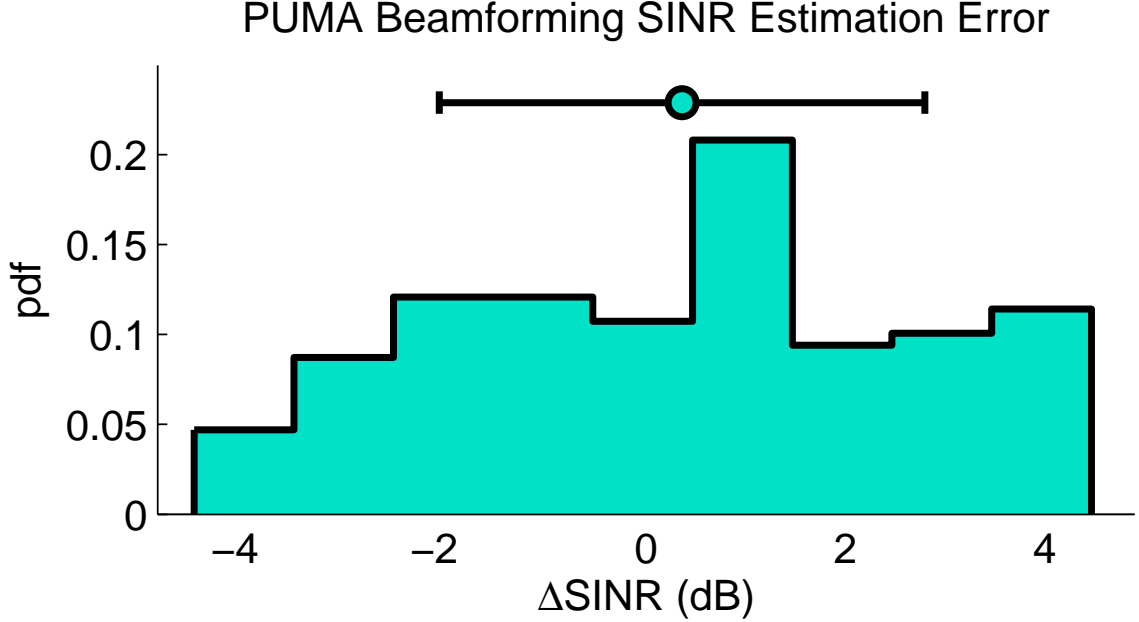
Figure 4.4: Measured estimation error of Eq. (4.3). μ=0.36, σ=2.43 dB.

4.3.3 PUMA Expected Datarate Calculation Accuracy

Using our trace-driven emulation methodology, we compare PUMA against a post-sounding, exhaustive search baseline. This baseline method forgoes the use of the SINR estimation algorithm and employs the actual measured MU-MIMO SINRs. This effectively represents the best case result of using Eq. (4.1) post-sounding after exhaustively measuring each potential receiver's CSI. Fig. 4.5 shows the comparative results of our emulation.

A perfect transmitter would send the incoming packets at a rate equivalent to their arrival at the AP. However, because of the overhead time T_{OH} in Eq. (4.4) required for each packet transmission, this is not possible. Instead, we show that PUMA transmits at the highest feasible portion of that rate by selecting transmission modes that maximize the SINR and thus the MCS while minimizing the transmission overhead. When the aggregate throughput saturates, the maximum possible throughput is achieved.

Observe in Fig. 4.5 that full knowledge of channel state from an exhaustive sounding process only results in a 3% increase in saturation throughput. Additionally, over all

aggregate offered loads, knowledge of full CSI only results in a maximum 7% increase in throughput. Note that the exhaustive search method's performance does not consider the overhead incurred from sounding all potential receivers (like PUMA, it only considers sounding overhead from the users actually served for each transmission). Had the full sounding overhead been considered, the performance of the exhaustive search method would perform significantly worse and its saturation throughput would be far lower.

Fig. 4.5 also highlights PUMA's datarate inference method's robustness to error. The variation between the measured and estimated SINR values (shown in Fig. 4.4) has a standard deviation of 2.5 dB, approximately equal to the SNR range of each MCS. However, even given this error, we observe that the performance difference between estimating the MU-MIMO SINR and measuring it is minimal. This is a direct result of the effective truncation of the estimated SINR metric, using the minimum SINR table. This truncation essentially smooths the estimated SINR metric and results in similar performance to full
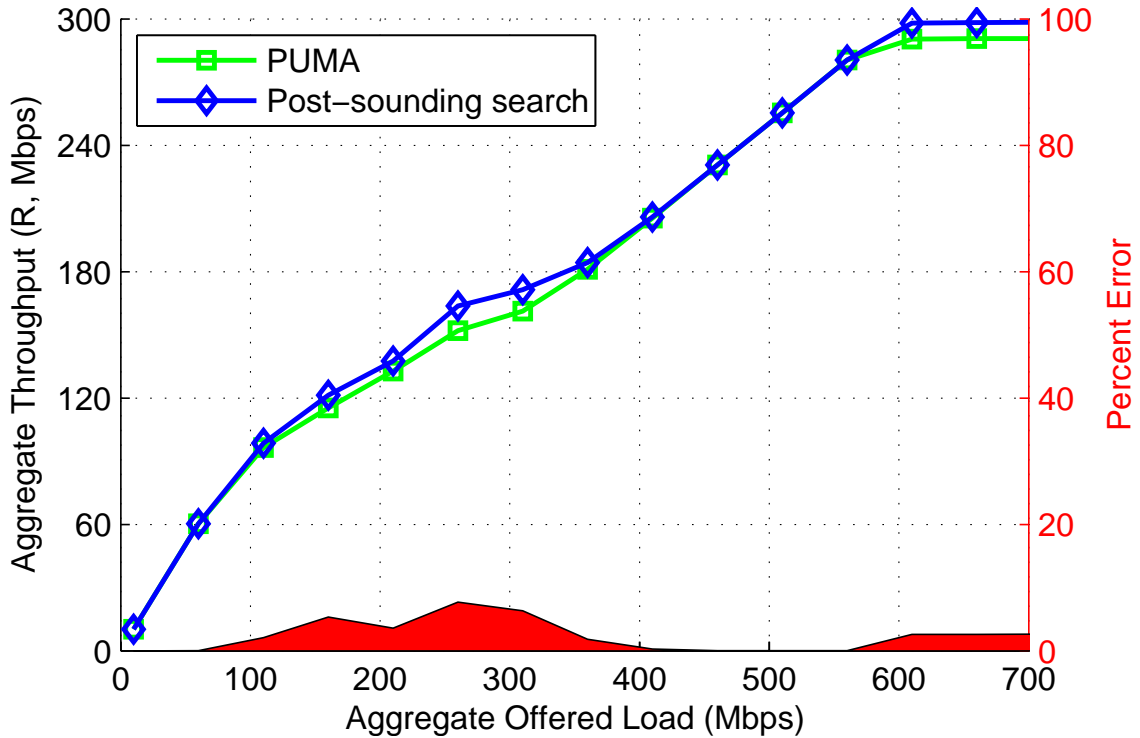
Figure 4.5: Comparison between PUMA and post-sounding, exhaustive search.
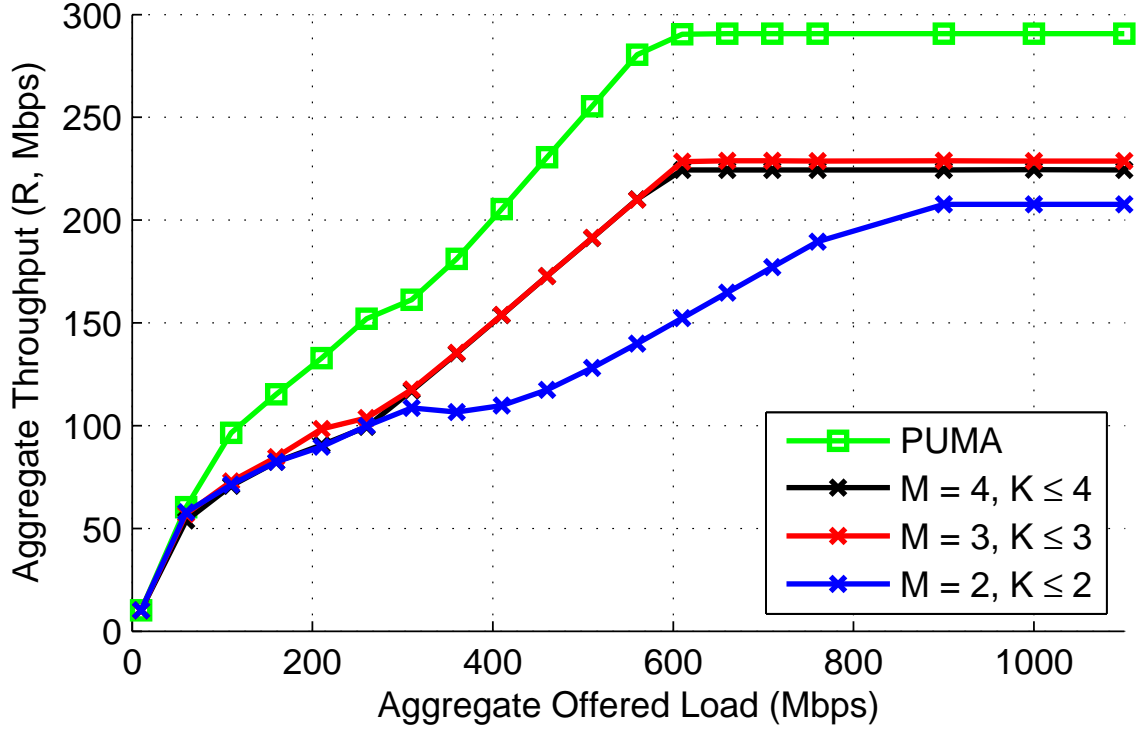
Figure 4.6: Comparison between PUMA and fixed modes ($[M_{2:4}, K_{1:4}]$).

knowledge of the MU-MIMO SINR without a post-sounding, exhaustive search.

4.3.4 Mode and User Selection Performance

We now compare PUMA to fixed mode selection indicative of the default method used in 802.11ac. Specifically, we compare the performance of PUMA to static $[M_{2:4}, K_{1:4}]$ topologies in Fig. 4.6 for a range of offered loads. For each possible M, we show the best fixed K value for sake of presentation. These static modes are all potential choices for PUMA's exhaustive search and thus observing how these parts contribute to our collective algorithm illustrates how well PUMA handles varying M, K, b_i, and measured omnidirectional SNR values.

So as not to unfairly disadvantage the fixed modes, we permit the fixed $[M, K]$ topologies to serve any number of users less than or equal to K. For example, given our omnidi-

rectional channel measurements, when $M=3$, K_{max} is 3 but individual transmissions can be K_2 or K_1, depending on how many users have packets available in their queues.

Although PUMA is inherently a combination of all fixed modes, observe that for any given offered load, no fixed mode performs comparably to PUMA (except at 10 Mbps where all methods are equal). This result suggests that it is not only enough to know the "best" $[M, K]$ combination given some offered load but also it is necessary to dynamically select between potential modes before every transmission, depending on user backlog and omnidirectional SNR.

The relationship between user backlog and omnidirectional SNR is the key interaction the fixed mode topologies fail to consider. Due to the sounding overhead incurred from MU-MIMO transmissions, larger frame aggregation rates are required to properly amortize the cost of employing parallel streams. Thus, users with lower omnidirectional SNRs (resulting in lower achievable datarates) must have more backlogged packets to be efficiently grouped in an MU-MIMO transmission. PUMA considers this interaction and thus does not transmit to low omnidirectional SNR users in higher order modes until the user's backlog is large enough. Thus, the MU-MIMO sounding overhead efficiently and dynamically amortizes the sounding overhead for each transmission.

A concern for PUMA is its potential to unfairly starve users that have consistently less backlogged traffic or poor omnidirectional SNRs. However, unfair scheduling is not a detrimental effect of PUMA, rather, it is a complementary issue. In the most basic sense, PUMA enumerates a list of selections (potential modes and user groups), assigns a metric to each, and selects the best. Our current metric is the aggregate throughput of the system determined by user backlog and expected datarate. To ensure a fair transmission system, an AP can employ any existing proportional fair scheduling algorithm (e.g., [34]) and adjust PUMA's metric accordingly. This would manifest as simply adding a scaling factor generated by a fair scheduling algorithm to step 4 of the PUMA protocol described

in Sec. 4.1.4.

Given the measured channels from our experiments, PUMA provides an aggregate saturation throughput increase of approximately 65 Mbps or 30% over the best fixed mode ($[M_3, K_{1:3}]$). The improvement in the saturation region with higher offered loads is of key importance for high congestion scenarios. While the other scenarios all saturate to similar aggregate throughputs (albeit at different rates), PUMA's 30% saturation throughput increase illustrates how efficient overhead amortization through adaptive mode selection based on user link state and backlog allows the same AP to improve the performance of a given topology.

Finally, PUMA's user selection mechanism is a direct result of step 4 of the algorithm discussed in Sec. 4.1.4. Once all potential mode and user groups are enumerated and assigned an aggregate throughput metric, the best mode and user combination is selected. However, the best mode and user combination is not guaranteed to be unique. PUMA's aggregate throughput metric is dependent on per-user omnidirectional SNR and backlog. Thus, multiple mode and user combinations may be assigned the same aggregate throughput metric (e.g., the simplified example shown in Fig. 4.1). In such cases, a random selection or additional fairness metric can be used.

Nevertheless, real systems rarely have homogenous traffic arrival rates or per-user omnidirectional SNRs. Thus, although PUMA cannot guarantee a unique mode and user combination, it provides a unique selection with high probability.

MU-MIMO PROTOCOLS FOR STABLE CHANNELS

The previously discussed protocol, PUMA, avoids the pitfalls of highly variable channel environments by utilizing omnidirectional SNR (as opposed to measured CSI) as a decision metric for MAC layer decisions. While the use of omnidirectional SNR is motivated by its temporal stability compared to measured CSI [27], it is inherently less accurate than using measured CSI.

In the case of variable MU-MIMO environments, given the overhead inherent to the channel sounding process (as discussed in Sec. 4.2), the benefits of avoiding channel sounding outweigh the use of a less accurate metric.

However, when information regarding the stability of the MU-MIMO channel is known beforehand, (e.g., when using the UHF band as discussed in Sec. 3.3, a different approach to MU-MIMO overhead management is vital to allow these multi-stream transmissions to achieve their full potential. Thus, for temporally stable MU-MIMO environments, we present **F**eedback **R**emoval with **O**pportunistic **Z**ero-overhead channel **E**stimatio**N** (FROZEN), an MU-MIMO protocol that avoids sounding overhead by leveraging the relative temporal stability of particular MU-MIMO environments and uplink packet transmissions to infer and track per-user CSI for concurrent transmissions without a costly sounding procedure.

The remainder of this chapter is organized as follows: Sec. 5.1 provides an overview of FROZEN, Sec. 5.2 discusses ICSIQLE, the novel channel estimate tracking method that necessary for informing FROZEN's MAC-layer decisions, and finally, Sec. 5.3 provides an experimental evaluation of FROZEN.

5.1 FROZEN Overview

In current MU-MIMO systems such as IEEE 802.11ac, CSI is obtained by first transmitting an explicit sounding packet in one direction, and then having the receiver decode and feed back the channel estimate as compressed data in the other direction [3] as shown in Fig. 5.1(a). Although such multi-stream transmissions are key to increasing spectral efficiency and therefore, system capacity, the overhead required for the measurement of CSI can easily overwhelm the spatial multiplexing gains of MU-MIMO precoding[1].

(a) Explicit MUBF   (b) Active Implicit MUBF   (c) Passive Implicit MUBF
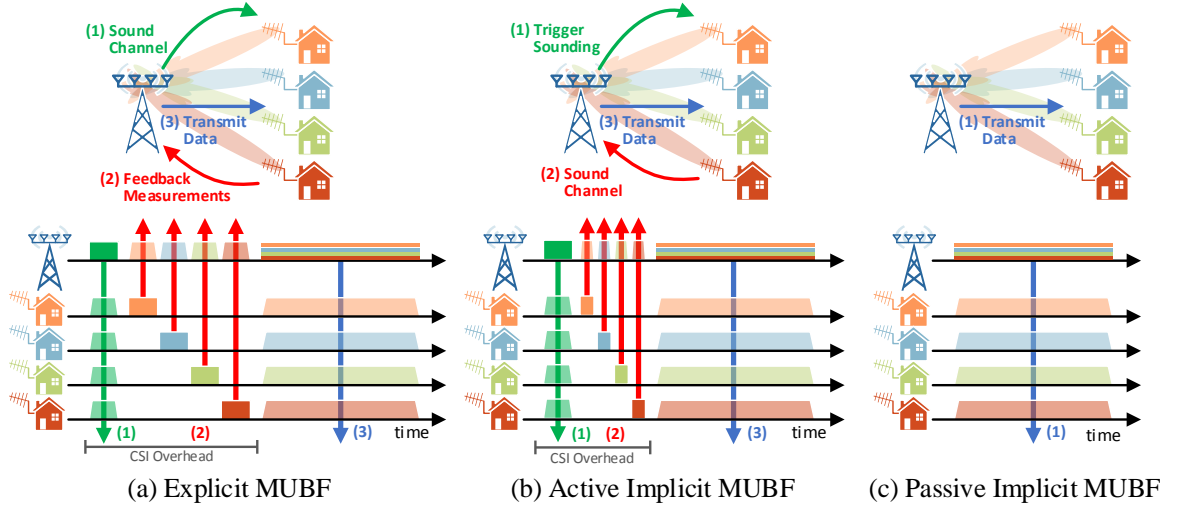
Figure 5.1: Generalized overhead diagram of different types of channel sounding.

This amount of overhead is a result of two key factors: *(i)* the significant amount of data required to feedback a digitized representation[2] of the sounding packet on a per-antenna

---

[1]802.11ac feedback size can range from 1.6 kb to 329 kb per client depending on the number of antennas on both AP and client [3, 2].

[2]While the training sequence is only a few OFDM symbols in length, the channel estimate is computed

and per-subcarrier basis and *(ii)* the potentially high frequency of CSI measurement (eg., a per-packet basis) due to the MU-MIMO transmitter's need for updated CSI.

To combat this overhead, FROZEN utilizes three key techniques:

1. **Implicit Sounding**: A process where the CSI between the AP and client is inferred by training sequences sent from the client nodes. This process, as shown in Fig. 5.1(b), reduces overhead because the sounding sequence does not have to be serially fed back to the AP per subcarrier and thus the sounding overhead per receiver is simply the amount of time required to send the sounding packet (for 801.11ac, the `NDP` discussed in Sec. 4.2.2). This sounding method has been studied extensively and novel reciprocal calibration methods [32] allow for its feasibility.

2. **Passive CSI Collection**: A CSI collection mechanism that implicitly infers CSI from all previously transmitted uplink packets (e.g., data, acknowledgment) using the channel training field of any OFDM transmission (for 802.11ac this is the VHT Long Training Field (VHTLTF)). This indirect sounding procedure results in no deliberate sounding packets being sent and is shown in Fig. 5.1(c).

3. **Indirect Sounding**: In the case that the CSI for a particular user either has not been collected or is outdated[1], a FROZEN enabled AP will indirectly sound the user by transmitting one packet in a single-user manner and implicitly measuring that user's CSI from the resulting acknowledgment. Thus, the AP can force a user's CSI update without any additional overhead.

While implicit sounding alone *reduces* sounding overhead, passive implicit sounding completely *removes* sounding overhead because no deliberate sounding packets are exchanged between the AP and the client. Channel sounding effectively piggybacks on all

---

on a per-subcarrier basis. The training sequence transmits energy on all of these subcarriers in parallel while the feedback packet must serialize the magnitude and phase measurements for each subcarrier to send it back.

[1]Determining if a user's CSI is outdated is discussed in Sec. 5.2.

existing uplink data transmissions. Additionally, when new CSI is required, FROZEN's indirect sounding mechanism *still* transmits data so no additional overhead is generated.

Because FROZEN allows for an AP to store (and track the validity of) a particular user's channel estimate, it allows for an AP to make offline MAC decisions. Thus, user selection algorithms based on Eq. (4.1) are feasible and thus the inherent receiver separability of a group of users due to a particular band/environment combination are irrelevant.

The efficacy of FROZEN is contingent on the ability of the AP to determine if and when a particular user's measured CSI is outdated and thus requires re-measurement. To that end, we develop a novel CSI tracking algorithm **I**mplicit **CSI Q**ua**L**ity **E**stimate (IC-SIQLE)[1], that not only tracks changes in magnitude but also changes in phase of the CSI measurement[2].

The method by which FROZEN uses ICSIQLE to actually track and estimate if subsequent transmissions require resounding is discussed in Sec. 5.2.3.

5.2 Implicit CSI QuaLity Estimate (ICSIQLE)

5.2.1 Absolute vs. Relative CSI Comparisons

As discussed in Sec. 5.1, FROZEN relying on previously measured CSI necessitates a method of determining if and when a particular client's previously collected CSI is still suitable for use in beamforming to that client. Outdated CSI results in improperly targeted beams that can result in inter-client interference.

However, tracking the change between subsequent channel measurements for a particular user (that user's $\vec{h}$ vector) is not as simple as comparing a particular user's CSI at some time t, $\vec{h}_t$, to its subsequent $\vec{h}_{(t+\Delta t)}$.

---

[1]pronounced "icicle"

[2]Existing CSI tracking methods such as autocorrelation or protocols like AFC ([35]) fail to accurately track phase changes due to phase wrapping, a key problem in MU-MIMO channel tracking [36] that ICSIQLE avoids (which will be discussed next in Sec. 5.2.2)

The aforementioned method compares the *absolute* difference between a user's measured $\vec{h}$'s. While absolute comparisons suffice for generalized channel statistics like autocorrelation (where a large number of measurements are taken as shown in Fig. 3.13), such comparisons fail to accurately track instantaneous changes in the measured channel due to the nature of Orthogonal Frequency-Division Multiplexing (OFDM) transceivers and the cyclic nature of phase.

A large absolute difference between two measured phases may have no effect on the performance of a received, beamformed transmission for two key reasons:

1. Arbitrary phase offsets may be present in any one measurement of a particular user's $\vec{h}_t$ due to receive gain control or timing estimation jitter, normal effects in OFDM transceivers that are compensated by the OFDM channel equalization step, yet become added to the measured $\vec{h}$ at the receiver [37].

2. Second, phase, unlike magnitude, is cyclic and thus it "wraps," reducing the utility of absolute comparisons. For example, if one element of $\vec{h}_t$ has a measured phase of 359° and the corresponding element in $\vec{h}_{(t+\Delta t)}$ has a measured phase of 1°, an absolute comparison of the two elements would result in a phase difference of 358° while in reality the phase could have only changed 2°. For this reason, the 802.11ac taskgroup has discussed how absolute phase comparisons are not useful for tracking CSI measurement variation [36] (tracking schemes require "phase immunity").

Current CSI tracking metrics that consider absolute phase (such as AFC's CN_0 [5]) fail to capture the true variation of the channel (as opposed to phase wrapping effects) and thus cannot accurately track the potential resulting performance loss. However, since our ICSIQLE metric is dependent on relative phase differences between individual elements of $\vec{h}$, it does accurately track channel variation and the resulting performance differential.

In Sec. 5.2.1 we show that relative phase differences are crucial to tracking channel variability, first on an element by element basis and then as the $\vec{h}$ as a whole. Then, in

Sec. 5.2.2 we show that using ICSIQLE, we can accurately track the potential performance loss due to channel variation thus enabling FROZEN.

5.2.2 ICSIQLE's Relative CSI Comparison Method

ICSIQLE's relative comparison method is based on tracking changes of relative phase between elements of a particular user's $\vec{h}$ measurement[1]. Thus, when comparing $\vec{h}_t$ and the subsequent $\vec{h}_{(t+\Delta t)}$, ICSIQLE's relative comparison method does the following:

Absolute (CN_0)	Relative (ICSIQLE)
$\vec{h}_t \leftrightharpoons \vec{h}_{(t+\Delta t)}$	$\vec{h}_t \leftrightharpoons \vec{h}_{(t+\Delta t)}$
$\angle\vec{h}_{t_1} \leftrightharpoons \angle\vec{h}_{(t+\Delta t)_1}$	$\Delta(\angle\vec{h}_{t_1}, \angle\vec{h}_{t_2}) \leftrightharpoons \Delta(\angle\vec{h}_{(t+\Delta t)_1}, \angle\vec{h}_{(t+\Delta t)_2})$
$\angle\vec{h}_{t_2} \leftrightharpoons \angle\vec{h}_{(t+\Delta t)_2}$	$\Delta(\angle\vec{h}_{t_2}, \angle\vec{h}_{t_3}) \leftrightharpoons \Delta(\angle\vec{h}_{(t+\Delta t)_2}, \angle\vec{h}_{(t+\Delta t)_3})$
$\angle\vec{h}_{t_3} \leftrightharpoons \angle\vec{h}_{(t+\Delta t)_3}$	$\Delta(\angle\vec{h}_{t_1}, \angle\vec{h}_{t_3}) \leftrightharpoons \Delta(\angle\vec{h}_{(t+\Delta t)_1}, \angle\vec{h}_{(t+\Delta t)_3})$

Table 5.1: CSI Phase Tracking Metric Methods (example for M_3 transmit antenna case).

To actually perform this relative phase comparison, we will use the product of the row vector $\vec{h}$ by itself as $\vec{h}^{\mathrm{H}}\vec{h}$ where $(\cdot)^{\mathrm{H}}$ is denotes the Hermetian operation (conjugate transpose). This multiplication is performed on $\vec{h}_t$ and $\vec{h}_{(t+\Delta t)}$ before comparing the two. For the same example as shown in Table 5.1, the aforementioned vector multiplication accom-

---

[1]Recall that each element of the $\vec{h}$ corresponds to the measurement between each of the AP's transmit antennas and the clients antenna.

plishes the following:

$$
\vec{h}^{\mathrm{H}}\vec{h} = \begin{pmatrix} h_1^* \\ h_2^* \\ h_3^* \end{pmatrix} \begin{pmatrix} h_1 & h_2 & h_3 \end{pmatrix}
$$

$$
= \begin{pmatrix} |h_1|e^{-j\angle h_1} \\ |h_2|e^{-j\angle h_2} \\ |h_3|e^{-j\angle h_3} \end{pmatrix} \begin{pmatrix} |h_1|e^{j\angle h_1} & |h_2|e^{j\angle h_2} & |h_3|e^{j\angle h_3} \end{pmatrix}
$$

$$
= \begin{bmatrix} |h_1||h_1|e^{j(\angle h_1 - \angle h_1)} & |h_1||h_2|e^{j(\angle h_1 - \angle h_2)} & |h_1||h_3|e^{j(\angle h_1 - \angle h_3)} \\ |h_2||h_1|e^{j(\angle h_2 - \angle h_1)} & |h_2||h_2|e^{j(\angle h_2 - \angle h_2)} & |h_2||h_3|e^{j(\angle h_2 - \angle h_3)} \\ |h_3||h_1|e^{j(\angle h_3 - \angle h_1)} & |h_3||h_2|e^{j(\angle h_3 - \angle h_2)} & |h_3||h_3|e^{j(\angle h_3 - \angle h_3)} \end{bmatrix}
$$

$$
= \begin{bmatrix} |h_1|^2 & |h_1||h_2|e^{\mathbf{\Delta}(\angle_{1,2})} & |h_1||h_3|e^{\mathbf{\Delta}(\angle_{1,3})} \\ |h_2||h_1|e^{\mathbf{\Delta}(\angle_{2,1})} & |h_2|^2 & |h_2||h_3|e^{\mathbf{\Delta}(\angle_{2,3})} \\ |h_3||h_1|e^{\mathbf{\Delta}(\angle_{3,1})} & |h_3||h_2|e^{\mathbf{\Delta}(\angle_{3,2})} & |h_3|^2 \end{bmatrix} \tag{5.1}
$$

We observe that this operation removes not only the measurement phase error, but also other common phase angles in the CSI that have no bearing on beamforming accuracy. The result populates the off-diagonal entries with all complex conjugate pairs such that all remaining angles are differences (as shown in Table 5.1 In general, this property scales to larger CSI vector sizes. An important feature of ICSIQLE is that it avoids selecting an arbitrary path as a reference and biasing the measurement based on that path's measurement quality as apparent from all combinations of phase differences in the $\vec{h}^{\mathrm{H}}\vec{h}$ expansion.

Since we wish to track channel variation with respect to time, we define ICSIQLE as the difference between the complex $M \times M$ matricies formed by the operation in Eq. (5.1) for the CSI estimated at two points in time. The output scalar value, $\text{ICSIQLE}_{t,(t+\Delta t)}$, defined at time index $(t + \Delta t)$ with respect to an initial CSI measurement at time t represents the "quality" of h_t if it were used to beamform at time $t + \Delta t$. The definition for ICSIQLE,

taken as a mean of the Frobenius norms[1] of the differences across N subcarriers, is:

$$\text{ICSIQLE}_{t,(t+\Delta t)} \equiv \frac{1}{2N} \sum_{n=1}^{N} \left\| \frac{h_{t,n}^{\text{H}} h_{t,n}}{\|h_{t,n}\|} - \frac{h_{(t+\Delta t),n}^{\text{H}} h_{(t+\Delta t),n}}{\|h_{(t+\Delta t),n}\|} \right\|_{\mathbf{F}} \tag{5.2}$$

5.2.3   Using ICSIQLE to determine CSI Staleness

When implemented on an AP, FROZEN relies on constant collection of CSI from uplink data and control packets from associated users. For each user, the access point tracks: *(i)* the last known implicit CSI estimate $\vec{h}_t$, *(ii)* the SNR or noise variance estimate for that CSI, *(iii)* the current Exponentially-Weighted Moving Average (EWMA) measure of previous ICSIQLE measurements, *(iv)* an arbitrary ICSIQLE threshold value $\mathbf{I}_{th}$, and *(v)* the last computed value of T_{VALID}. When a new CSI estimate becomes available at $(t+\Delta t)$, these values are updated online with the value of T_{VALID} used to tag the new CSI with a "use by" time:

$$\mathbf{I}_1 := \text{ICSIQLE}_{t_0,t_1},$$

$$\text{EWMA}(\mathbf{I})_1 := \mathbf{I}_1 + \alpha \cdot \text{EWMA}(\mathbf{I})_0$$

$$T_{\text{VALID}} := \mathbf{I}_{th}/\text{EWMA}(\mathbf{I})_1. \tag{5.3}$$

The key insight behind FROZEN is that at time t_1, the moment a new implicit CSI estimate is acquired by the system, FROZEN cannot exactly predict the channel at some future t_2 when that CSI estimate may be used to precode a new MU-MIMO transmission. However, based on the observed rate of change of a particular user's MIMO channel due to client or environmental mobility, FROZEN *can* bound the expected performance of a future transmission by limiting the amount of time, T_{VALID}, a collected CSI measurement

---

[1]The Frobenius norm of the difference is the sum of the magnitudes of each difference so the effect on change of each element of $\vec{h}^{\text{H}}\vec{h}$ is considered.

can be used.

When FROZEN determines that a user's CSI estimate is stale, it has two choices that are left to the MAC-layer. First, the CSI can be discarded resulting in the removal of the user from any future beamformed transmissions until an uplink packet can refresh the CSI. We discuss the use of this policy in more detail in Section 5.3.3.

Secondly, FROZEN also offers the possibility of a soft decision in terms of treating the accuracy of implicit CSI as a continuum rather than a binary decision. Multiple T_{VALID} values can be defined for different confidence bounds on the expected beamformed performance using an old CSI value. Specifically, depending on the available DoFs, less accurate CSI could still be acceptable for MU-MIMO performance as shown in PUMA's use of MU-MIMO scaling properties (Sec. 4.1.2). This could drive an MCS or group-selection algorithm that can utilize less reliable CSI.
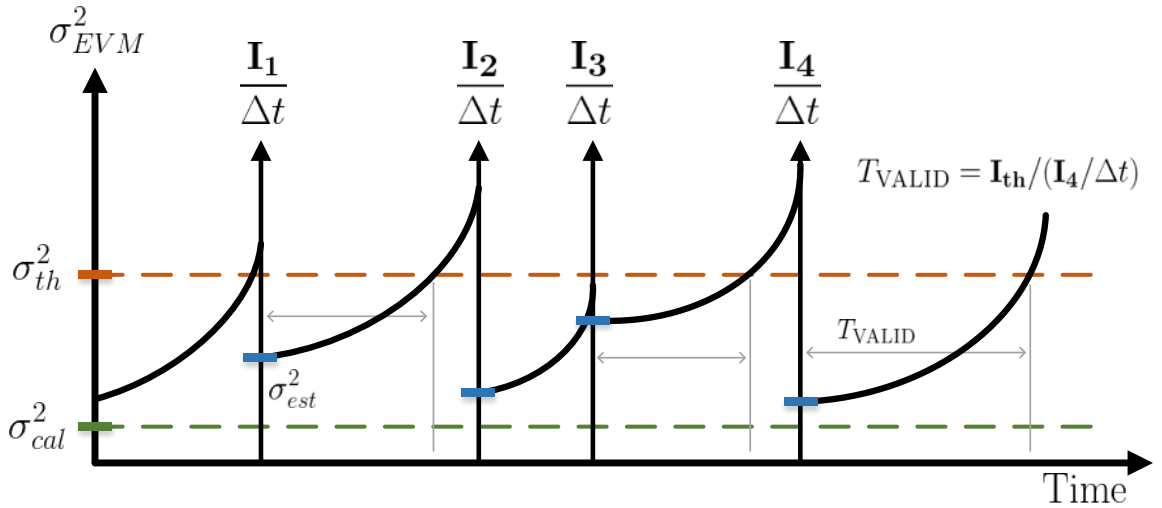
Figure 5.2: A timeline of FROZEN, showing how new opportunistic CSI estimation events are used, how T_{VALID} is applied, and how calibration variance σ^2_{cal} and CSI estimation variance σ^2_{est} can be used to adjust FROZEN's decisions based on the performance of the underlying radio hardware and the accuracy of the CSI measurement.

In Fig. 5.2, we present a theoretical timeline showing how the FROZEN evolves in a working system and how thresholds for CSI quality are set. FROZEN predicts future confidence bounds of implicit CSI based on the current estimated rate of change of ICSIQLE,

represented by $\mathbf{I}_t$. We represent opportunistic receptions of a given user's packets as delta functions, providing a new implicit CSI estimate and trigger re-calculation of EWMA($\mathbf{I}_t$). For each reception, a new T_{VALID} for the last received CSI is calculated. If the quality of CSI is changing rapidly due to high channel variability, then FROZEN can not utilize a previous CSI estimate for very long. At t_4, the most-recent CSI estimate is overheard, and the estimation variance σ^2_{est}, calculated according to the methods discussed in Sec. 5.3.1, is low, yielding a T_{VALID} that is relatively high. Now, if the access point's scheduler ever wishes to transmit an MU-MIMO transmission including the user within T_{VALID}, it can be reasonably certain that the packet will be successful using the previously collected CSI.

The Error-Vector Magnitude (EVM) variance threshold that determines $\mathbf{I}_{th}$ is determined by the host system; for example, the variance after which the transmitter must change MCS to guarantee reception or after which transmissions have unacceptable probability of failing.

FROZEN's ICSIQLE decision threshold is designed such that it is adaptable to both the underlying accuracy of the radio hardware used and the instantaneous SNR of the CSI estimate itself. Given these parameters, the system designer can set the appropriate threshold to maximize performance *i.e.,* ensure that a FROZEN enable AP re-sounds only when necessary.

5.3   FROZEN Evaluation

5.3.1   Multi-Stream Noise and Channel Measurements

We implement a beamforming and channel-sounding framework based around the 802.11af WiFi standard, in order to compare against 2.4/5 GHz implementations of 802.11ac since the physical OFDM layers of 802.11ac VHT 40 MHz and 802.11af TVHT 5 MHz are identical aside from the radio sampling rate [38, 3]. Over-the-air transmissions are conducted

using the WARPLab framework discussed in Sec. 3.1.

Standards-compliant piloted phase correction is implemented to compensate for sample timing drift during the payload, necessary to yield accurate measurements of received EVM. Two copies of the cyclicly-shifted L-STF provide ample time to account for AGC settling and the WARPLab trigger jitter. Timing recovery and carrier-frequency offset correction operates on the L-LTF field, while we utilize the TVHT-LTF symbols with the 802.11af spatial spreading matrix for channel estimation.

Our analysis of FROZEN requires updated CSI for each downlink beamformed packet as well as CSI noise variance. On the multi-stream (one stream per client) downlink beamformed packet, we send an extra *non-beamformed* TVHT-LTF immediately following the beamformed TVHT-LTF as shown in Fig. 5.3. When scaled according to the expected degradation in beamformed SNR compared to broadcast SNR [39] to avoid AGC clipping, this extra TVHT-LTF provides an estimate of the wireless channel at the moment that the beamformed packet is sent. This addition is not necessary for FROZEN's operation, but it allows us to validate the protocol.
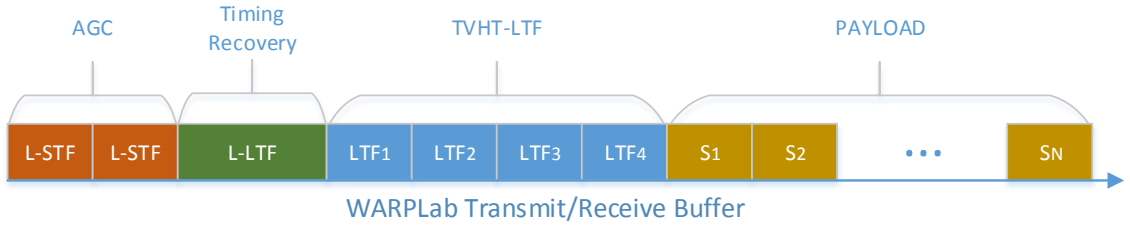
Figure 5.3: Multi-stream beamformed packet format implements the 802.11af TVHT 5 MHz PHY, utilizing the TVHT-LTF for multi-stream channel sounding.

5.3.2 ICSIQLE Performance Evaluation

5.3.2.1 The Necessity of Relative Phase for CSI Tracking

Before directly evaluating ICSIQLE we first evaluate its premise: relative angle comparisons of $\vec{h}$'s is key to actualy tracking the variability of the channel. Using the method

discussed in Sec. 5.3.1 and the hardware setup discussed in Sec. 3.1, we measured a static $[M_4, K_1]$ channel in a laboratory environment.

We show the measured phase over a 30 second time interval in Fig. 5.4. For simplicity, we only show the measurements for an arbitrary subcarrier; however, the conclusions drawn subsequently would be identical regardless of choice.
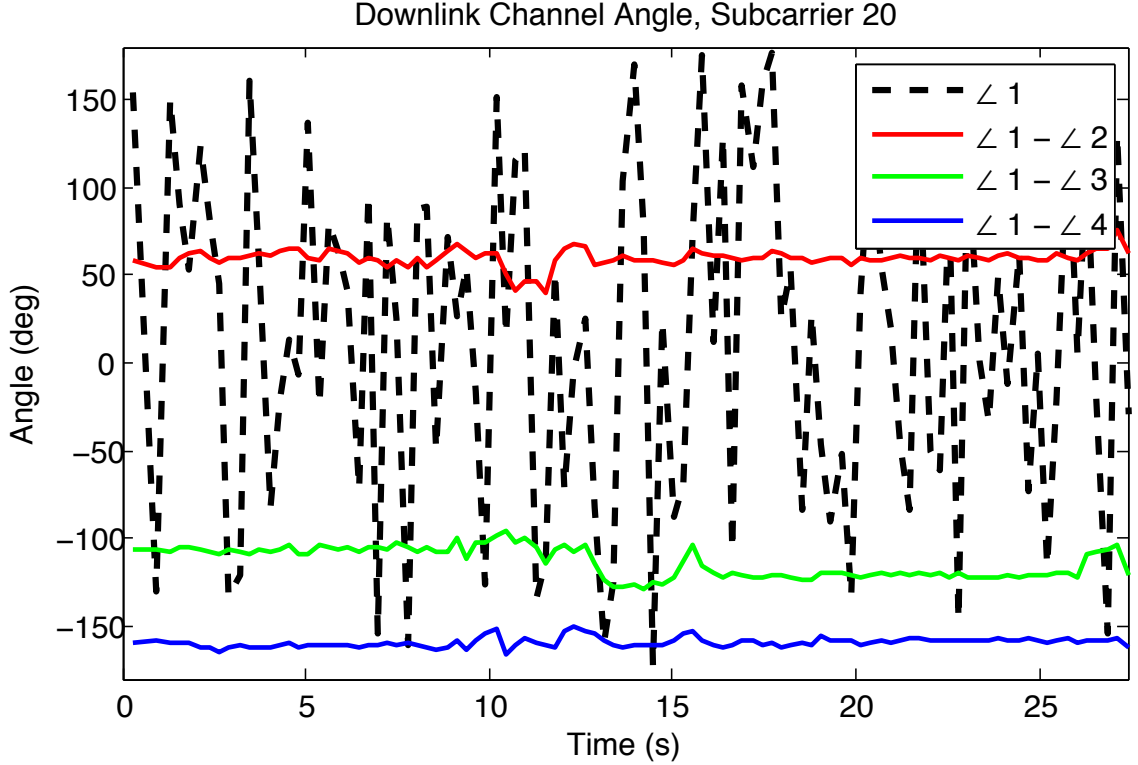
Figure 5.4: Using the difference in phase between different transmit antennas is much more reliable than depending on absolute measures of channel phase.

The dashed black line is the absolute measured phase of the antenna path from the AP's first antenna to the client. Observe the high variation of the measured phase when considering absolute phase differences. If tracking absolute phase, this static, laboratory environment channel appears to be highly variable and any MAC protocol based on this metric would be unnecessarily resounding the channel.

However, when considering *relative* phase, as shown by the red, green, and blue lines (antennas 2, 3, and 4 relative to antenna 1), we see that the phase variability of the channel

is actually low as we hypothesized given the measurement environment. This motivating experiment clearly shows how relative phase measurements are key to truly tracking the variability of the MU-MIMO channel and can solve the problems discussed in the 802.11ac task group discussed in [36].

5.3.2.2 ICSIQLE: MU-MIMO Performance Tracking

In order to actually validate the predictive power of ICSIQLE, we perform a series of over-the-air experiments where we explicitly beamform in a $[M_4, K_1]$ office environment to a set of nodes. The experiments consist of a single sounding packet at $t = 0$, followed by a set of downlink beamformed transmissions utilizing the same CSI from the original sounding packet for precoding.

We chose the Single-User Beamforming (SUBF) $[M_4, K_1]$ topolgy instead of an MU-MIMO topology because the purpose of a per client tracking metric is to measure the CSI variability for a *particular user*. As discussed in Sec. 2.2, the two key performance determining factors for an MU-MIMO transmission are channel variability and receiver separability. Because ICSIQLE or any other CSI tracking method's goal is to measure channel variability, we remove the effect of receiver separability from this portion of the evaluation to accurately judge how well ICSIQLE tracks channel variability.

Each downlink beamformed transmission also estimates the current *broadcast* downlink channel as discussed in Section 5.3.1. Using this information, we are able to calculate ICSIQLE values between $t = 0$ and subsequent $(t + \Delta t)$'s and correlate that with corresponding measured EVM estimated from the decoded payload of the beamformed packet. This experiment is repeated thousands of times in different static and pedestrian-speed mobile nodes in an office environment. For comparison, we also compute corresponding CN_0 values (based on absolute phase) and show the performance of both metrics in Fig. 5.5.

Fig. 5.5 plots the considered metric and the corresponding measured EVM from trans-
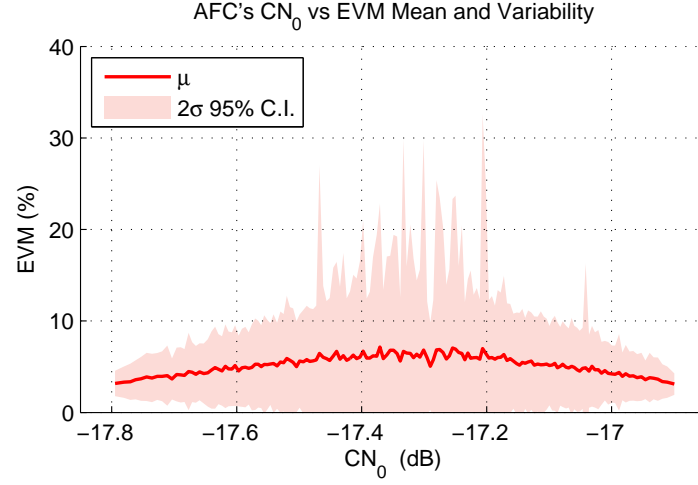
(a) AFC's CN_0 metric [5] employs element-by-element *absolute* phase difference to track the user's CSI and thus cannot accurately predict the performance of MU-MIMO transmissions. Neither the measured performance (EVM) nor the variance of performance are correlated with CN_0.
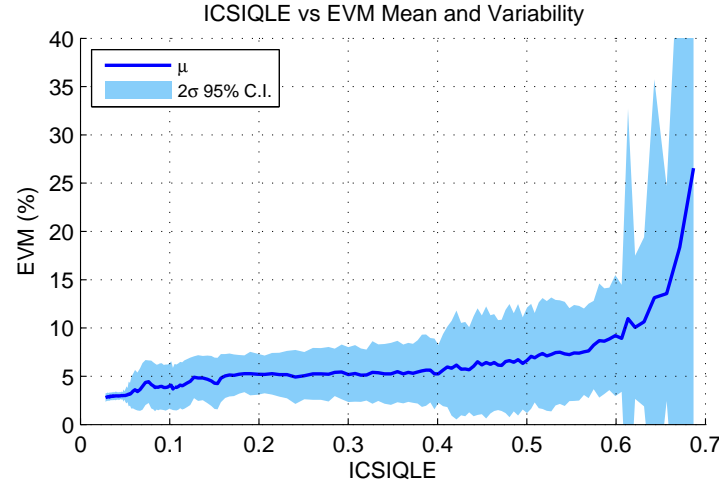
(b) ICSIQLE's leveraging of *relative* phase allows for accurate tracking of beamforming performance. Larger ICSIQLE values indicate larger variations between $\vec{h}_t$ and $\vec{h}_{(t+\Delta t)}$ and the experimentally measured beamformed EVM shows the performance drop at higher variability.

Figure 5.5: Relative Angle (ICSIQLE and Absolute Angle (CN_0) CSI tracking metrics vs. mean beamformed EVM and 2σ beamformed EVM bounds, representing a 95% confidence interval on EVM.

missions resulting from a sounded $\vec{h}_t$ over the channel $\vec{h}_{(t+\Delta t)}$. The shaded regions correspond to the 2σ 95% confidence interval of the measured beamforming performance in terms of EVM. We expect that the larger difference between $\vec{h}_t$ and $\vec{h}_{(t+\Delta t)}$ reported by a

particular metric, the worse (larger) EVM should be.

Observe the uncorrelated performance of the absolute difference based CN_0 metric in Fig. 5.5(a). Not only is mean EVM monotonically increasing as expected, but it also (for this experiment) has a parabolic shape implying that the worst case performance should be at centrally measured CN_0 values. Using this metric for CSI tracking and thus to determine when to resound the channel would result in a significant number of false positives and thus would not actually reduce sounding overhead even when possible.

However, observe the correlated, monotonically increasing shape of our relative-phase-based ICSIQLE metric in Fig. 5.5(b). This result shows two key factors of ICSIQLE's ability to predict beamforming performance. First, ICSIQLE tracks the mean of the expected beamformed EVM as evidenced by the increase in mean EVM with an increase in $\vec{h}$ variation. Second, and more importantly, ICSIQLE also tracks EVM variance as shown by the increase in the shaded 2σ bounds with increasing ICSIQLE values.

This ability to track EVM variance is key as it allows for the system designer to utilize FROZEN with ICSIQLE to set relative thresholds on acceptable ICSIQLE values. As the true channel CSI drifts (it becomes more and more outdated), our ability to bound the beamformed EVM utilizing that stale CSI also decreases. Without the luxury of re-sounding CSI, we are unable to guarantee that the eventual beamformed packet will meet a given threshold EVM necessary for accurate decoding in all topologies and circumstances but we can estimate how much the EVM will vary.

For example, if the current 16-QAM-$^1/_2$, MCS selected for client k is optimal for an $[M_4, K_2]$ transmission with fresh CSI, the access point scheduler may know that a degradation of EVM from 5% to 8% (or, equivalently, degradation of SNR from 26 to 22 dB) would correspond to a reduction in MCS to QPSK-$^3/_4$. The controller could declare that rate drop unacceptable and send the next enqueued packet to client k as an $[M_4, K_1]$ transmission, where the ACK would refresh the user's implicit CSI. An alternative strategy would

be to instead reduce the MCS rate to ensure (with 95% confidence) that the transmission is successful, and again refresh the implicit CSI upon the client's ACK.

Regardless, the key is that ICSIQLE provides the necessary information to allow the MAC to make these decisions by linking measured ICSIQLE rates to expected future EVM performance bounds.

5.3.2.3   ICSIQLE: MU-MIMO Channel Variability

To validate ICSIQLE's ability to track CSI variation due to environmental mobility, we construct an $[M_4, K_2]$ topology, use the measurement method outlined in Sec. 5.3.1, and perform a 300 second long experiment. This experiment consisted of three discrete time sections: *(i)* static environment, *(ii)* mobile environment at pedestrian speeds, and finally *(ii)* returning to a static environment.

Our measurement method allows for the collection of instantaneous MU-MIMO CSI and thus calculate instantaneous ICSIQLE values for subsequent beamformed transmissions. The purpose of this experiment is to see if ICSIQLE can accurately track the differences between static and mobile environments to emulate the use of this metric in a real system.

Fig. 5.6 shows the rate of change of ICSIQLE and CN_0 over time. We plot the measured metric values for the two receivers and super-impose the EWMA curve on top of the measured points.

Fig. 5.6(a) shows CN_0's absolute phase based method's ability to track the difference between static and mobile channels. There is no correlation nor discernible difference between the measured CN_0 values in the static and mobile regions. As stated by the 802.11ac task group, such metrics *require* "phase immunity" [36] which such absolute phase based methods do not have.

Unlike CN_0, ICSIQLE's relative angle based metric's ability to track channel variation
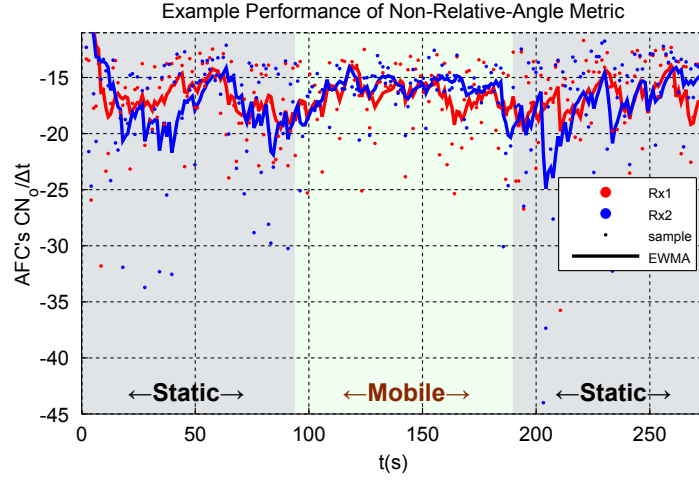
(a) The AFC CN_0 metric's [5] reliance on absolute phase does not allow for the tracking of static and mobile environments.
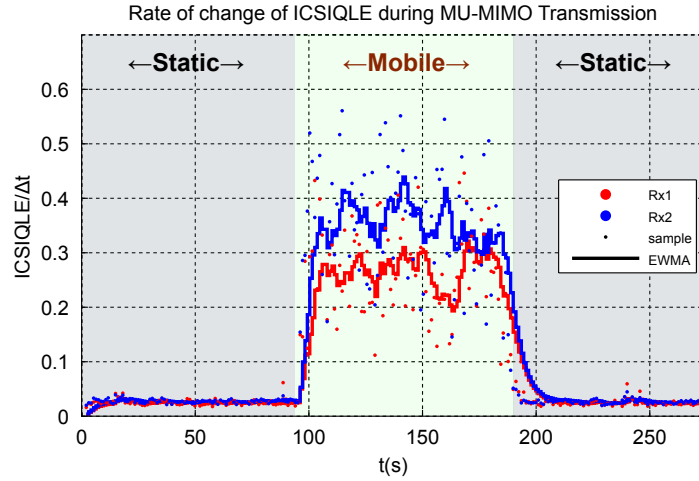
(b) ICSIQLE's relative phase calculation allows for accurate tracking of static and mobile MU-MIMO environments

Figure 5.6: Rate of change of ICSIQLE over time for two clients in an $[M_4, K_2]$ system, plotted during a trial where the environment was static until 100 seconds, pedestrian-speed channel mobility was induced until 180 seconds, and then the channel remained static for the remainder. An exponentially-weighted moving average is super-imposed.

is shown in Fig. 5.6(b). Observe that as soon as the transmission environment switches from static to mobile, the ICSIQLE metric immediately increases for both receivers and remains high for the duration of the mobile period. Once the environment becomes static again, the ICSIQLE value for both receivers immediately drops back down to the level it was in the initial static time period. While the underlying principle of estimating the noise

power between two CSI estimates is sound, it is clear that an accurate system metric must also be robust to measurement phase error in practice.
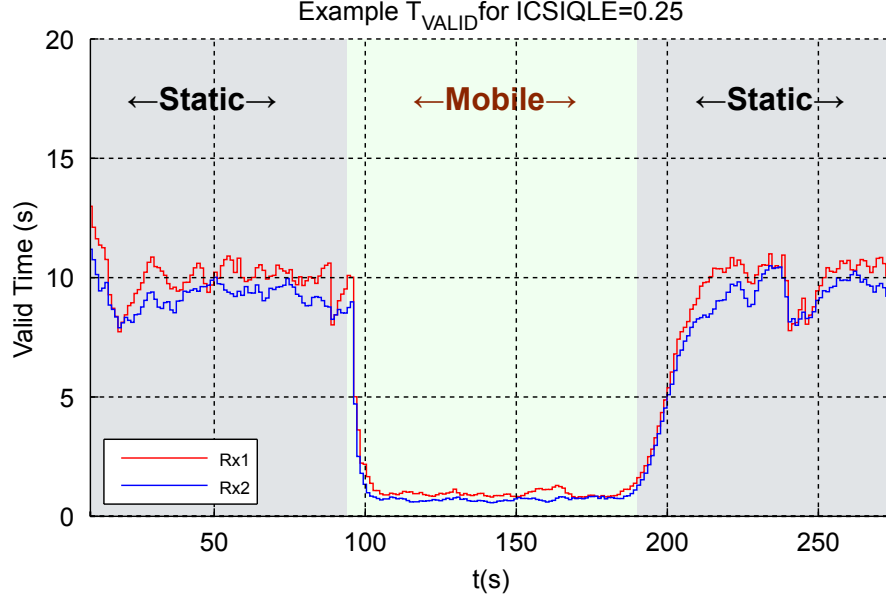
Figure 5.7: A plot of the calculated T_{VALID} for the two users in the channel mobility experiment.

Finally, using the same channel trace from the previous experiment and the method shown in Sec. 5.2.3, we present an example computation of T_{VALID} for an $\mathbf{I}_{th} = 0.25$ in Fig. 5.7. Note how for this chosen $\mathbf{I}_{th}$, a FROZEN enabled AP, the static time period allows for approximately 10 seconds of valid time while the mobile time period allows for less than one second of valid time. This plot emulates how a FROZEN enabled AP would actually use $\text{EWMA}(\mathbf{I})_t$ to make MAC-layer resounding decisions.

5.3.3 Numerical Analysis of FROZEN

Now that we have verified the efficacy of ICSIQLE with regards to channel tracking, we now analyze the potential gains of a FROZEN-enabled AP compared to a standard 802.11ac, explicit beamforming system.

We consider the operating regions of 802.11ac explicit beamforming and FROZEN for

different MCS indexes and observe the performance for different packet aggregation rates.

The results of this analysis are shown in Fig. 5.8.
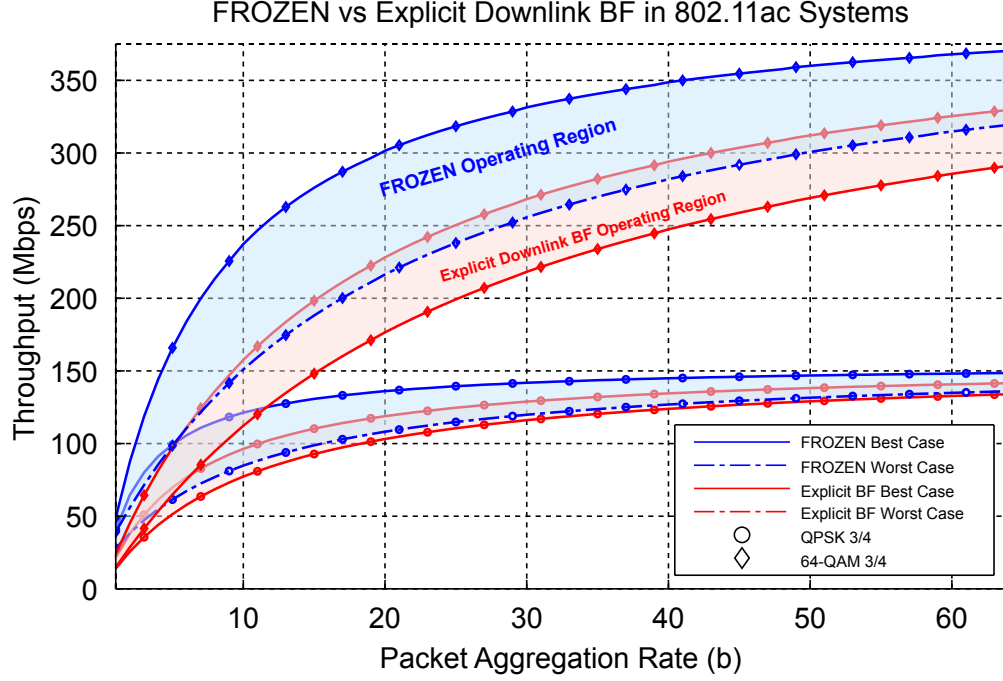
Figure 5.8: Performance of FROZEN with respect to packet aggregation rate and MCS selection, compared to 802.11ac with MU-MIMO overhead for an $[M_4, K_4]$ system.

For two different modulation rates, we show the operating regions for FROZEN and 802.11ac explicit beamforming. These shaded regions are bounded by the best and worst case performance for the two beamforming methods. The differences between the best and worst case performance for these beamforming methods differ due to the method by which they collect CSI and are outlined below:

- **Explicit 802.11ac:** For 802.11ac explicit feedback, as outlined in Sec. 4.2.2, the controlling factors for the amount of CSI feedback are the number of bits per subcarrier and the number of subcarriers fed back (due to subcarrier grouping). While a lower grouping factor and larger number of bits per subcarrier would result in a more accurate measurement (and thus better performance), that effect is not considered here; only the amount of feedback overhead.

 – *Best Case:* Subcarrier grouping = 4, Bits per subcarrier = 14.

 – *Worst Case:* Subcarrier grouping = 1, Bits per subcarrier = 16.

- **FROZEN:** While FROZEN does not utilize the explicit sounding mechanism (*i.e.,* the `CBFR` from Sec. 4.2.2), it may execute non MU-MIMO data transmissions (to force an acknowledgment to collect an implicit CSI estimate). These packets still contain data but as they are not parallel transmissions, they may result in lower overall throughput than an MU-MIMO system that does not contain an explicit sounding sequence.

  – *Best Case:* CSI is always up-to-date from to the previous transmission's acknowledgment. Effectively, an 802.11ac MU-MIMO transmission without an `NDP`, `NDPA`, or `CBFR`.

  – *Worst Case:* Every single MU-MIMO transmission requires an indirect resounding of the channel. Thus, every $[M_4, K_4]$ transmission is interleaved with four single-user packets to force the CSI update. For a conservative evaluation of FROZEN, we assume that the single-user transmissions are all at the same MCS as the multi-user transmissions (although in reality they would be higher; see Sec. 4.1.2).

Now observe how the operating region differences between FROZEN and 802.11ac in Fig. 5.8 are heavily dependent on the chosen MCS and the amount of transmitted data (packet aggregation rate). These factors determine how much time and how efficiently a particular MU-MIMO transmission spends transmitting data as opposed to overhead. At high packet aggregation rates and low MCS (QPSK-$3/4$ is shown), the benefit of FROZEN appears minimal. However, modulation rate does not have the largest effect on FROZEN's performance increase. Instead, FROZEN's benefit is immediately clear at lower packet aggregation rates. Note in Fig. 5.8 how both modulation rates (QPSK-$3/4$ and 64-QAM-$3/4$)

show larger relative increases in throughput when using FROZEN in the low packet aggregation regime.

To explore this further, we will look at these two modulation rates among others and specifically observe the relative throughput increase of FROZEN with respect to packet aggregation rate in Fig. 5.9.
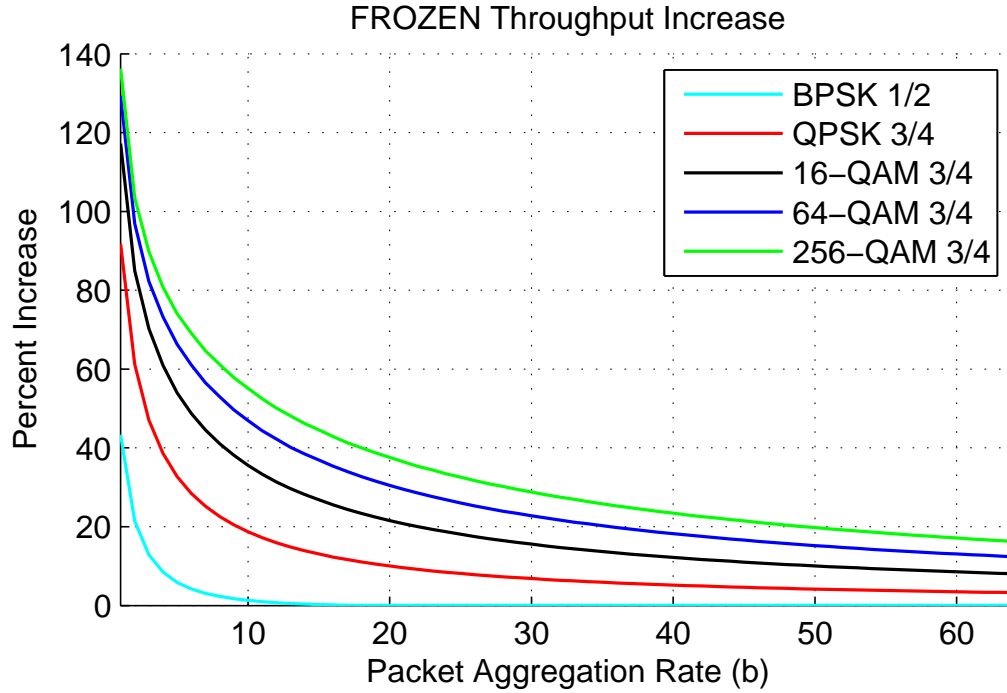
Figure 5.9: Percent gain of FROZEN with respect to 802.11ac with MU-MIMO overhead for an $[M_4, K_4]$ system.

For the given modulation rates, the percent increase of the mean expected throughput (the curve bisecting each transmission type's operating region in Fig. 5.8) is shown from 802.11ac to FROZEN.

The lowest modulation rate, BPSK-$1/2$, shows the least benefit of using FROZEN because the transmission time of the data is far greater than that of the overhead required even by 802.11ac. At approximately 10 aggregated packets per transmission, the benefit is almost zero; however, in cases of low packet aggregation (when a user's transmit queue is minimally backlogged), the use of FROZEN allows an AP to achieve a 10 to 40% increase

in throughput.

However, at higher modulation rates where the transmission time of the data is significantly less (although more efficient), the overhead from the 802.11ac sounding mechanism begins to overwhelm the gains of MU-MIMO. In these cases, FROZEN provides a much higher throughput increase. In fact, in the lower aggregation rates where normally an MU-MIMO transmission would not be feasible due to the overhead required, FROZEN truly shows its potential. These minimally backlogged cases show anywhere from 60 to 80% throughput increases for FROZEN due to its opportunistic removal of overhead.

Minimally backlogged flows are common in low data-rate applications such as Voice over IP (VoIP) or simple internet browsing. FROZEN allows for MU-MIMO transmissions to be used in these cases in addition to fully backlogged scenarios. Additionally, applications such as VoIP require low latency and FROZEN allows an AP to quickly serve those users without switching to SISO or waiting for the queue to fill to the point where the MU-MIMO transmission is feasible in terms of overhead.

The explicit beamforming mechanism of 802.11ac incurs significant overhead to the point where MU-MIMO transmissions are only efficient at very high aggregation rates and low modulation rates. FROZEN bridges the gap to the large remainder of transmission scenarios and allows MU-MIMO transmissions to function efficiently in all cases.

RELATED WORK

6.1 SDR Platforms

A number of common development platforms are capable of some degree of frequency-agility and programmability, e.g., [40, 41, 42]. However, these platforms are generally limited to either narrow-bandwidth applications when used for real-time applications or lack the open hardware and software stack required for research. None of them contain high-power amplifiers for long-range experiments. The form-factor currently required for real-time operation of platforms performing DSP operations on a CPU [40, 41] becomes a limitation when measuring wideband channel statistics for long periods of time with high temporal granularity, as such experiments often require many mobile user nodes.

Furthermore, existing systems do not integrate all components (specifically, a high-power analog front-end or highly dynamic AGC subsystem) necessary for high-bandwidth, long-range experiments. Off-the-shelf UHF amplifiers are not designed for frequency-flat, wideband operation between 470-698 MHz and their size and external power requirements further hinder the mobility of multiple radio nodes.

WURC is designed to work interchangeably with any digital baseband and only draws power from the expansion card slot available on most FPGA development boards while

integrating the remaining components necessary for a high-powered wideband transceiver. In combination with WURC, the Wireless Open-access Research Platform (WARP) digital baseband platform contains a complete real-time layer 2 network stack and large experimental log storage capabilities (2 GB DDR3 RAM) within a small form-factor board, making it feasible to build and deploy a large number of wireless, mobile nodes for UHF MU-MIMO experiments.

6.2 Over-the-Air Characterization Studies

6.2.1 MU-MIMO 2.4/5 GHz Characterization.

While previous work emphasizes the importance of channel coherence time for MU-MIMO systems [10] and theoretical results suggest that center frequency is directly related to channel coherence time [21], these works do not provide the information necessary to perform a comparison based on center frequency. Such an investigation is necessary as MU-MIMO theoretical models for UHF and 2.4/5 GHz WiFi bands are parametrized for different environments (outdoor and indoor respectively). Models suggest that UHF band MU-MIMO exhibits increased temporal correlation at the cost of *increased* spatial correlation compared to 2.4/5 GHz WiFi (which would be detrimental to MU-MIMO due to the difficulty in providing orthogonal streams to the user [10]). However, we show that this tradeoff is not a result of the frequency band; instead, it is a result of the transmission environment. Thus, this discrepancy is not an inherent flaw to existing MU-MIMO channel models; rather, it is a result of incomplete parametrization and comparison of the MU-MIMO channel for all band/environment combinations.

6.2.2   SISO UHF Characterization

Several works explore the propagation characteristics of UHF transmissions in a variety of environments and topologies, e.g., [43, 44, 45, 46]. These works exhaustively analyze the performance of packetized UHF transmissions through different materials and in various environments. However, they focus on single-antenna, single-user transmissions and thus the characterization is restricted to metrics such as path loss, delay profile, and attenuation through materials. In contrast, our work focuses on the aggregate effects of these metrics with respect to MU-MIMO transmissions, namely temporal and spatial correlation in outdoor and indoor environments. Additionally, our work focuses on comparing these characteristics to 2.4/5 GHz bands where MU-MIMO techniques are used prevalently.

6.2.3   MIMO UHF Characterization

Other works exhaustively characterize MIMO transmissions in the UHF band [47, 48, 49, 50, 51]. However, they focus on outdoor, Single-user MIMO transmissions and thus focus on point to point transmissions with a single transmitter/receiver pair, each equipped with multiple antennas. While single-user and multi-user MIMO transmissions can have an equivalent number of transmit/receive antenna paths, the co-located receive antennas in the single-user case drastically reduces the variability in the temporal and spatial correlation with respect to environmental factors. Thus, the usage scenario of distinct MU-MIMO user nodes separated by some distance is not represented in the existing work. Instead, our work focuses on *multi-user* MIMO transmissions and specifically characterizes the effects of separated receivers.

Lastly, uplink MU-MIMO channels were studied in the UHF-band in a rural outdoor environment [52]. In contrast, we focus on downlink transmissions, consider both indoor and urban outdoor environments and provide channel characterization and spatial correlation of groups of users. Additionally, we evaluate *both* UHF and 2.4/5 GHz band MU-MIMO

performance to comparatively characterize the performance of a UHF-band MU-MIMO system and provide an open-source platform.

6.3 MU-MIMO Protocols for Overhead Amortization

6.3.1 Frame Aggregation

Numerous works consider the effects of frame aggregation in MU-MIMO systems. For example, [53] develops a frame aggregation technique and [54] examines the effects of frame aggregation specifically with 802.11ac.

While PUMA selects the most efficient mode based on the number of packets in each receiver's queue, frame aggregation techniques are complementary to PUMA since these techniques can weight or rule out the initial set of potential users. Implementing a frame aggregation technique with PUMA simply requires the modification of step 1 outlined in Sec. 4.1.4. The selection of K_R from K_{max} can be based on a rule more sophisticated than user packet availability. Additionally, once the list is generated, the values of b_i can be weighted accordingly to implement a frame aggregation protocol.

However, such frame aggregation techniques, are not necessary with a FROZEN enabled AP. The purpose of frame aggregation is to amortize the MU-MIMO sounding overhead at the cost of increased latency when serving a particular user. When transmitting in the appropriate operating regime, FROZEN opportunistically removes sounding overhead and allows an AP to transmit any length frame with any number of aggregated packets without penalty.

6.3.2 User Grouping and Selection

Several works focus on user grouping and selection based on channel state and/or unequal transmission length [55, 56, 57, 58, 35]. These works use theoretical expressions similar

to Eq. (4.1) to estimate the aggregate capacities of potential user sets. The information required to employ these modeling techniques is the channel state.

Environments that exhibit high channel variability preclude the ability to use previously collected CSI and continually measuring the channel for all possible users would significantly diminish the performance of MU-MIMO transmissions due to overhead. Thus, in this operating regimes, PUMA allows these MAC-layer decisions to occur before sounding. While the aforementioned protocols are potentially more accurate than PUMA due to the additional information they require, procuring this information given protocol overhead render them impractical to deploy. Also, due to the volatile nature of indoor, 2.4/5.8 GHz WLANs, even once that information is collected, it has a high probability of being outdated further reducing its accuracy. PUMA balances the tradeoff between obtaining the necessary information and the time taken to obtain that information, and thus is more plausible for a true 802.11ac WLAN deployment.

On the other hand, these types of Eq. (4.1)-based protocols *can* be beneficial in the low channel variability operating regime and their techniques are complementary to FROZEN. However, they all lack a channel tracking metric to actually enable their functionality. FROZEN's ICSIQLE metric would allow these user grouping protocols to determine of previously collected CSI is valid and thus make them feasible to implement.

The authors of [5] propose a method to reduce the rate of channel sounding by attempting to track the variation in the MU-MIMO channel. In the highly variable operating regime, this method provides limited utility as the result of their tracking would result in per-packet re-sounding. For these regimes, PUMA is required to make user selection decisions before sounding occurs. In regimes that exhibit low channel variability, [5] could be beneficial; however, as extensively shown in Chapter 5, their channel tracking method, CN_0, attempts to track absolute phase instead of relative phase and thus does not provide any tracking ability whatsoever. FROZEN's ICSIQLE metric, however, allows for accurate

tracking of channel variation and thus allows a FROZEN enabled AP to accurately deter-mine when the channel must be re-sounded. Additionally, through FROZEN's passive-implicit sounding method, the sounding overhead is not just reduced but completely re-moved.

6.3.3 Mode Comparisons

Survey works exist that compare different MIMO modes such as MU-MIMO and frame aggregation [59] or MU-MIMO and multiple SU-MIMO [60]. Both works highlight the tradeoffs between these schemes but do not provide algorithms for exploiting those trade-offs such as PUMA or FROZEN. Additionally, neither work verifies these differences with measured over-the-air transmissions.

CONCLUSION

In this thesis, we show that the development of separate MU-MIMO protocols for different bands and environments is necessary to ensure efficient MU-MIMO performance. Existing MU-MIMO protocols and the existing 802.11ac standard are all-purpose solutions for MU-MIMO transmissions that must function in the pathological case of MU-MIMO (high channel variability and difficult user separability) as shown in Fig. 7.1.
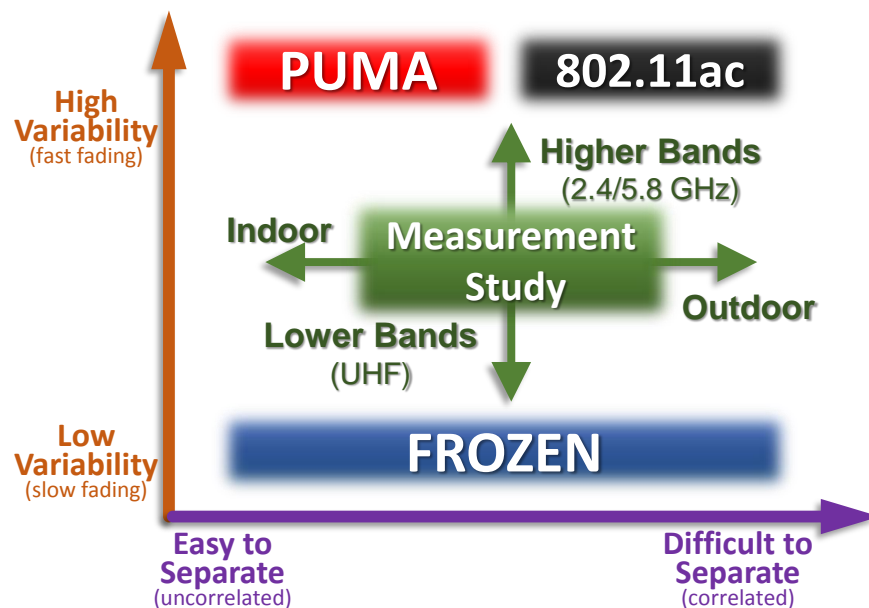

Figure 7.1: Thesis Contributions.

While these all-purpose protocols can certainly work for all operating regimes, their efficiency is significantly diminished compared to protocols that leverage the characteristics of different operating band and environment combinations.

In this work, we first identify the band and environment factors that contribute to the MU-MIMO performance determining characteristics (channel variability and user separability) through an extensive measurement study. Then, we develop **P**re-sounding **U**ser and **M**ode selection **A**lgorithm (PUMA) and **F**eedback **R**emoval with **O**pportunistic **Z**ero-overhead channel **E**stimatio**N** (FROZEN), protocols whose development was driven by our measurement study and that leverage the appropriate characteristics of each MU-MIMO operating region.

This method of MU-MIMO protocol design, which does not target the pathological MU-MIMO operating region, allows for transmission techniques that increase MU-MIMO efficiency and allows this transmission technique to reach its full potential.

REFERENCES

[1] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, Mar. 2006.

[2] O. Bejarano, M. Park, and E. Knightly, "IEEE 802.11ac: From channelization to multi-user MIMO," *IEEE Communications Magazine*, vol. 51, no. 10, Oct. 2013.

[3] IEEE Std 802.11ac-2013, "Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz," 2013. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11ac-2013.pdf

[4] O. Bejarano, E. Magistretti, O. Gurewitz, and E. W. Knightly, "MUTE: Sounding Inhibition for MU-MIMO WLANs," *Proc. IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2014.

[5] X. Xie, X. Zhang, and K. Sundaresan, "Adaptive feedback compression for MIMO networks," in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2013, pp. 477–488.

[6] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, May 1983.

[7] C. Zhong, M. McKay, T. Ratnarajah, and K. Wong, "Distribution of the Demmel condition number of Wishart matrices," *IEEE Transactions on Communications*, vol. 59, no. 5, May 2011.

[8] F. Kaltenberger, D. Gesbert, R. Knopp, and M. Kountouris, "Correlation and capacity of measured multi-user MIMO channels." in *Proc. IEEE Personal, Indoor, and Mobile Radio Communications Conference (PIMRC)*, Cannes, France, Sept. 2008.

[9] A. Greenbaum and T. Chartier, *Numerical Methods: Design, analysis, and computer implementation of algorithms.* Princeton University Press, 2012.

[10] E. Aryafar, N. Anand, T. Salonidis, and E. Knightly, "Design and experimental evaluation of multi-user beamforming in Wireless LANs," in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, Chicago, IL, Sept. 2010.

[11] Wurc documentation. [Online]. Available: http://www.vuum.com/wurc/

[12] A. Grebennikov, *RF and Microwave Power Amplifier Design.* McGraw-Hill, 2004.

[13] B. Yarman, *Design of ultra wideband power transfer networks.* John Wiley & Sons, 2010.

[14] A. Flores, R. Guerra, E. Knightly, P. Ecclesine, and S. Pandey, "IEEE 802.11af: A standard for TV white space spectrum sharing," *IEEE Communications Magazine*, vol. 51, no. 10, Oct. 2013.

[15] Lime Microsystems LMS6002D. [Online]. Available: http://www.limemicro.com/products/LMS6002D.php

[16] Rice University WARP project. [Online]. Available: http://warp.rice.edu

[17] Mango Communications 802.11 Reference Design. [Online]. Available: http://mangocomm.com/802.11

[18] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Communications Magazine*, vol. 19, no. 6, Dec. 2012.

[19] M. Zhu, G. Eriksson, and F. Tufvesson, "The COST 2100 channel model: Parametrization and validation based on outdoor MIMO measurements at 300 MHz," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, Feb. 2013.

[20] J. Poutanen, K. Haneda, L. Lingfeng, C. Oestges, F. Tufvesson, and P. Vainikainen, "Parameterization of the COST 2100 MIMO channel model in indoor scenarios," in *Proc. EurAAP European Conference on Antennas and Propagation (EuCAP)*, Rome, Italy, Apr. 2011.

[21] T. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 2001.

[22] J. Wallace, M. Jensen, L. Swindlehurst, and B. Jeffs, "Experimental characterization of the MIMO wireless channel: Data acquisition and analysis." *IEEE Transactions on Wireless Communications*, vol. 2, no. 2, Mar. 2003.

[23] G. Breit, "Coherence Time Measurement for TGac Channel Model," in *IEEE 802.11-09/1173r1*, Nov. 2009.

[24] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[25] F. Adib, S. Kumar, O. Aryan, S. Gollakota, and D. Katabi, "Interference alignment by motion," in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, Miami, Florida, Sept. 2013.

[26] F. Kaltenberger, M. Kountouris, D. Gesbert, and R. Knopp, "On the trade-off between feedback and capacity in measured MU-MIMO channels," *IEEE Transactions on Communications*, vol. 8, no. 9, Sept. 2009.

[27] J. Camp and E. Knightly, "Modulation rate adaptation in urban and vehicular environments: Cross-layer implementation and experimental evaluation," in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, San Fransisco, California, Sept. 2008.

[28] M. Sharif and B. Hassibi, "A comparison of time-sharing, dpc, and beamforming for mimo broadcast channels with many users," *IEEE Transactions on Communications*, vol. 55, no. 1, Jan. 2007.

[29] M. Lopez, "Multiplexing, scheduling, and multicasting strategies for antenna arrays in wireless networks," Ph.D. dissertation, M.I.T., Aug. 2002.

[30] A. Edelmen, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, M.I.T., May 1989.

[31] H. Yang and T. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE Journal on Selected Areas in Communications*, 2013.

[32] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, Istanbul, Turkey, Aug. 2012.

[33] N. Anand, S.-J. Lee, and E. Knightly, "STROBE: Actively securing wireless communications using zero-forcing beamforming," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, Florida, Mar. 2012.

[34] H. Kim, K. Kim, Y. Han, and S. Yun, "A proportional fair scheduling for multicarrier transmission systems," in *Proc. IEEE Vehicular Technology Conference (VTC)*, Sept. 2004.

[35] X. Xie and X. Zhang, "Scalable user selection for MU-MIMO networks." in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, Toronto, Canada, Apr. 2014.

[36] G. Breit, "Coherence Time Measurement for TGac Channel Model," *Presentation Slides, IEEE 802.11-09/1173r1*, Nov. 2009.

[37] C. K. B. Park, H. Cheon, "A Novel Timing Estimation Method for OFDM Systems," *IEEE Communications Letters*, vol. 7, no. 5, May 2003.

[38] IEEE Std 802.11af-2013, "Amendment 5: Television white spaces (tvws) operation," 2013. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11af-2013.pdf

[39] N. Anand, J. Lee, S.-J. Lee, and E. W. Knightly, "Mode and User Selection for Multi-User MIMO WLANs without CSI," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*. IEEE, 2015.

[40] Ettus USRP. [Online]. Available: https://www.ettus.com

[41] Microsoft SORA. [Online]. Available: http://research.microsoft.com/en-us/projects/sora/

[42] Nutaq. [Online]. Available: http://nutaq.com/en

[43] I. Collings, H. Suzuki, and D. Robertson, "Ngara broadband access system for rural and regional areas," *Telecommunications Journal of Australia*, vol. 62, no. 1, 2012.

[44] J. Hampton, N. Merheb, W. Lain, D. Paunil, R. Shuford, J. Abrahamson, and W. Kasch, "Propagation characteristics of ground based urban communications in the military UHF band," in *Proc. IEEE Military Communications Conference (MIL-COM)*, Atlantic City, NJ, Oct. 2005.

[45] V. Pham and J.-Y. Chouinard, "A study on the channel and signal cross correlation of UHF DTV channels," in *Proc. IEEE International Symposium on Signals, Systems and Electronics (ISSSE)*, Montreal, Canada, Jul. 2007.

[46] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chen, and R. Chandra, "Exploring indoor white spaces in metropolises." in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, Miami, FL, Sept. 2013.

[47] J. Boyer, P. Brown, K. Hayler, M. Garcia, J. Mitchell, P. Moss, and M. Thorp, "MIMO for Broadcast–results from a high-power UK trial," in *Proc. International Broadcasting Convention (IBC)*, Amsterdam, The Netherlands, Sept. 2007.

[48] G. Eriksson, S. Linder, K. Wiklundh, P. Holm, P. Johansson, F. Tufvesson, and A. Molisch, "Urban peer-to-peer MIMO channel measurements and analysis at 300 MHz," in *Proc. IEEE Military Communications Conference (MILCOM)*, San Diego, CA, Nov. 2008.

[49] A. Hammons, J. Hampton, N. Merheb, and M. Cruz, "Cooperative MIMO field measurements for military UHF band in low-rise urban environments," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Darmstadt, Germany, Jul. 2008.

[50] M. Jung, J.-H. Kim, Y.-K. Yoon, and H.-J. Hong, "Multipath characteristics of MIMO channel at the UHF band for wireless systems in the urban area," in *Proc. IEEE International Conference on Advanced Communication Technology (ICACT)*, Phoenix Park, South Korea, Feb. 2011.

[51] R. Parviainen, J. Ylitalo, J-P.Nuutinen, P. Talmola, J. Henriksson, H. Himmanen, R. Ekman, and E. Huuhka, "Experimental investigations on MIMO radio channel characteristics on UHF band," in *Proc. IEEE International Conference on Communications (ICC)*, Dresden, Germany, Jun. 2009.

[52] H. Suzuki, D. Robertson, N. Ratnayake, and K. Ziri-Castro, "Prediction and measurement of multiuser MIMO-OFDM channel in rural Australia," in *Proc. IEEE Vehicular Technology Conference (VTC)*, Yokohama, Japan, May 2012.

[53] B. Bellalta and M. Oliver, "A space-time batch-service queueing model for multi-user MIMO communication systems," in *Proc. ACM Conference on Modeling, analysis and simulation of wireless and mobile systems (MSWiM)*, Oct. 2009.

[54] B. Bellalta, J. Barcelo, D. Staehle, A. Vinel, and M. Oliver, "On the Performance of Packet Aggregation in IEEE 802.11ac MU-MIMO WLANs," *IEEE Communications Letters*, Jul. 2012.

[55] M. Esslaoui, F. Riera-Palou, and G. Femenias, "A fair MU-MIMO scheme for IEEE 802.11ac," in *International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2012.

[56] Y. Jang and H. K. Y. Lee, "Adaptive mode selection for multiuser MIMO downlink systems," in *Proc. IEEE Vehicular Technology Conference (VTC)*, Montreal, Canada, Sept. 2006.

[57] Z. Shen, R. Chen, J. Andrews, R. Heath, and B. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 2005.

[58] T. Tandai, H. Mori, and M. Takagi, "Cross-Layer-Optimized User Grouping Strategy in Downlink Multiuser MIMO Systems," in *Proc. IEEE Vehicular Technology Conference (VTC)*, Apr. 2009.

[59] J. Cha, H. Jin, B. C. Jung, and D. K. Sung, "Performance comparison of downlink user multiplexing schemes in IEEE 802.11ac: Multi-user MIMO vs. frame aggregation," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2012.

[60] A. Thapa and S. Shin, "A MAC protocol to select optimal transmission mode in very high throughput WLAN: MU-MIMO vs. multiple SU-MIMO," in *Proc. IEEE Asian Himalayas International Conference on Internet (AH-ICI)*, Nov. 2012.