University of California
Santa Barbara

Millimeter wave picocellular networks: capacity analysis and system design

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy

in

Electrical and Computer Engineering

by

Zhinus Marzi

Committee in charge:

Professor Upamanyu Madhow, Chair
Professor Yasamin Mostofi
Professor Kenneth Rose
Professor Haitao Zheng

March 2019

The Dissertation of Zhinus Marzi is approved.

_____

Professor Yasamin Mostofi

_____

Professor Kenneth Rose

_____

Professor Haitao Zheng

_____

Professor Upamanyu Madhow, Committee Chair

September 2018

Millimeter wave picocellular networks: capacity analysis and system design

Copyright © 2019

by

Zhinus Marzi

To my parents, with love and gratitude

and to

Venus and my beloved little Viana
My heart beats for you!

Acknowledgements

Curriculum Vitæ
Zhinus Marzi

Education

2017	Doctor of Philosophy, Electrical and Computer Engineering, University of California, Santa Barbara, USA.
2013	Master of Science, Communications and Signal Processing, Sharif University of Technology, Tehran, Iran.
2010	Bachelor of Science, Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.

Publications

Zhinus Marzi, Upamanyu Madhow, "Capacity analysis of mm-wave outdoor picocells," *to be submitted shortly.*

Soorya Gopalakrishnan*, **Zhinus Marzi***, U. Madhow and Ramtin Pedarsani, "Toward Robust Neural Networks via Sparsification," *to be submitted shortly.*

(C7) Soorya Gopalakrishnan*, **Zhinus Marzi***, Upamanyu Madhow and Ramtin Pedarsani, "Combating Adversarial Attacks Using Sparse Representations," *6th International Conference on Learning Representations (ICLR)*, Vancouver, BC, Canada, Apr. 2018

(C6) **Zhinus Marzi***, Soorya Gopalakrishnan*, Upamanyu Madhow and Ramtin Pedarsani, "Sparsity-based Defense against Adversarial Attacks on Linear Classifiers," *IEEE International Symposium on Information Theory (ISIT)*, Vail, Colorado, Jun. 2018. (*Joint first authors)

(C5) **Zhinus Marzi**, Upamanyu Madhow and Joao Hespanha, "On the information in spike timing: neural codes derived from polychronous groups," **invited paper**, *2018 Information Theory and Applications Workshop (ITA)*, San Diego, CA, 2018.

(C4) Maryam Eslami Rasekh, **Zhinus Marzi**, Yanzi Zhu, Upamanyu Madhow and Haitao Zheng, "Compressive path tracking for 60GHz picocells," *The 18th International Workshop on Mobile Computing Systems and Applications (ACM HotMobile)*, Sonoma, CA, Feb. 2017.

(J1) **Zhinus Marzi**, Dinesh Ramasamy, Upamanyu Madhow, "Compressive channel estimation and tracking for large arrays in mm wave picocells," *IEEE Journal of Selected Topics in Signal Processing (JSTSP)*, vol. 10, no. 3, pp. 514-527, April, 2016.

(C3) **Zhinus Marzi**, Upamanyu Madhow and Haitao Zheng "Interference analysis of 60GHz Outdoor Picocells," *IEEE Global Communications Conference,* San Diego, CA, Dec. 2015.

(C2) Yibo Zhu, Zengbin Zhang, **Zhinus Marzi**, Chris Nelson, Upamanyu Madhow, Ben Y. Zhao and Haitao Zheng, "Demystifying 60GHz Outdoor Picocells," *proceedings of the 2014 ACM SIGMOBILE International Conference on Mobile Computing and Networking (MobiCom)*, Sept. 2014, Hawaii, USA.

(C1) Reza Taherkhani, **Zhinus Marzi**, Saeed Karimi, Alireza M. Shahri, "Local Positioning System Using Simple Ultrasonic Sensors," *Proceedings of the 10thInternational Workshop on Research and Education in Mechatronics (REM2009)*, Glasgow, UK.

Abstract

Millimeter wave picocellular networks: capacity analysis and system design

by

Zhinus Marzi

The explosive growth in demand for wireless mobile data, driven by the proliferation of ever more sophisticated handhelds creating and consuming rich multimedia, calls for orders of magnitude increase in the capacity of cellular data networks. Millimeter wave communication from picocellular base stations to mobile devices is a particularly promising approach for meeting this challenge because of two reasons. First, there is a large amount of available spectrum, enabling channel bandwidths of the order of Gigahertz (GHz) which are 1-2 orders of magnitude higher than those in existing WiFi and cellular systems at lower carrier frequencies. Second, the small carrier wavelength enables the realization of highly directive steerable arrays with a large number of antenna elements, in compact form factors, thus significantly enhancing spatial reuse. Hence, we propose to employ the 60 GHz unlicensed band for basestation to mobile communication in outdoor picocells.

We first investigate the basic feasibility of such networks, showing that 60GHz links are indeed viable for outdoor applications. For this purpose, we provided link budget calculations along with preliminary simulations which show that despite the common concerns about higher oxygen absorption and sensitivity to movement and blockage, picocloud architecture provides availability rate of more than 99%.

Next, we explore the idea of increasing spatial reuse by shrinking picocells hoping that interference is no longer the bottleneck given the highly directive antenna arrays at this band. Our goal is to estimate the achievable capacity for small picocells along an

urban canyon. We consider basestations with multiple *faces* or sectors, each with one or more antenna arrays. Each such array, termed *subarray* can employ Radio Frequency (RF) beamforming to communicate with *one* mobile user at a time. We first focus on characterization and modeling the inter-cell interference for one subarray on each face. Our analysis provides a strong indication of very large capacity (in the order of Tbps/km) with a few GHz of bandwidth.

Following this, we explore the impact of adding multiple subarrays per face. This leads to intra-cell interference as well as additional inter-cell interference. While the effect of additional inter-cell interference can be quantified within our previous framework, intra-cell interference has inherently different features that call for new approaches for analysis and design. We propose a cross-layer approach to suppress the intra-cell interference in two stages: (a) Physical layer (PHY-layer) method which mitigates interference by joint precoding and power adaptation and (b) Medium Access Control layer (MAC-layer) method which manages the residual interference by optimizing resource allocation. We then estimate the capacity gain over conventional LTE cellular networks and establish that 1000-fold capacity increase is indeed feasible via mm-wave picocellular networks.

Lastly, we examine fundamental signal processing challenges associated with channel estimation and tracking for large arrays, placed within the context of system design for a mm-wave picocellular network. Maintainance of highly directive links in the face of blockage and mobility requires accurate estimation of the spatial channels between basestation and mobile users. Here we develop the analytical framework for *compressive* channel estimation and tracking. We also address the system level design discussing link budget, overhead, and inter-cell beacon interference. Simulation results demonstrate that our compressive scheme is able to resolve mm-wave spatial channels with a relatively small number of compressive measurements.

Contents

List of Figures

xiii

List of Tables

Chapter 1

Introduction

The demand for cellular data is projected to increase 1000-fold by 2020 [2], driven to a large extent by mobile video. According to Shannon's capacity limit equation, there exist three dimensions to boost up the cellular networks overall capacity:

- increasing data rate per user by allocating more channels to them (time/frequency)

- increasing data rate per channel use (spectral efficiency)

- increasing the number of coexisting users (same time, frequency and space)

Existing cellular systems below 5 GHz are fundamentally constrained by available bandwidth [3], and are approaching their limits of spectral efficiency [4]. Moreover, deploying smaller and denser cells (in order to serve a greater number of subscribers per region) has already been explored in current cellular networks [5, 6]. However, this approach is *fundamentally limited* by interference constraints for the carrier frequencies employed in today's cellular systems [7]. This is because of the hard limit on the size of the antenna employed at basestation or the mobile device. Therefore, even advances in MIMO technology can only avoid interference to a limited degree at low carrier frequencies which is not enough for realizing significant capacity gain [7].

The promise of mm-wave.   However, higher carrier frequencies, offer an attractive alternative. We take 60 GHz band as an example which has 7 GHz of unlicensed spectrum. This can already enable channel bandwidths of the order of GHz, 1-2 orders of magnitude higher than those in existing systems at lower carrier frequencies (more channels). Moreover, shorter carrier wavelength allows for larger number of antenna elements to fit within compact form factors [1]. The highly directional beams of which will drastically reduce interference (higher spectral efficiency) and enable highly overlapping links with minimal mutual interference (greater number of users).

After all, there is no doubt about the potential for the mm-wave band to provide significant capacity gain. Even though most of the attention so far has focused on using 60 GHz for *indoor* applications, the goal of this work is to show that it is a strong candidate for delivering the 1000-fold capacity increase demanded for *outdoor* cellular networks in dense urban environments.

We propose using picocells with base stations opportunistically deployed on lampposts or rooftops 1.1. As we will show in section 1.1, the radius of 100 meters or less is a sweet spot for mm-wave picocells where the impact on link budget due to oxygen absorption (if using the 60 GHz unlicensed band) and rain (if using any mm-wave band) is relatively small. The additional losses due to oxygen absorption and heavy rain are about 16 and 20 dB/km respectively. Hence, the total loss over 100m is 3.6 dB which can be handled within our desired link budget.

Given the highly directive nature of the beam and the limited diffraction at short wavelengths, mm-wave links are more prone to blockage and limiting range also reduces the likelihood of blockage.

Given the high demand for wireless data in dense urban environments, we focus our

---

[1]A 100-element array at 60GHz band where carrier wavelength is only 5mm easily fits in no more than a square inch
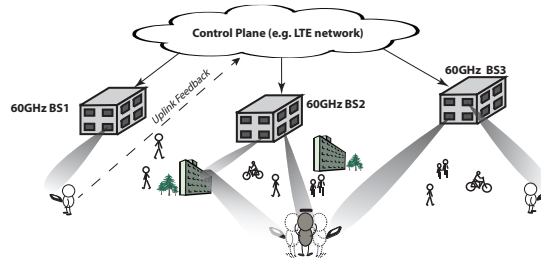
Figure 1.1: A potential 60GHz picocell architecture for outdoor mobile broadband. Base stations are densely deployed, each of which contains a few antenna arrays, forming highly directional beams simultaneously to multiple users. The beams follow users through direct or reflected paths, creating high-bandwidth downlink transmissions [8].

modeling and performance evaluation on the urban canyon setting depicted in Figure 1.2, with streets flanked by buildings on both sides. Picocellular basestations are deployed on lampposts (at lamppost height of $6m$) in a zig-zag pattern on both sides of the street.

We want to use pencil beams but provide omnidirectional coverage. To accomplish this, we propose the use of base stations with multiple "faces" where each face can have multiple antenna arrays. A nominal field of view for each face might be 120 degrees. Thus, two faces suffice for an urban canyon, where the overall field of view required is 180 degrees. For the east-west street shown, each base station has two faces, facing east and west, respectively. Thus, a user in the street segment shown in Figure 1.2 is covered by the eastward-facing face of BS2 and the westward-facing face of BS1. Denoting the east-west distance between adjacent basestations as d, the street segment shown can be termed a picocell of width d. Thus, each picocell is covered by one face each from two base stations. Mobile stations are modeled as either pedestrian walking along sidewalks, or cars moving along the street.

Multiple antenna systems at lower carrier frequencies have a relatively small number of elements, each with its own radio frequency (RF) chain. This provides control of the individual baseband signals associated with each element, enabling sophisticated adaptation, including frequency-selective spatiotemporal processing (e.g., per subcarrier

3
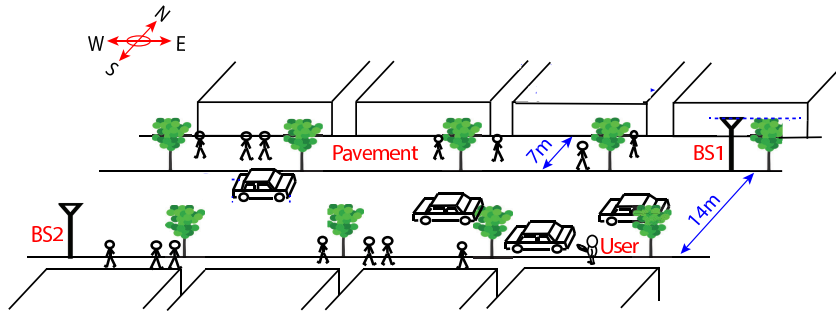
Figure 1.2: Simulation environment: a typical street in New York City (observed from the Google Map), where two basestations serve users on both sidewalks along the street [8].

beamforming in OFDM systems). This approach does not scale when we have a large number of antenna elements packed into a tiny form factor. Instead, we consider RF beamforming, in which a common baseband signal is routed to/from the antenna elements, and we can only control the amplitude and phase for each element. In such an architecture, each antenna array can only communicate with a single user at a time. Thus, resource allocation for a single antenna array is using time division multiplexing, but multiple users can be supported by deploying multiple antenna arrays per face (we call each such array a subarray).

Mm-wave bands have also been studied toward developing the next generation (5G) of wireless networks. 5G will use spectrum in the existing LTE frequency range (600 MHz to 6GHz) as well as mm-wave bands (24 - 86 GHz). One ambitious goal set for 5th generation wireless networks is the 1000x capacity increase. However, a critical bottleneck is the limited available spectrum, especially at costly and crowded sub-6 GHz bands. Hence, exploiting shared and especially unlicensed spectrum (including mm-wave bands) is essential for meeting 5G goals [9]. Early mm-wave prototypes operate at 28 GHz to avoid oxygen absorption and deliver extreme data rates over a line-of-sight range of 350 meters and non-line-of-sight coverage of 150 meters in a dense urban deployment [10]. However, in this work we study the unlicensed 60GHz band since (a) it has already

been standardized for indoor use, which can be leveraged for seamless indoor-outdoor coverage (b) the amount of spectrum is massive and recently increased from 7GHz to 14GHz in the United States [11] and (c) the use of unlicensed spectrum potentially enables new entrants into a market currently dominated by deep-pocketed carriers. We, therefore, focus on 60GHz in our numerical results, although our general framework for analysis and design applies to millimeter wave bands in general.

1.1   Basic Feasibility of mm wave Outdoor Picocells

Now, we review some of the results from our previous study which address the common concerns about 60GHz via extensive measurements and simulations, showing that the potential for 60GHz outdoor picocells can be realized without fighting physics [8]

1.1.1   Communication range

Link budget calculations leveraging IEEE 802.11ad: The IEEE 802.11ad standard for 60GHz [12] outlines a number of modulation and coding schemes to adapt to different range-rate tradeoffs. We have leveraged 802.11ad physical layer for the purpose of our link budget calculations by which we show that link ranges of the order of 100m are attainable (see Table 1.1).

The receiver sensitivity for each data rate (*i.e.,* the minimum received power required to sustain that data rate) is obtained from the 802.11ad standard and already accounts for large values of implementation loss (5 dB) and noise figure (10 dB). Our link budget also accounts for the FCC EIRP restriction of 40 dBm and loss due to free space propagation and oxygen absorption. On top of that, we add a link margin of 15 dB, which is accounting for reflection loss (in case we are using a bounce rather than a LoS path, which typically leads to at most 5-7 dB loss) and for rain. That is, the receiver sensitivity as a function

of range R (in meters) is computed as

$$P_{RX}(dBm) = EIRP(dBm) + G_{RX}(dBi) - L_{ploss}(R) - L_{margin} \qquad (1.1)$$

where $G_{RX}(dBi) = 10\log_{10} N_{RX}$ is the receiver antenna directivity as a function of the number of receive elements N_{RX} (we consider $N_{RX} = 16, 64, 100$ for the 4×4, 8×8 and 10×10 arrays). The path loss L_{ploss} in dB is

$$L_{ploss}(R) = 10\log_{10} \frac{16\pi^2 R^2}{\lambda^2} + \alpha R \qquad (1.2)$$

with $\alpha = 0.016$ dB/m (16 dB/km) accounting for oxygen absorption, and $\lambda = 0.005$m at 60 GHz.

Table 1.1: Outdoor range/rate tradeoffs with 802.11ad PHY

Data rate	RX sensitivity	4×4 RX array	8×8 RX array	10×10 RX array
385 Mbps	-68 dBm	63m	114m	138m
1.155 Gbps	-64 dBm	41m	77m	94m
2.31 Gbps	-61 dBm	30m	57m	70m

As shown in Table 1.1, within FCC's power regulations, and taking into account oxygen absorption and heavy rain, 60GHz offers coverage area that exceeds 130m for 385Mbps and 70m for 2.31Gbps.

1.1.2   Sensitivity to blockage and user motion

Given the highly directive nature of the beam and the limited diffraction at small wavelengths [13], 60 GHz links can indeed be blocked easily by human body or other obstacles. However, this issue can be solved by an electronically steerable antenna array that can point its beam towards one of several available reflected paths or by the *picocloud* architecture that can handle the handover of user among basestations seam-

lessly. Next, we conduct simulations to show that making use of reflected paths together with the picocloud architecture, provide robust connectivity even in obstacle-rich urban environments.

Simulation Setup: We consider a street canyon as in Figure 1.2. We considers pedestrians, cars and trees as obstacles that can potentially block or attenuate 60GHz links. We model each pedestrian as a cylinder of radius $0.3m$ with height uniformly distributed between $1.5m$ and $2m$, and trees as cyclinders with radius $0.4m$ and height $3m$. We model cars as rectangular cuboids of size $2m \times 4.5m \times 1.5m$. We consider a moderate pedestrian density of $0.08 person/m^2$ [14]. The link budget in our simulations is as in eq. (1.1)-(1.2).

We also consider regular square arrays of various sizes with standard radiation patterns [15]. However, since we fix the EIRP to 40dBm, the size of the transmit array does not impact the link budget, but can have a significant effect on interference. The receive array size though, impacts both link budget and interference. The 60 GHz channel is well characterized by a small number of dominant rays [16]. Path strengths will decay rapidly with multiple bounces. Hence, in our model, we consider the LoS ray and one-bounce paths.

Link Availability: We consider several factors that may affect link availability:

- User's orientation: Four possible orientations (North, South, East and West) are considered and when facing basestation 1, the user's body blocks the LoS path from basestation 2 and vice versa.

- Blockage by other obstacles *e.g.* other pedestrians, trees, and cars. We consider a moderate pedestrian density of 0.08 per square meter[2]. We find that, in general,

---

[2]We obtain similar results for much higher densities since it turns out that blockage of user body is the key source of link impairment.

Data rate (Mbps)	385	1155	2310
LoS (BS1)	66.0%	66.0%	66.0%
LoS + 2 NLoS (BS1)	94.2%	93.8%	93.5%
LoS (BS1& BS2)	94.5%	94.5%	94.5%
6 paths (BS1&BS2)	100%	99.95%	99.95%

Table 1.2: Effect of picocloud and reflection paths on availability, 10×10 RX array, 100m separation between the two basestations.

cars are not tall enough to block LoS and wall reflections to users on sidewalks, but may disrupt ground reflections. Hence our simulations do not utilize any ground reflections from the street or sidewalk.

- The presence of gaps between buildings (see Figure 1.2) which might eliminate the reflected paths from walls.

Taking into account these factors, we run simulations to estimate the distribution of RSS at different receiver locations, accounting for different receiver orientations and locations of other pedestrians and cars. We use this to compute the *availability rate* for three 802.11ad data rates listed in Table 1.1 (a data rate is "available" if the RSS exceeds its required receiver sensitivity).

Table 1.2 summarizes the availability rates for four different scenarios in terms of the capability of switching among available paths from a single or multiple basestations. As shown in Table 1.2, using LoS path from a single basestation cannot provide robust connectivity, but when both basestations are involved, the link availability jumps from 66% to 94.5% at 2310 Mbps. Adding just two wall-reflected paths further increases availability to 99.95-100%.

Basestation Spacing and RX Array Size. Assuming a picocloud of two basestations and the use of both LoS and reflected paths, Table 1.3 reports the availability rate for different basestation spacing and RX array size. We see that for the 10×10 RX array and a BS spacing of 200m, we have high availability even at 2 Gbps (99.3% in clear,

Table 1.3: Availability rates (blockage scenario).

RX array	BS spacing	Clear			Rainy		
		Data rate (Mbps)					
		385	1155	2310	385	1155	2310
4×4	100m	99.9%	99.5%	99%	99.8%	99.4%	98%
	200m	99.4%	94.6%	85%	98.3%	89.2%	75.7%
8×8	100m	99.9%	99.75%	99.6%	100%	99.75%	99.7%
	200m	99.5%	99.45%	98.9%	99.5%	99%	97.2%
10×10	100m	100%	99.95%	99.95%	99.9%	99.9%	99.8%
	200m	99.5%	99.5%	99.3%	99.5%	99.5%	98.8%
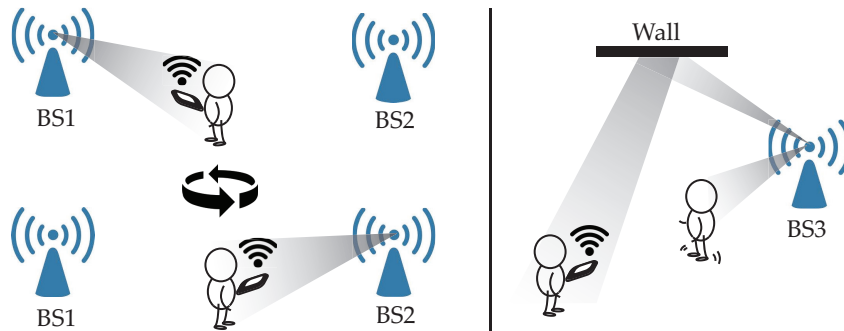
Figure 1.3: Available path changes as a user rotates her body or gets blocked by another pedestrian [1]

and 98.8% in rainy conditions). This implies that even with a sparse deployment, the network offers robust connectivity and coverage.

User Motion: maintaining high rate connections toward *moving* users also require switching among multiple paths from a single base station as well as handing over to another basestation (Figure 1.3). However, most of the time we only need to realign beams once every few seconds to keep up with outdoor pedestrian motion which is practical since the time constant of electronic adaptation is small compared to the time constants of motion. Later on in Chapter 4 we will introduce a fast and efficient algorithm that make it possible to estimate and track the available paths to mobile users.

1.2   Massive Capacity Gain of mm wave Outdoor Pic-
ocells

The small wavelengths in mm-wave bands enable the realization of highly directive antennas in compact form factors, thus producing significant gains in spatial reuse. In Chapter 2 and 3 of this thesis, we conduct an interference analysis that accounts for the unique characteristics of mm-wave communication, and use this to provide rough esti-mates of the attainable network capacity. Our analysis are tailored for outdoor picocells along an urban canyon. More specifically, we consider picocellular base stations deployed on lampposts on each side of the street along an urban canyon (e.g. a typical street in New York City), as depicted in Figure 1.2. Each BS has two "faces," facing east and west. Each face may have multiple antenna arrays, each with a large number of elements.

In our prior work [8], we have obtained promising results regarding the basic fea-sibility and massive spatial reuse achievable in the proposed architecture. While there are significant design challenges in realizing the proposed architecture (e.g.,see the dis-cussion in [8]), our focus in this work (Cahpter 2 and 3) is to *quantify the achievable network capacity* as we shrink 60 GHz picocells, assuming that these challenges can be surmounted.

1.2.1   Contributions

Prior work at lower carrier frequencies shows that interference becomes a fundamental limiting factor [17] in picocellular settings, but as we show here, the narrow beams synthesized using large arrays at 60 GHz alleviate this problem. Here is a brief overview of our roadmap to estimate the capacity gain of mm-wave picocellular networks:

In **Chapter 2**, we obtain a rough estimate of the overall achievable capacity by

shrinking cells. First, we mainly focus on *inter-cell* interference by considering only one subarray per face of a base station. In our analysis, we account for the geometry of the urban canyon, and obtain the following results.

• We show that the inter-cell interference caused by the main beam directed at the desired user vanishes after a number of cells determined by the heights of the BS and the user.

• We develop analytical expressions for the inter-cell interference due to sidelobes, accounting for oxygen absorption and reflections.

• We simulate the statistics of the signal-to-interference-plus-noise ratio (SINR) and attainable data rates under some simple resource management schemes, estimating the attainable network capacity to be of the order of Terabits/sec per km along the canyon, using 2 GHz of spectrum.

In **Chapter 3**, we have further explored mm-wave picocellular networks potential by letting multiple subarrays per base station. This will evidently cause an excessive amount of interference by (a) inducing *intra-cell interference* originated from the other transmitters on the same base station and (b) raising inter-cell interference caused by additional transmitters in neighboring cells.

Our previous analysis for inter-cell interference could still be applicable to this scenario with minor modifications. Intra-cell interference however, has different features that calls for new approaches to deal with. For example, *joint* precoding, power adaptation and resource allocation schemes that require coordination among interfering transmitters are easier to employ for combating intra-cell interference where transmitters are co-located, as opposed to the inter-cell interference where the interfereing transmitters are installed on different base stations.

We have proposed a *cross-layer* approach to deal with the intra-cell interference in which we have combined techniques from two broad areas that have been studied in

the literature: (a) downlink linear precoding and power control [18, 19, 20, 21, 22] (b) powerful optimization approaches recently developed for network-level resource allocation [23, 24]. Here is a brief description of our two-step method:

1. Given that a resource block is assigned to a pre-defined set of users, we develop a PHY-layer building block which employs an optimal linear method (i.e., LMMSE) for beamforming and power allocation to supress the LoS intra-cell interference among them.

2. We then incorporate the PHY-layer block in designing the MAC-layer protocol which solves an optimization problem to determine the set of active users on each resource block.

Lastly, we evaluate our proposed scheme via comprehensive simulations of picocells along an urban canyon in which both inter- and intra-cell interference are taken into account. According to our simulation results, as we shrink cells (down to the cell width of 20(m)), users spectral efficiency is mostly ($\geq 97\%$) limited by hardware limitations (even with spectral efficiency of s_M=6 bps/Hz corresponding to 64-QAM). Larger cells however, are more prone to interference, yet our proposed scheme provides users with sufficient spectral efficiency for employing smaller constellations like QPSK. We then computed the overall capacity per square kilometer for a typical region in Manhattan area and demonstrated that dense mm-wave picocellular networks can actually deliver the promised 1000-fold capacity increase over the today's cellular networks.

1.3   Compressive channel estimation and tracking

In **Chapter 4**, we address fundamental signal processing challenges associated with channel estimation and tracking for large arrays, placed within the context of system

design for a mm-wave picocellular network.

While the signal processing and system design concepts presented here are of rather general applicability, our numerical results are for the particular setting that we discussed in prior chapters 1.2.

At the base station, we consider very large 32×32 arrays (such 1000-element arrays are still only palm-sized at a carrier wavelength of 5 mm) targeting the long term, as well as "moderately sized" 8×8 arrays (which can fit within an area of about half a square inch) which are currently realizable. Note that 16-element arrays were reported several years ago [25], and are already deployed in existing 60 GHz products, while 32-element arrays have been prototyped [26]. We assume that mobile devices are equipped with smaller 4×4 antenna arrays.

We focus on downlink 60 GHz communication, with the goal of enabling base station arrays to perform transmit beamforming towards mobile devices, despite the challenges posed by mobility and blockage (which occurs more easily at smaller wavelengths). We do not count on reciprocity. The uplink could be a standard LTE or WiFi link at lower carrier frequencies, used both for uplink data (not modeled here) and feedback for enabling spatial channel estimation at the transmitter.

We consider RF beamforming, in which a common baseband signal is routed to/from the antenna elements, and we can only control the amplitude and phase for each element. Indeed, we go even further, assuming that the amplitude for each element is fixed, and that we can only apply coarse four-phase control for each element. Standard least squares array adaptation and channel estimation techniques, which require access to the baseband signals associated with each element, do not apply in this setting. Instead, we consider here a *compressive* approach which exploits the sparsity of the mm-wave channel, so that relatively few measurements are required for channel estimation despite the large number of array elements.

Random phases
from $\pm 1, \pm j$

Mobiles
feedback

$v_{1,1} v_{2,1}$ $v_{1,2} v_{2,2}$

BS estimates channel
compressively

1st Rx measurement weight

L Rx measurement
weights

M Tx beacons

Same M beacons

$\bullet\ \bullet\ \bullet$

$v_{1,1}\ \ v_{2,1}$ $v_{1,2}\ \ v_{2,2}$
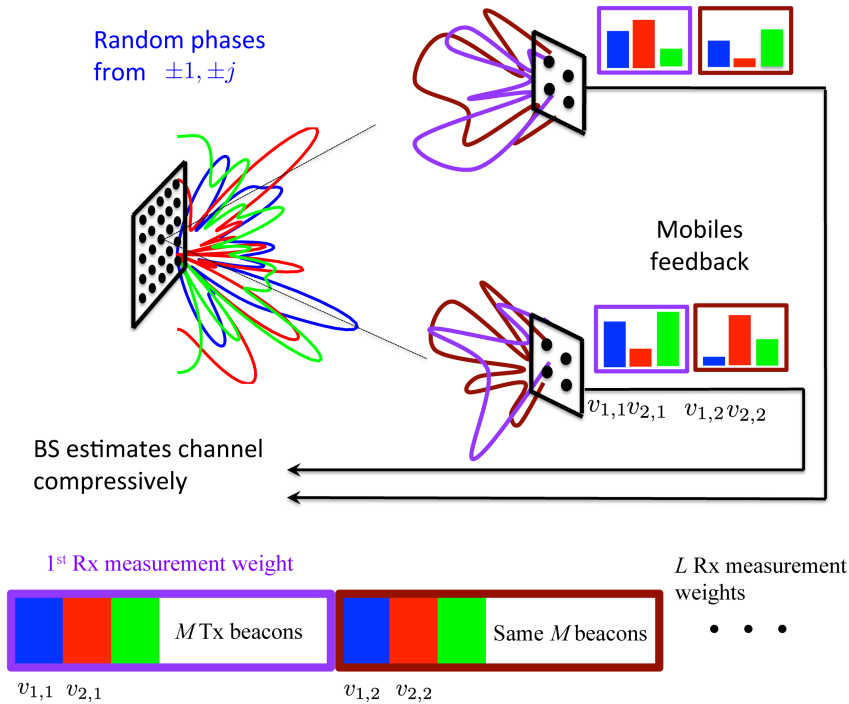
Figure 1.4: Channel sounding scheme: The transmitter repeats the M transmit bea-
cons L times so that receivers can measure the channel $v_{i,j}$ corresponding to the ith
setting at the transmitter and the jth at the receiver.

An alternative approach to spatial channel estimation with RF beamforming is ex-
haustive search, but its overhead does not scale for the large arrays of interest to us.

Hierarchical codebook search is more efficient [27, 28, 29, 30], but it cannot guarantee
overhead reduction in a multiuser setting. Furthermore, the multi-resolution beam pat-
terns used in hierarchical search require more than one RF chain [30], and these methods
only provide binned estimates of path directions. In contrast, the method presented here
provides fine-grained estimates while using only a single RF chain.

Chapter 4 of this thesis builds on prior conference papers of our group on compressive
array adaptation [31, 32], but goes well beyond them in several respects. In addition
to a more detailed development of the analytical framework underlying our estimation
algorithm, we now explicitly model the receive array at the mobile, which requires a
generalization of the beaconing and feedback strategy. We also address system level
design for compressive tracking in far greater detail, discussing link budget and overhead,

and accounting for inter-cell beacon interference. The previous work on compressive array adaptation [31] subsequently led to a general framework for compressive estimation [33, 34], which identify the isometries required to preserve fundamental bounds such as the Ziv-Zakai (ZZB) and Cramer-Rao (CRB), and use the relationship between these bounds to provide criteria for determining the minimum number of compressive measurements required to preserve geometry and to permit accurate parameter estimation based on a signal corrupted by an AWGN. We now adapt these general results here to develop guidelines for system-level parameter choices.

1.3.1   Contributions

Our contributions in this chapter are summarized as follows:

Architecture: We propose a novel architecture in which base stations send out compressive beacons, with a different set of pseudorandom phases used to transmit each beacon. Each mobile measures the complex gains associated with each beacon compressively, using pseudorandom control of the phases of its receive array. The scheme, described in more detail later, is depicted in Figure 1.4: the base station sends M beacons, repeated L times, which permits the mobile to use L different settings of its own array to measure the associated complex gains. Each mobile feeds back appropriately chosen functions of its measurements to the base stations on the uplink (which can be a standard LTE link). Each base station use this information to estimate and track the dominant paths to each mobile that it receives feedback from.

Algorithms: For the regular two-dimensional (2D) arrays considered here, directions of arrival/departure map to 2D spatial frequencies. The base station estimates the spatial frequencies to each mobile using a simple sequential algorithm, shown to be near-optimal (in terms of approaching the Cramer-Rao Bound) in related publications. The algorithm

exploits the geometric continuity of the channel across successive beaconing intervals to reduce the required number of compressive measurements.

System Design: While we do not provide a complete system design centered around our compressive strategy, we do provide preliminary results addressing some of the most important issues. We show that the overhead incurred by our beaconing scheme is very small (less than 1%). Furthermore, while compressive beacons are essentially omnidi-rectional (in contrast to the highly directive beams employed for communication), the link budget suffices for accurate channel estimation, and a simple beacon reuse strategy suffices to control inter-beacon interference across picocells.

Chapter 2

Capacity of mm wave picocells: a first estimate

2.1 Introduction

Millimeter (mm) wave picocellular networks are a promising approach for delivering the 1000-fold capacity increase required to keep up with projected demand for wireless data: the available bandwidth is orders of magnitude larger than that in existing cellular systems, and the small carrier wavelength enables the realization of highly directive antenna arrays in compact form factor, thus drastically increasing spatial reuse. In this chapter, we carry out an interference analysis for mm wave picocells in an urban canyon, accounting for the geometry associated with the sparse multipath characteristic of this band. While we make some modeling simplifications, our analysis provides a strong indication of the very large capacity, of the order of *Terabits/sec per km,* provided by such networks, using system bandwidths of the order of a few GHz.

2.1.1   Related Work

There are a number of prior papers investigating interference in mm wave networks. The effect of directional links on mesh networks is investigated in [35], while indoor environments are considered in [36, 37]. Coverage and attainable data rates in outdoor mm wave networks are investigated in [38] using stochastic geometry models, with base stations, users and obstacles placed in the plane according to Poisson point processes. To the best of our knowledge, this is the first attempt to account for the geometry of an urban canyon to quantify interference and capacity in mm-wave networks.

2.2   Interference analysis for mm-wave picocells

In the next section, we provide a geometric characterization of the inter-cell interference in our model of an urban canyon 1.2. Two important simplifications: (a) we ignore interference across parallel urban canyons, as well as interference which might leak from cross streets; (b) we do not consider potential reflections from horizontal ledges. However, while more detailed modeling and simulations are needed to refine the interference and capacity estimates provided here to account for such effects, we expect the qualitative conclusions to remain unchanged.

2.2.1   Inter-cell Interference Characterization and Modeling

We investigate the inter-cell interference caused by the main lobe and side lobes separately, since they have different characteristics. Since we consider a large number of antenna elements, the main beam is narrow and is well modeled by a single ray. Side lobes are much weaker, but their directions are difficult to predict, hence we must be more careful in bounding their effect.

Main lobe interference

We consider basestation antenna arrays with a large number of elements forming a pencil beam towards the desired user. This "desired" beam can be along the LoS, or it can be a single bounce from a wall or the ground (e.g., when steering around an obstacle blocking the LoS). We seek to understand the interference such a beam creates for neighboring basestations. We can use ray tracing for this purpose, given the highly directive nature of the beam and the limited diffraction at small wavelengths [13].

We assume that each face only creates interference in the direction it is facing. The following theorem proves that the main beam will escape to the sky after a few bounces (Figure 2.1), assuming that we can ignore the effect of potential reflections from horizontal ledges. Specifically, we bound the distance (from the transmitting BS) over which the main beam can create interference. We denote by h_{max} the maximum height of users, by H_{BS} the height of a BS, and by d the width of a picocell shared among two opposite facing antennas on adjacent BSs (Figure 2.4).
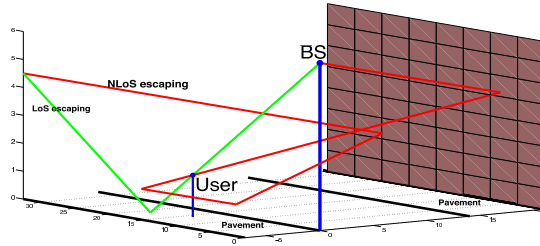
Figure 2.1: Mainlobe will escape to sky after a few bounces

Theorem 1 *The maximum range over which the main beam can create interference is bounded by $\frac{H_{BS}+h_{max}}{H_{BS}-h_{max}}d$. Thus, the main beam from a face creates interference for at most $N_{max} = \lceil \frac{H_{BS}+h_{max}}{H_{BS}-h_{max}} \rceil$ adjacent BSs in the direction it is facing.*

19

In order to prove this theorem, we need to introduce the "virtual" ray. When the "real" (i.e., physical) ray is reflected from any surface (ground or walls), the corresponding virtual ray is the straight line which is the mirror image of the real ray with respect to the reflecting surface. Thus, the direction of the virtual ray is unchanged by the reflection, and it "goes through" walls and the ground. Figure 2.2 shows an example of a real ray, and the corresponding virtual ray for a single reflection.
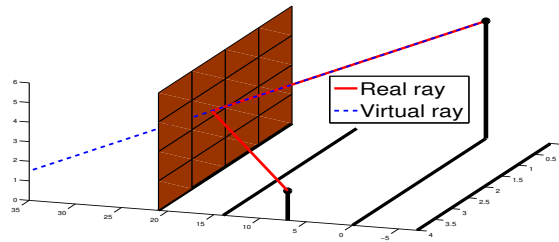
Figure 2.2: Real and virtual rays for a wall reflection

Proof: The proof is based on tracing the virtual ray associated with the main beam. Since the BS height is larger than that of any user, the main beam must go downward to reach the target user. It is easy to see that, in order to maximize the distance over which this beam creates interference, we must have a ground reflection.

Now, suppose that the main beam undergoes a ground reflection. The "real" reflected ray can only go upward, under our assumption that there is no horizontal surface above the ground that can reflect it again to create a downward trajectory. After the ground reflection, the virtual ray is as far below the ground as the real ray is above it (Figure 2.3). Setting $Z = 0$ as the ground surface, once the virtual ray crosses the $Z = -h_{max}$ plane, the real ray will have gone to a height of more than h_{max}, and can no longer create interference, since a potential victim user has height at most h_{max}.

Figure 2.4 provides a bird's eye view of the urban canyon, and lays out the coordinate system. The line equation of the virtual ray corresponding to the main beam is as follows:

20
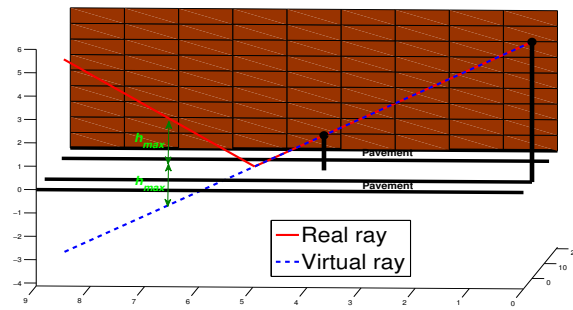
Figure 2.3: Ground reflection
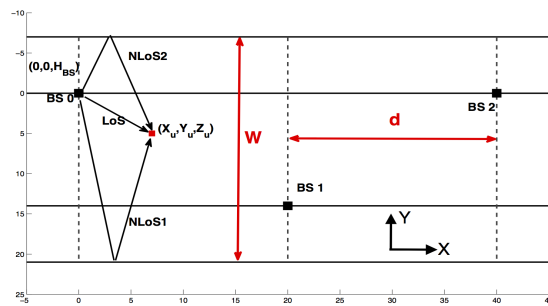
Figure 2.4: Problem geometry

$$\frac{X - X_{BS}}{a} = \frac{Y - Y_{BS}}{b} = \frac{Z - Z_{BS}}{c} \tag{2.1}$$

where $X_{BS} = Y_{BS} = 0$ and $Z_{BS} = H_{BS}$. As this virtual ray hits $Z = -h_{max}$, the actual ray has gone beyond the region in which it could interfere with users of height at most h_{max}.

$$X = \frac{a}{c}(z - H_{BS})|_{z=-h_{max}} = \frac{a}{c}(-h_{max} - H_{BS}) \tag{2.2}$$

in which a and c are as follows for LoS and different NLoS paths to the target user, specified in Figure 2.4.

	LoS	NLoS1	NLoS2	NLoS3
a	X_u	X_u	X_u	X_u
c	$Z_u - H_{BS}$	$Z_u - H_{BS}$	$Z_u - H_{BS}$	$-Z_u - H_{BS}$

Thus, we have the following expression for X:

$$X = \frac{X_u}{\mp Z_u + H_{BS}}(h_{max} + H_{BS}) \tag{2.3}$$

Since each BS is serving a user inside the immediate picocell, the maximum value of X corresponds to the maximum value of $X_u = d$ (at the edge of coverage for the face), as follows:

$$\max_u X = \frac{d(H_{BS} + h_{max})}{H_{BS} - h_{max}} \tag{2.4}$$

This is less than or equal to the width of $\frac{H_{BS}+h_{max}}{H_{BS}-h_{max}}$ picocells covered by $\lceil \frac{H_{BS}+h_{max}}{H_{BS}-h_{max}} \rceil$ adjacent BSs.   ■

For typical values of $H_{BS} = 6(m)$ and $h_{max} = 2(m)$ employed in our simulations, Theorem 1 implies that the main beam interferes with two adjacent BSs in the direction of the face producing the beam.

Sidelobe interference

While the main beam points towards a user inside the picocell, the direction of emission of sidelobes is highly variable, hence it is not possible to limit side lobe interference to a finite number of adjacent picocells. However, as stated in Theorem 2 below, the cumulative sidelobe interference seen within a given picocell is bounded (to a relatively small value), because of the geometric decay (with distance) of the strength of the interference from a distant picocell caused by oxygen absorption and reflection losses, along with the quadratic decay due to path loss. We consider a reference cell 0, and seek to quantify interference from cells $k > 0$ to its east and $k < 0$ from its west.

Denote by P the smallest received power over the desired link, which is given by

$$P = P_{Tx} G_{Tx} G_{Rx} \left(\frac{\lambda}{4\pi L_{max}} \right)^2 e^{-\beta L_{max}} \tag{2.5}$$

where P_{Tx},G_{Tx} and G_{Rx} are the transmitter power and the gains of Tx and Rx antenna arrays, respectively. The parameters λ, β and L_{max} denote, respectively, the wavelength, oxygen absorption coefficient (16 dB/km) and maximum length of a link inside a picocell.

Theorem 2 *For a user in cell 0, the sidelobe interference due to the BSs $[K, \infty)$ and $(-\infty, K]$ is bounded by $\alpha_K P$, where P is the smallest received power over the desired link.*

$$\alpha_K = \frac{\sum\limits_{k=K}^{\infty} I_k + \sum\limits_{n=-\infty}^{-K} I_k}{P} \tag{2.6}$$

where α_K decays geometrically with K, and is estimated below in (2.10).

Proof: The total interference introduced by the kth BS to a user inside $cell_0$ known as I_k is given by

$$I_k = \sum_{n=0}^{\infty} I_{k,n} \tag{2.7}$$

23

where $I_{k,n}$ is the interference component that travels from the kth BS undergoing n number of bounces before being captured by the Rx antenna array. It can be evaluated as follows:

$$I_{k,n} \approx N_n P_{Tx}(g_{Tx})_{k,n}(g_{Rx})_{k,n}(\frac{\lambda}{4\pi r_{k,n}})^2 e^{-\beta r_{k,n}}(\frac{1}{l_{k,n}})^n \tag{2.8}$$

where $(g_{Tx})_{k,n}$, $(g_{Rx})_{k,n}$ and $r_{k,n}$ are the Tx gain, Rx gain and the length of corresponding n times reflected path from kth BS, respectively. Moreover, $\frac{1}{l_{k,n}}$ is the corresponding reflection loss coefficient which depends on the incidence angle and the reflecting surface material. The factor N_n accounts for the number of possible paths (corresponding to different possible ordering of reflectors). While the interference may differ across such paths, we can replace them by a common value because we seek a pessimistic estimate for the interference. By virtue of Theorem 1, for $k > \lceil \frac{H_{BS}+h_{max}}{H_{BS}-h_{max}} \rceil$, $(g_{Tx})_{k,n}$ is the gain of the Tx antenna array outside the main beam.

Note that $r_{k,0}$ is actually the distance between the kth BS and the target user in $cell_0$, and can be roughly approximated by kd. By Pythagoras's theorem, $r_{k,n}$ can be interpreted based on $r_{k,0}$ and W which is the distance between parallel reflectors (buildings or walls):

$$r_{k,n}^2 = r_{k,0}^2 + n^2 W^2 \approx (kd)^2 + (nW)^2 \tag{2.9}$$

This approximation applies to bounces between two walls. By using arguments similar to those in Theorem 1, there can be at most one ground reflection, and we can ignore its effect on the path length since we are after a pessimistic estimate of the interference. We have at most 4 different possible n-times reflected paths, accounting for whether we have the ground reflection and the order of walls' reflection: $N_n = 4$ for $n > 1$, $N_1 = 3$ and $N_0 = 1$. Thus, the total interference from cells $[K, \infty)$ or $(-\infty, -K]$, divided by P,

can be written as follows:

$$\frac{\alpha_K}{2} = \frac{\sum\limits_{k=K}^{\infty} I_k}{P} = \sum_{k=K}^{\infty}\sum_{n=0}^{\infty}\left(N_n \frac{(g_{Tx})_{k,n}(g_{Rx})_{k,n}}{G_{Tx}G_{Rx}}\left(\frac{L_{max}}{r_{k,n}}\right)^2\right.$$
$$\left. e^{-\beta(r_{k,n}-L_{max})}\left(\frac{1}{(l_{k,n})}\right)^n\right) \tag{2.10}$$

By substituting (2.9) above, we obtain

$$\frac{\sum\limits_{k=K}^{\infty} I_k}{P} \le 4L_{max}^2 e^{\beta L_{max}} \sum_{k=K}^{\infty}\sum_{n=0}^{\infty}\left(\frac{1}{(kd)^2 + (nW)^2}\right.$$
$$\left. e^{-\beta\sqrt{(kd)^2+(nW)^2}}\left(\frac{1}{(l_{min})}\right)^n\right) < 4\sum_{k=K}^{\infty} e^{-\beta kd}\sum_{n=0}^{\infty}\frac{1}{(l_{min})^n} \tag{2.11}$$
$$= 4\frac{e^{-\beta K d}}{1 - e^{-\beta d}}\frac{l_{min}}{l_{min} - 1}$$

The same computation holds for interference introduced by cells $(-\infty, -K]$. Thus, α_K defined in equation 2.6 is bounded and decreases exponentially with increasing K.   ∎

We now compute α_K using (2.6) and (2.8), and plot α_K (dB) versus K in Figure 2.5. The results are for a 32×32 array at the BS TX, and a 4×4 array at mobile RX. For $K \ge 3$, we only capture interference from sidelobes (Theorem 1), hence $\frac{(g_{Tx})_{k,n}}{G_{Tx}}$ has been substituted by average sidelobe strength relative to the main beam, averaged over the different directions that the main beam can take for users in the picocell. The normalized receive gain $\frac{(g_{Rx})_{k,n}}{G_{Rx}}$, however, is set to the average antenna gain strength in all directions relative to the main beam. We fix the reflection loss at 5 dB, to avoid detailed modeling of material and angle of incidence. This is the smallest reflection loss encountered in outdoor environments for most surfaces, according to measurements reported in [39], hence it is expected to give pessimistic estimates of the interference. We see from Figure 2.5 that α_K decreases exponentially with K, with higher rate of decrease

for wider picocells. We also note that the interference is dominated by the contribution from LoS and first order reflections ($n = 0, 1$); the interference due to these are plotted as dashed lines, and falls on top of the net interference curves obtained by summing over n. Thus, in the simulations in the next section, we restrict attention to $n = 0, 1$.
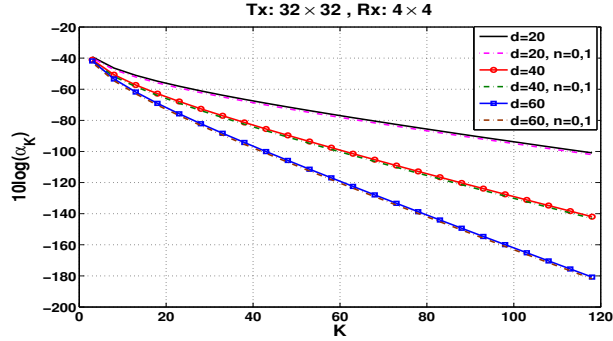
Figure 2.5: Normalized Intercell interference

By Theorem 1, if we wish to avoid main beam interference, then $\lceil \frac{H_{BS}+h_{max}}{H_{BS}-h_{max}} \rceil$ adjacent BSs have to coordinate. For $H_{BS} = 6(m)$ and $h_{max} = 2(m)$, this means that every 3 adjacent BSs have to coordinate. Suppose, for example, that we orthogonalize transmissions among such sets of 3 BSs (i.e., with a frequency reuse of 3). Now, from the computations associated with Theorem 2 shown in Figure 2.5, the cumulative interference caused by sidelobes from BSs beyond this set ($K \geq 3$) is at least 40 dB weaker than the desired received power. Thus, a frequency reuse of 3 leads to very large SINR, so that our spectral efficiency is expected to be bounded only by hardware considerations. This is verified by Monte Carlo simulations in the next section. However, we also show in the next section that such orthogonalization is wasteful. Given the interference reduction due to narrow beams, much larger network capacity can be obtained (at the expense of a small collision rate) with spatial reuse one.

2.3   An Initial Capacity Estimation

We now use Monte Carlo simulations for evaluation of inter-cell interference and capacity. We wish to quantify the potential capacity gains from shrinking the picocell width (e.g., down to 20 m). We consider an urban canyon of length 1 km, and consider the interference seen by a typical user in a picocell in the middle of this canyon, which would see the most interference. By virtue of Theorems 1 and 2, we ignore interference coming from outside the 1 km segment.

Since a user in the target picocell can be served by one of two BSs on two different sides, it is unlikely for her body to block the LoS path from both. Furthermore, as we shrink the picocell width, the LoS path slants more steeply downward, hence it is difficult for other obstacles (e.g., pedestrians, cars) to block it. Thus, in our computations, we assume for simplicity that the LoS path is available to the desired user. Of course, both LoS and first order NLoS paths are accounted for when computing interference from other BSs. (as noted earlier, Figure 2.5 shows that the interference from higher order reflections is negligible in comparison).

Figure 2.6 is the complementary CDF (CCDF) of the achievable SIR(dB) for $20(m)$ width picocells. We consider 32×32 Tx and 4×4 Rx antenna arrays, the same settings as for the analytical computations shown in Figure 2.5. We note that, for a frequency reuse of three, orthogonalizing every three adjacent picocells, the SIR CCDF is consistent with our observation that the main lobe interference has been eliminated (Theorem 1) and sidelobe interference is at least 40 dB weaker than the desired received power as in Figure 2.5.

In order to evaluate overall network performance, we now scale back to a smaller 8×8 BS TX array, while still keeping the mobile RX array at 4×4. These values are chosen because they are close to the current state of the art (32 element arrays are already
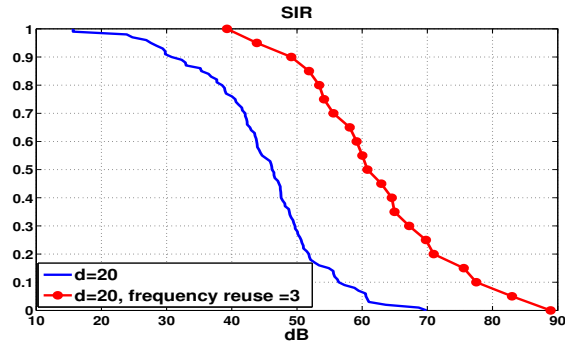
Figure 2.6: CCDF of achievable SIR for a 32×32 Tx and 4×4 Rx antenna arrays.
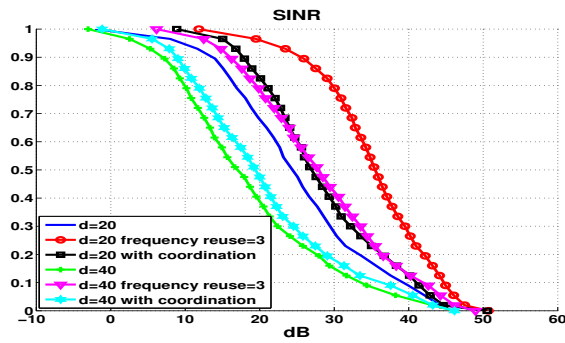
Figure 2.7: CCDF of achievable SINR for 8×8 Tx and 4×4 Rx antenna arrays.

deployed in commercial 60 GHz products), and it turns out that they suffice to provide full spatial reuse as we scale down cell sizes.

 Figure 2.7 shows the CCDF of the achieved SINR(dB) for three different scenarios:

1) Frequency reuse of one and no coordination: each BS randomly chooses to serve a single user in its own picocell, and hence may interfere with the target user through LoS or NLoS paths.

2) Frequency reuse factor of three, orthogonalizing every set of three adjacent BSs. As already noted, this eliminates main beam interference and drastically reduces sidelobe interference.

3) Multi-User Diversity (MUD) with sets of three adjacent BSs coordinating (minimally). If the SINR of the target user is less than a threshold (set at a fairly high 40 dB in these simulations), out of the two adjacent BSs, the one which introduces the most interference is asked to change its main beam's direction and choose another user to serve (we assume that such a user is always available and is randomly located within the coverage area). If this results in more interference at the target user, the interfering BS reverts to its previously chosen user.

It is possible to consider far more complicated scheduling rules, and optimally choose users to serve among groups of neighboring BSs. However, we leave detailed exploration of medium access control and scheduling to the next section, seeking here just to provide some simple capacity benchmarks.

As seen in Figure 2.7, a frequency reuse of three increases SINR by up to 20dB, mainly because of elimination of main beam interference. However, the corresponding 3X reduction in signaling bandwidth leads to a significant penalty in achievable data rates; see Figure 2.8. The third scenario, which utilizes multiuser diversity, offers relatively small improvement. This is because our minimal coordination strategy only allows one possible switch. (In addition to simplifying medium access control, another reason for

our minimalism is so as not to count on a large pool of users for multiuser diversity within a small picocell.)

We also notice from Figure 2.7 that the SINR is worse for larger cell widths ($d = 40(m)$ compared to $d = 20(m)$). This might seem surprising, especially because the results in Figure 2.5 show that sidelobe interference decreases with d. However, the increased interference is due to the main beam from neighboring BSs: for larger cell widths, the target user can be farther away from the BS, and, since the BS height is fixed, the main beam slants down less and therefore interferes with a larger region in the adjacent cell.

In order to estimate capacity, we use the following approximation for the spectral efficiency (in bps/Hz)

$$s = \frac{1}{F}\min\left(s_M, \log_2(1 + SINR)\right) \tag{2.12}$$

where F denotes the frequency reuse factor, and s_M the maximum spectral efficiency supported by the system. In our simulations, we (somewhat arbitrarily) set $s_M = 6$ bps/Hz, corresponding to uncoded 64QAM (in practice, we would use light coding). Such large constellations may be a stretch with today's hardware, given the phase noise in mm-wave radios and the difficulty of high-precision digitization at large bandwidths, but we hope that such hardware limitations would be overcome in the future.

The achievable data rate for a BS face is therefore given by $R_{picocell} = sB$, where B denotes the available system bandwidth. Figure 2.8 shows the achievable data rates in a picocell for system bandwidth $B = 2$ GHz. We do not account for excess bandwidth in our data rate estimates. While frequency reuse of 3 (which could be implemented via either time division or frequency division) yields more deterministic performance, the achievable rate is about 4 Gbps, which is significantly smaller than the 90% availability rate of 12 Gbps for our third scheme (full reuse, minimal coordination). This data rate corresponds to saturation of spectral efficiency at $s_M = 6$ bps/Hz, and would scale down

30

if we used smaller constellations (e.g., 4 Gbps for QPSK). Thus, as long we allow a small outage probability for occasional collisions across adjacent picocells, hardware rather than interference is the bottleneck.
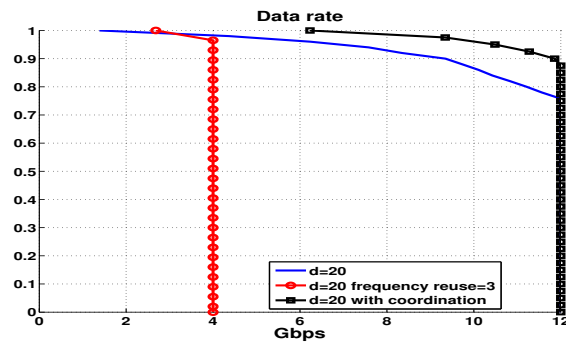
Figure 2.8: CCDF of achievable data rates over 2GHz bandwidth for 8×8 Tx and 4×4 Rx antenna arrays

We can now estimate the capacity of a picocellular network deployed over a length of urban canyon. Since each BS has two faces, there are twice the number of active links as BSs. For $d = 20(m)$, there are 50 BSs per km, and hence 100 active links. According to Figure 2.8, each link can support 12 Gbps at 90% availability, so that the entire network capacity is estimated at 1.2 Tbps per km using 2 GHz of spectrum.

Chapter 3

Enhancing capacity via managed intra-cell interference

3.1 Introduction

In addition to cell densification, one can attain further spatial reuse *within* the cell by increasing the number of subarrays on each base station. However, this benefit comes with the pitfall of *intra-cell inteference*, i.e., when a transmitter interferes with receivers in the *same* cell that it does not target. This could significantly reduce the spectral efficiency of spatially correlated users.

In this section, we first characterize intra-cell interference in our system model and then propose a *cross-layer* approach to deal with it. To this end, we combined techniques from two broad areas that have been studied in the literature: (a) downlink linear precoding and power control [18, 19, 20, 21, 22] (b) powerful optimization approaches recently developed for network-level resource allocation [23, 24]. Here is a brief description of our two-step method:

1. Given that a resource block is assigned to a pre-defined set of users, we develop

a building block at the PHY-layer, which employs an optimal linear method (i.e., LMMSE) for beamforming and power allocation to suppress the LoS intra-cell interference among them.

2. We then incorporate the PHY-layer block in designing the MAC-layer protocol, which determines the set of active users on each of the resource blocks.

We then evaluate our proposed scheme via comprehensive simulations of picocells along an urban canyon in which both inter- and intra-cell interference are taken into account. Our simulation results demonstrate that, as we shrink cells (down to the cell width of 20m), users' spectral efficiency is mostly ($\geq 97\%$) limited by the hardware limitations (even with large constellations such as 64-QAM). A quick-glance comparison with our previous results (from Chapter 2) indicates that we are able to increase the capacity by a factor of K (at least for small number of subarrays per face i.e., K=2) in small cells. Larger picocells are more prone to interference and do not enjoy multiple subarray architecture, yet our proposed scheme provides users with sufficient spectral efficiency for employing smaller constellations like QPSK. Lastly, we computed the overall capacity per square kilometer for a typical region in Manhattan area and demonstrated that dense mm-wave picocellular networks can actually deliver the promised 1000-fold capacity increase over the conventional LTE networks.

3.2   System Model

In this chapter, we consider the same street canyon as before (Figure 1.2) except that we consider K subarrays placed each face of a basestation (Figure 3.1).

We now describe the channel model. Since mm-wave channels can accurately be estimated by efficient algorithms proposed in Chapter 4, we assume that channel knowledge
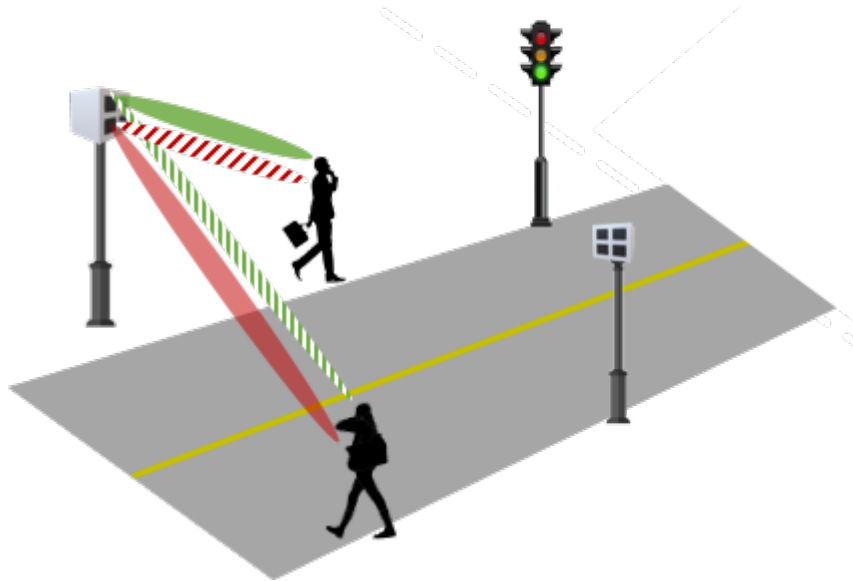
Figure 3.1: Multiple subarrays placed on each face of a basestation which leads to intra-cell interference.

is available at both the base station and mobile users. We assume that the transmitters have learned channel matrices $\{\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_K\}$, each of size $M \times N$ where M is the antenna size of the mobile user and N that of the transmitter. $\mathbf{H}_k$ is characterized by the path loss and spatial frequencies between the k-th transmitter and its corresponding mobile user.

Note that link distances are large enough that all transmitters installed on a face could be approximated as co-located from the users point of view. Therefore, the channel matrix from all transmitters to the k-th user is same as $\mathbf{H}_k$. We have also made the assumption that target users are served through the LoS path, and hence each mobile user automatically beamforms towards all the transmitters.

3.3 Characterization and Mitigation of Intra-cell Interference

3.3.1 Intra-cell Interference Characterization

Similar to the inter-cell case, intra-cell interference is composed of LoS and NLoS components (depicted in Figure 3.2). However, the LoS component is expected to be the dominant one for three reasons:

1. Since the target users are served via the LoS path, the receiver's main lobe is unlikely to encompass the NLoS components of interference. The LoS component would in contrast get *amplified* by the same amount as the desired signal.

2. NLoS components are subject to higher path loss.

3. NLoS components suffer from reflection loss induced by reflecting surfaces.

Our simulation results for the same urban canyon scenario also validate this assumption (depicted in Figure 3.2(b).)
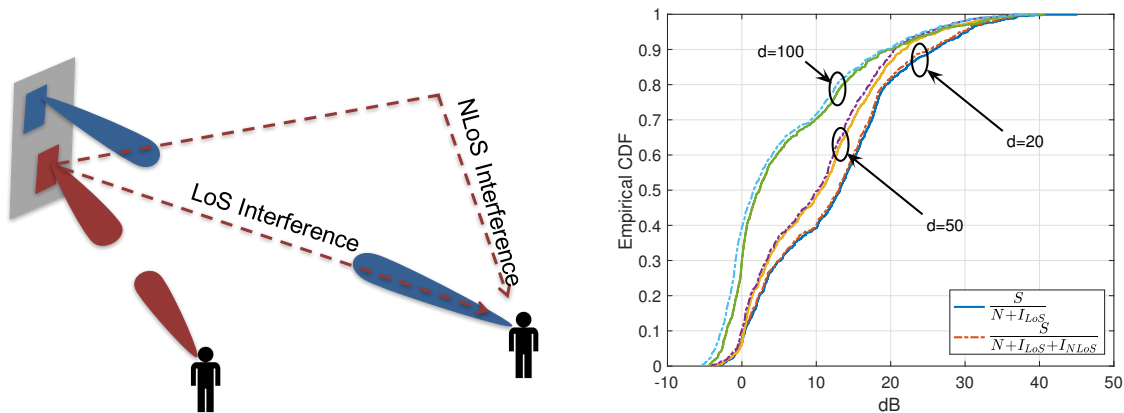
Figure 3.2: (a) Intracell interference composed of LoS and NLoS components, (b) CDF of Signal to Intracellular Interference

Therefore, we assume that intra-cell interference can be alleviated by suppressing the LoS component only. For the rest of this section, by the term interference we refer to the LoS component of intra-cell interference.

3.3.2 PHY layer design: Power allocation and beamforming

Mitigation of co-channel interference in multiuser MIMO has been extensively studied in the literature [18, 19, 20, 21, 22]. Different approaches such as precoding, transmitter or/and receiver beamforming, power adaptation, etc. have been explored. In this section, we restrict ourselves to RF beamforming and power control to avoid the hardware complexity of digital precoders.

In the context of power control and beamforming, there are two classical optimization problems: (a) sum-rate maximization and (b) minimum-rate maximization, subject to the power constraint(s). The former is often studied in the context of information-theoretic capacity, and does not guarantee fair sharing of resources among users. We therefore focus on the latter, which guarantees a minimum level of QoS (Quality of Service) for each of the streams.

The minimum-rate optimization can be translated to the following problem:

$$\mathcal{S}(P_T) = \begin{cases} \max_{\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \cdots, \boldsymbol{\omega}_K\}} & \min_i \text{SINR}_i \\ \text{s.t.} & \sum_{k=1}^{K} \|\boldsymbol{\omega}_k\|_2^2 \leq P_T \end{cases} \tag{3.1}$$

where $\boldsymbol{\omega}_k \in \mathbb{C}^N$ is the transmit beamforming vector aimed at the k-th user, $\|\boldsymbol{\omega}_k\|_2^2$ is the power consumed by the k-th antenna array, and SINR_k is the signal to interference

ratio at k-th receiver

$$\text{SINR}_k = \frac{\left|\boldsymbol{\omega}_k^H \mathbf{h_k}\right|^2}{\sum_{\substack{i=1 \\ i \neq k}}^{K} \left|\boldsymbol{\omega}_i^H \mathbf{h_k}\right|^2 + \sigma_k^2}$$

A straightforward argument shows that (3.1) will result in the same SINR for all the users, and hence the maximum index of fairness is guaranteed.
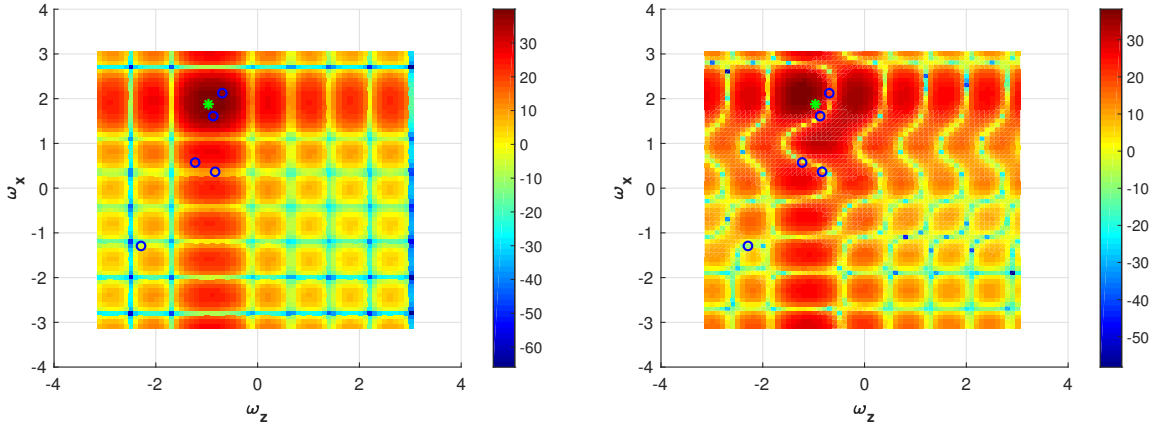
Figure 3.3: Transmit antenna patterns causing intra-cell interference (left) and the new antenna patterns after employing interference suppression via Algorithm 1 (right). The spatial frequency of the target user is marked by a green star, and the remaining users are marked by blue circles. Employing Algorithm 1 (right) aligns the non-targeted users with the null directions.

Our solution to problem (3.1) builds on previous work in [18, 19]. We start with the related power optimization problem

$$\mathcal{P}(\gamma) = \begin{cases} \min_{\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \cdots, \boldsymbol{\omega}_K\}} \sum_{k=1}^{K} \|\boldsymbol{\omega}_k\|_2^2 \\ \\ \text{s.t.} \qquad \min_i \text{SINR}_i \geq \gamma \end{cases} \tag{3.2}$$

It was shown in [19] that (3.1) and (3.2) are inverse problems, meaning that $\mathcal{S}(\mathcal{P}(\gamma_0)) = \gamma_0$ and $\mathcal{P}(\mathcal{S}(P_T)) = P_T$. Furthermore, (3.2) has an iterative solution [18]. We leverage these

observations to formulate Algorithm 1, which iteratively solves (3.2) for increasing values of γ until the power constraint in (3.1) is saturated. The solution to (3.2) employs LMMSE (Linear Minimum Mean Square Error) to estimate the transmit beamforming vector (lines 8-15) , followed by power allocation to enforce the minimum SINR constraints (line 16).

Algorithm 1 PHY layer design

1: **Input:** $\{p_i^0, \mathbf{h}_i\}$ $\forall i,$ $\gamma,$ $\Delta\gamma$

2: **Output:** $\{\boldsymbol{\omega}_i\}$ $\forall i,$ γ

3: **procedure** BEAMFORMING AND POWER ADAPTATION

4:      Compute normalized channels: $\tilde{h}_k = h_k/\sigma_k^2$ $\forall k$

5:      **while** $G_{max}\|\boldsymbol{\omega}_k\|_2^2 \leq EIRP,$ $\forall k$ **do**

6:          $\gamma = \gamma + \Delta\gamma$

7:          $n \leftarrow 0$

8:          **repeat**

9:             **for** $k \in \{1, 2, \cdots, K\}$ **do**

10:               $\hat{\boldsymbol{\omega}}_k^n = \mathrm{argmin}_{\boldsymbol{\omega}_k} \sum_{\substack{j=1 \\ j\neq k}}^{K} p_j^n|\boldsymbol{\omega}_k^H\tilde{h}_j|^2 + \|\boldsymbol{\omega}_k\|_2^2,$

                    s.t. $\boldsymbol{\omega}_k^H\tilde{h}_k = 1$

11:               $p_k^{n+1} = \gamma \sum_{\substack{j=1 \\ j\neq k}}^{K} p_j^n|(\hat{\boldsymbol{\omega}}_k^n)^H\tilde{h}_j|^2 + \gamma\|\hat{\boldsymbol{\omega}}_k^n\|_2^2$

12:               $\tilde{p}_k^{n+1} = \gamma \sum_{\substack{j=1 \\ j\neq k}}^{K} \tilde{p}_j^n|(\hat{\boldsymbol{\omega}}_j^n)^H\tilde{h}_k|^2 + \gamma$

13:             **end for**

14:             $n \leftarrow n+1$

15:          **until** convergence

16:          $\boldsymbol{\omega}_k = \sqrt{\tilde{p}_k}\hat{\boldsymbol{\omega}}_k$ $\forall k$

17:      **end while**

18: **end procedure**

Figure 3.3 illustrate how the algorithm distorts the transmitter antenna pattern by pushing nulls toward the users that the transmitter does not target. This improves SINR but might cause SNR degradation (as shown in Figure 3.4).

Some remarks on the algorithm:

- Intuitively, the goal of the optimization problems in (3.1) and (3.2) is to manipulate the transmitters antenna pattern to minimize the induced interference toward the non-targeted users while maintaining constant gain along the desired direction. In addition, power adaptation is employed to cope with link distance variations.

- In practice, we have individual power constraints on the Equivalent Isotropically Radiated Power (EIRP), which impose the following constraint:

$$G_{max}\|\boldsymbol{\omega}_k\|_2^2 \leq EIRP \quad \forall k \in \{1, 2, \cdots, K\}$$

where G_{max} is the maximum array gain provided by the antenna and EIRP is the limit established by FCC (Federal Communications Commission) for different frequencies (e.g., EIRP=40 dBm at 60 GHz). Our iterative solution allows us to impose the individual power constraints by setting the stopping criteria as when any of the transmit powers has reached the threshold (line 5).

- We have omitted the effect of the receiver antenna array in our formulation. Specifically, the channel matrix $\mathbf{H}_k$ has been replaced by a vector $\mathbf{h}_k$. This is for two reasons:

  1. In order to limit te complexity of mobile receivers, interference suppression is employed at the base station alone.

  2. For intra-cell interference, the receiver antenna provides an array gain of M for both signal and interference. Thus, it does not affect performance in an interference-limited scenario.

  That said, we note that in a noise-limited regime (which is most likely the case

once interference suppression algorithms are applied) receiver antenna arrays boost SINR by a factor of M. Therefore, we take the receiver arrays back into account for our simulation results (Section 3.4).
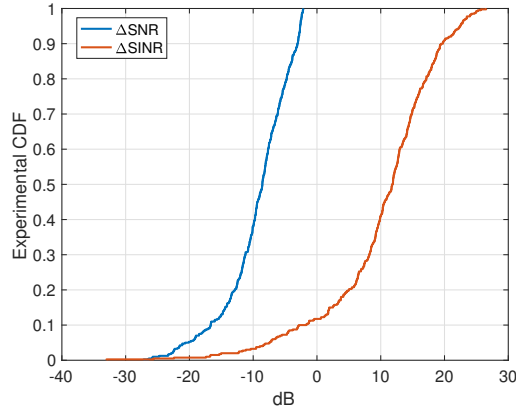
Figure 3.4: CDF of ΔSNR and ΔSINR, showing how employing Algorithm 1 improves SINR values while it might cause SNR degradation. By nullforming along the undesired directions, we boost the SINR (which essentially determines the data rate). However, there might be an SNR degradation due to sidelobe enhancement. ΔSNR denotes $\text{SNR}_{\text{after nullforming}} - \text{SNR}_{\text{before nullforming}}$.

3.3.3   MAC layer design: Resource allocation

The preceding PHY layer optimization is for sharing a single resource block among a pre-defined set of users. In this section, we consider interference management in the MAC layer, where resources are divided into blocks (resource granularity) and only certain users allowed to operate in each block (user selection). Intuitively, these additional degrees of freedom can be exploited in the following manner: by selecting spatially separated users to operate in the same block, we can mitigate interference and increase spectral efficiency.

Preliminaries:

Consider a cell with Q users sharing frequency band B over a frame of duration T. We make two assumptions:

1. The frame duration T is small enough that mobile users can be considered to be quasi-stationary over a frame.

2. The directive antenna arrays employed on both transmitter and receiver suppress multipath fading sufficiently that we may approximate the channel as frequency non-selective.

We consider resource allocation via time division, so that at every point in time each active user utilizes the entire bandwidth B. For simplicity we allow an infinite time granularity.

We need to allocate each frame to a subset of users. Denoting by $\mathcal{Q}$ the set of all users, we define $\mathcal{P}_{\leq K}(\mathcal{Q})$ as the set of all possible subsets of users (*configurations*) that can be served simultanously by (up to) K antenna arrays:

$$\mathcal{P}_{\leq K}(\mathcal{Q}) = \{U_c \subset \mathcal{Q} \mid |U_c| \leq K\}$$

We wish to find the fraction of resources that should be allocated to each of these configurations in order to maximize sum (or minimum) spectral efficiency. More specifically, let x_c represent the portion of resources allocated to the c-th configuration. We want to find $\mathbf{x} = [x_1, x_2, \cdots, x_C]^T$ where $C = \sum_{k=0}^{K} \binom{Q}{k}$ is the cardinality of $\mathcal{P}_{\leq K}(\mathcal{Q})$.

The spectral efficiency for the q-th user under policy $\mathbf{x}$ is then defined as

$$r_q = \sum_{c=1}^{C} x_c \, \log(1 + \gamma_c^q) \quad \text{(bits/sec/Hz)} \tag{3.3}$$

41

where γ_c^q is the SINR of the q-th user under c-th configuration (where U_c is the set of active users.) Clearly, we set $\gamma_c^q = 0$, for $q \notin U_c$).

The resource allocation problem:

Like the optimization problems for beamforming and power adaptation, the resource allocation problem could also be formulated to maximize either the sum-rate or the min-rate. In order to provide fairness among users, we focus on the min-rate version, which can be formulated as follows:

$$\max_{\mathbf{x}} \ \min_q \ r_q \tag{3.4}$$
$$\text{s.t.} \quad S^T \mathbf{x} = \mathbf{r}$$
$$\mathbb{1}^T \mathbf{x} = 1$$
$$\mathbf{x} \succeq \mathbb{0}$$

In the first constraint, we have rewritten the equations in (3.3) in a matrix form by defining $S_{C \times Q} = [s_{cq}]$ where $s_{cq} = \log_2(1 + \gamma_c^q)$ is the spectral efficiency of the q-th user under the c-th configuration and $\mathbf{r} = [r_1, r_2, \cdots, r_Q]^T$ is the vector of resultant spectral efficiency over all resource blocks. The last two conditions ensure that sum of the portions allocated to different configurations add up to one and neither of them can be negative.

Theoretically, allocation policies resulting from (3.4) should maximize the min-rate among users. However, in practice we might not be able to attain the theoretical rate due to hardware constraints. If s_M is the hardware-constrained spectral efficiency limit, the maximum min-rate will be bounded by $(K/Q)\, s_M$. This corresponds to the *saturation* point where all transmitters operate at their highest modulation rate.

Figure 3.5 shows the empirical CCDF of maximum min-rate for different cell sizes, along
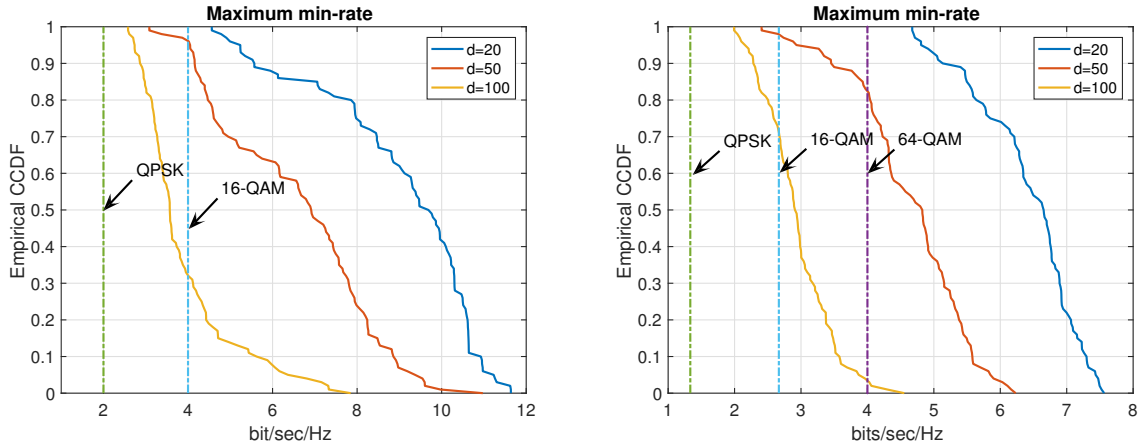
Figure 3.5: Empirical CCDF of the maximum min-rate for (a) Q=K=4 (b) Q=6, K=4. In scenarios with smaller cell sizes, maximum min-rates are larger than and hence limited to the saturation point imposed by modulation (vertical lines). This is because smaller cells have (almost) vertically aligned beams which illuminate a smaller region and hence induce intra-cell interference for a smaller number of users. Moreover, larger number of users could increase the attainable spectral efficiency by enabling us to utilize multiuser diversity to avoid interference.

with the saturation point imposed by the various modulations (i.e., $(K/Q) s_M$). As depicted in Figure 3.5, for smaller picocells with larger number of users ($d \leq 20(m)$ and $Q > K$) spectral efficiency is limited by hardware rather than noise or interference.

Some remarks:

- The optimization problem in (3.4) maximizes the worst users' spectral efficiency and therefore will result in equal rate for all users in $\mathcal{Q}$. Its performance is therefore inherently bounded by that of the worst user. However, there are certain scenarios where we can maximize the sum-rate as well: for example, when we have surplus resources after providing all users with the minimum required datarate r_{min}. Therefore, if the resultant min-rate provided by the allocation policy in (3.4) is greater than r_{min}, we employ the following optimization problem to maximize the sum-rate

by utilizing multiuser diversity.

$$\max_{\mathbf{x}} \; \mathbb{1}^T S^T \mathbf{x} \qquad\qquad (3.5)$$

$$\text{s.t.} \quad S^T \mathbf{x} \succeq r_{min} \mathbb{1}$$

$$\mathbb{1}^T \mathbf{x} = 1$$

$$\mathbf{x} \succeq \mathbb{0}$$

- An important observation is that an optimal allocation policy typically allocates more resource blocks to configurations with a larger number of users. This is because the overall datarate is linearly proportional to the number of simultaneous users, whereas the dependence on SINR is logarithmic. However, there are settings in which time multiplexing leads to a higher data rate than spatial multiplexing (for example, when users are highly spatially correlated such that by eliminating their mutual interference, higher data rates can be attained even over smaller number of resource blocks.)

- Figure (3.6) demonstrates this phenomenon by showing a few examples for the solution to the resource allocation problem. The optimal solution tends towards serving maximum number of users (K) simultaneously (blue portions) unless the induced interference is so large that only a subset of them are served (green or red portions).

3.4   Capacity Estimation

We now demonstrate via simulations that mm-wave cells enjoy a significant gain in capacity over conventional LTE cellular networks, despite the increased amount of inter-
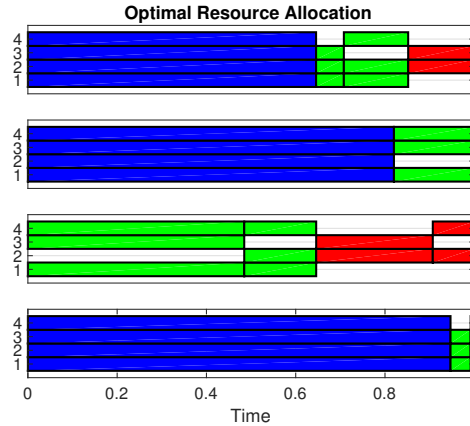
Figure 3.6: Optimal solution of the resource allocation problem for different realizations of mobile users. The picocell parameters are d=50m and $K = Q$=4. Optimal allocation policies tend to serve the largest possible number of users (blue portions) while in some cases it is better to turn off a subset of subarrays, i.e., green/red portions.

and intra-cell interference.

3.4.1   Preliminaries

Our interference analysis in the preceding sections is partially geometry dependent and specifically tailored for cells along an urban canyon. Hence for our simulations, we consider an urban canyon of length 1 km and investigate a picocell in the middle of this canyon, where users would see the most interference (Figure 3.7). By virtue of Theorems 1 and 2 in Chapter 2, we ignore interference coming from outside the 1 km segment.

Since a user in the target picocell can be served by one of two BSs on two different sides, it is unlikely for her body to block both LoS paths. Furthermore, as we shrink the picocell width, the LoS path slants more steeply downward, hence it is difficult for other obstacles (e.g., pedestrians, cars) to block it. Thus, in our computations, we assume for simplicity that at least one LoS path is available to every user. Of course, both LoS and first order NLoS paths are accounted for when computing interference from other base stations. (As noted in [40] interference from higher order reflections is negligible in

comparison.)

We consider 8×8 base station TX array and 4×4 mobile RX arrays. These values are chosen because they are close to the current state of the art (32 element arrays are already deployed in commercial 60 GHz products), and it turns out that they suffice to provide high spectral efficiency as we scale down cell sizes.

By Theorems 1 and 2 (Chapter 2), for a typical user served by BS_0, the interference induced by the base stations farther than $2d$ away from BS_0, is negligible. Specifically, in the scenario depicted in Figure 3.7, the following sources would interfere with the shaded user served with one of the eastward facing antenna arrays of BS_0:

1. *inter-cell* interference from K eastward facing antenna arrays on BS_{-2}

2. *inter-cell* interference from K eastward facing antenna arrays on BS_{-1}

3. *intra-cell* interference from K-1 eastward facing antenna arrays on BS_0

4. *inter-sector* interference from K westward facing antenna arrays on BS_1

5. *inter-cell* interference from K westward facing antenna arrays on BS_2

each of which is composed of LoS and multiple NLoS components.
Note that the *inter-sector* interference refers to the interference caused by the immediate neighboring base station on the opposite side of the street (BS_1) which serves users inside the cell. This should have been accounted as intra-cell interference but since it has different characteristics from item three above, we put a different name on it.

In our simulations, we have employed frequency reuse of two which automatically eliminates items two and four above. We also attenuate the LoS interference of item three by employing Algorithm 1 introduced in section 3.3.2. We compute the overall spectral efficiency, $\log_2(1+\text{SINR})$, for the users served by BS_0 by taking into account the
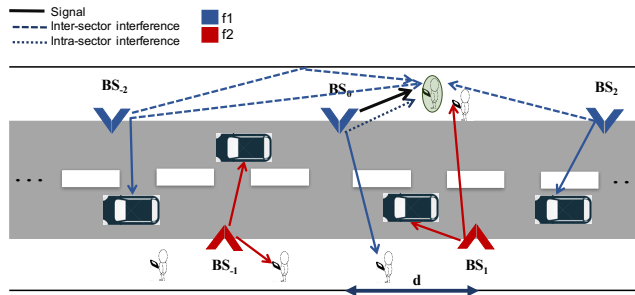
Figure 3.7: Simulation scenario

residual interference from items one, three and five. The resultant matrix S is then fed into the optimization problem (3.4) to obtain the maximum min-rate obtained by the optimal time allocation.
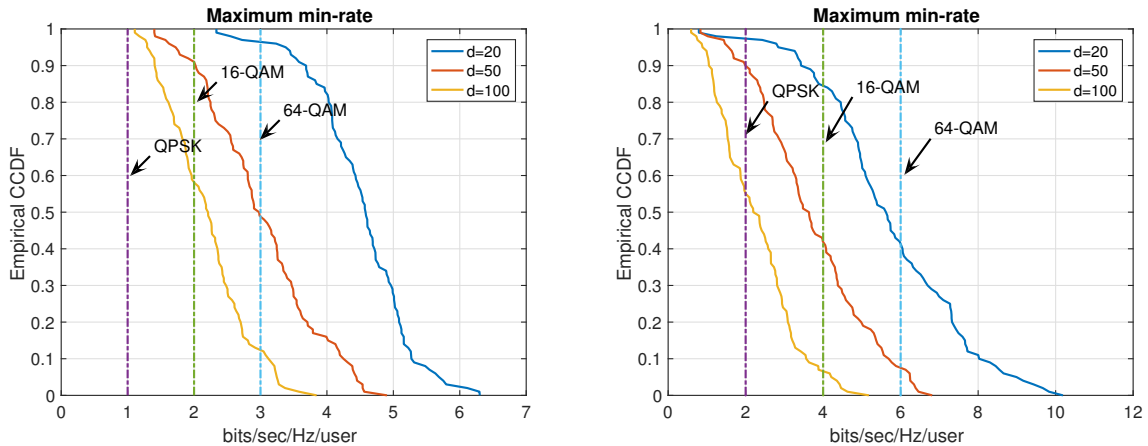
Figure 3.8: Empirical CCDF of the maximum min-rate for users in a picocell in the middle of an urban canyon for (a) K=2, Q=4 and (b) K=4, Q=4 .

Figure 3.8 shows the empirical CCDF of the maximum min-rate provided for each of the users. Note that the hardware saturation points corresponding to QPSK, 16-QAM and 64-QAM modulations are $(K/Q)\, s_M =$ 1, 2 and 3 respectively for the case K=2 and Q=4.

3.4.2   Capacity calculations

We now are ready to estimate the attainable capacity gain over the conventional LTE networks. The downlink capacity of LTE network is estimated as 0.6 Gbps/km^2 over a total bandwidth of 255 MHz in [41]. However, the available bandwidth for downlink cellular networks is 500 MHz, hence the total capacity could be further increased by adding more channels per base station. Therefore, we estimate the total downlink capacity of LTE networks as 1.2 Gbps/km^2. For the sake of comparison, we consider a one square kilometer region in Manhattan area (Figure 3.9), which encompasses 15 urban canyons. Thus, we can get a rough estimate of the overall capacity per square kilometer of our approach via the following computations:

$$\text{Capacity}\,(bps/km^2) = \text{Maximum min-rate}(bps/Hz/user)\times \qquad (3.6)$$
$$\frac{B}{F}(Hz) \times 2Q\ (\text{Num. users}/cell) \times n_c\ (\text{Num. cells}/km^2)$$

where B, F and n_c are the total bandwidth, the frequency reuse factor and the number of picocells per square kilometer respectively. Note that $2Q$ in (3.6) refers to the number of users served within the picocell [1] which are covered by either eastward facing antennas of BS_0 or westward facing antennas of BS_1. In our example of a 1km^2 region in Manhattan shown in Figure 3.9, there are a total of 15 street canyons of length 1km (in both directions), each of which encompasses 1km/d cells. Hence, we get $n_c \approx 150,300$ and 750 for picocell widths of d=100,50 and 20 meters respectively.

We have summarized the above results in table 3.1 specifying the overall attainable capacity for different scenarios. Note that the Maximum-min-rate in equation 3.6 is

---

[1]This requires $2Q \times n_c$ users/km^2 = 9000 users/km^2 in our most extreme case: d=20m and Q=6 which is still much smaller than the population density of Manhattan area: 27,826 persons/km^2 [42].
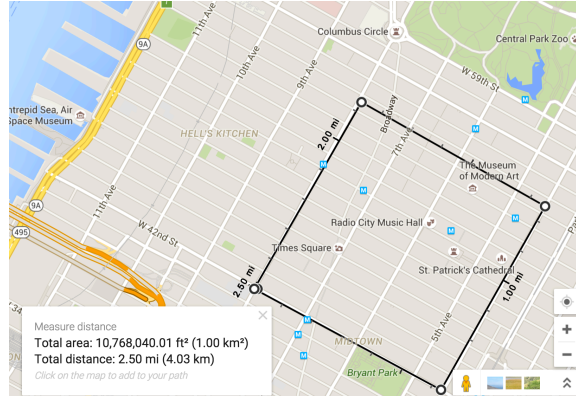
Figure 3.9: 1 km^2 in Manhattan area, encompassing 15 street canyons.

replaced with the average of datarates attained in 3.8 after quantizing to the hardware limited values of $(K/Q)\,s_M$.

Table 3.1: Capacity (Tbps/km^2) over a total bandwidth of 2GHz for a rural area in New York employing 8×8 and 4×4 antenna arrays as transmitter and receivers.

Capacity (Tbps/km^2)	$K=1$	$K=2$	$K=1$	$K=2$	$K=4$
	$F=1$		$F=2$		
d=100 m	1.3	1.8	1.6	2.6	2.7
d=50 m	5.3	8.9	3.3	6.4	8.9
d=20 m	17.6	33.1	8.9	17.9	30.9

Smaller picocells (d=20) are less prone to interference since the antenna beams aiming their target users, are slanted more steeply and hence will illuminate (induce interference to) an smaller area around them. This feature, along with the increased spatial reuse attained with smaller cell sizes, leads to massive estimated capacity of 30.9 Tbps/km^2. Users in larger picocells experience higher interference levels, and hence smaller constellations are employed to serve them. This still suffices for achieving capacity as large as

49

2.7 Tbps/km^2.

Note that, users in larger picocells are not gaining much from more subarrays per face (Table 3.1). This is because almost horizontally aligned beams in wide cells, lead to much closer spatial frequencies for which our interference suppression algorithm is not as effective. Possible approaches to solve this problem are (a) increasing the number of antenna elements which provides more degrees of freedom for employing the interference suppression algorithm or (b) increasing basestation height which will draw users away in the spatial frequency domain.

Table 3.2: Comparing convention LTE and mm-wave cellular networks

	LTE	mm-wave		Gain
		d=20	d=100	
Capacity	1.2Gbps	30.9 Tbps	2.7 Tbps	$\geq$ 2250X
Bandwidth	500 MHz	2GHz		4X
Spatial reuse	–	–		$\geq$ 500X

Table 3.2 compares the resultant capacity for mm-wave picocells computed via simulations with the benchmark capacity of LTE networks. As shown below, the promised 1000-fold capacity increase is immediately reachable even with the largest picocell size (d=100m) considered here. Excluding the 4X gain from the larger bandwidth of 2GHz employed in our system (which is still a small fraction of the 14GHz of available bandwidth at 60GHz), the remaining gain ($\geq$ 500) is attained through the larger spatial reuse from small cells and pencil beams.

Chapter 4

Compressive channel estimation and tracking for large arrays in mm-wave picocells

4.1 Introduction

In this chapter, we propose and investigate a compressive architecture for estimation and tracking of sparse spatial channels in mm-wave picocellular networks. Since the base stations in our system employ radio frequency (RF) beamforming, so that standard least squares adaptation techniques (which require access to individual antenna elements) are not applicable. We focus on the downlink, and show that "compressive beacons," transmitted using pseudorandom phase settings at the base station array, and compressively processed using pseudorandom phase settings at the mobile array, provide information sufficient for accurate estimation of the two-dimensional (2D) spatial frequencies associated with the directions of departure of the dominant rays from the base station, and the associated complex gains. This compressive approach is compatible with coarse

51

phase-only control, and is based on a near-optimal sequential algorithm for frequency estimation which approaches the Cramér Rao Lower Bound. The algorithm exploits the geometric continuity of the channel across successive beaconing intervals to reduce the overhead to less than 1% even for very large (32×32) arrays. Compressive beaconing is essentially omnidirectional, and hence does not enjoy the SNR and spatial reuse benefits of beamforming obtained during data transmission. We therefore discuss system level design considerations for ensuring that the beacon SNR is sufficient for accurate channel estimation, and that inter-cell beacon interference is controlled by an appropriate reuse scheme.

4.1.1   Related Work

We assume that the channel is well described by a relatively small number of discrete rays with delays and angles of arrival/departure taking values in a continuum. The key contribution of this work is to provide a super-resolution framework for estimating and tracking these rays, with model-based estimation allowing us to go beyond (spatial) bandwidth based resolution limits. To the best of our knowledge, the existing literature on mm-wave measurements does not attempt to super-resolve channels in this fashion, hence we do not know, for example, whether the continuous power-delay profiles reported in [43] are consistent with a parsimonious channel model such as ours. However, preliminary experimental results [44], which use compressive measurements to successfully recover power-angle profiles for a controlled experiment (a small number of reflectors in an anechoic chamber), indicate that a simple ray-based model like ours may well suffice. Validating this assertion would require application of the more sophisticated compressive estimation techniques discussed here, as opposed to the standard basis pursuit algorithms employed in [44]. Note that existing models for simulation-based evaluations, such as the

channel models standardized for IEEE 802.11ad indoor 60 GHz channels, typically *assume* more complex models which are variants of the Saleh-Valenzuela model, with a number of clusters, each consisting of multiple closely spaced rays. Such cluster-based models could be motivated by the roughness of reflecting surfaces such as walls, but they have not been experimentally demonstrated. While these issues fall beyond the scope of the present work, which aims to make a fundamental contribution to signal processing for sparse spatial channels, a sustained effort in measurement-based validation of our model and approach is a critically important topic for future work.

4.2   System Model

We focus our modeling and performance evaluation on the urban canyon setting depicted in Figure 1.2. We consider mm-wave transmission on the downlink (for beaconing and downlink data) and LTE or WiFi at lower carrier frequencies on the uplink (for feedback and uplink data). In terms of channel estimation and tracking, this could be viewed as a worst-case assumption, since two-way transmission on the same mm-wave band could potentially be exploited using channel reciprocity. For simplicity, we consider a single antenna array for each face, used for both compressive beaconing and downlink data communication. Mobile stations are modeled as either pedestrians walking along sidewalks, or cars moving along the street.

In our simulations, we model $K = 4$ dominant paths from base station to mobile in our simulations: the line of sight (LoS), and the single bounce reflections from the ground or the side walls. Some of these paths may be blocked by obstacles (diffraction around obstacles is limited for small carrier wavelengths). We ignore multiple bounces, since such paths get attenuated significantly, especially because each bounce sees a smaller angle of incidence than a typical single bounce. However, our compressive estimation

algorithm does not use the preceding assumptions on number of dominant paths as prior information, and automatically discovers and tracks paths.

For a regular d-spaced square $N_{1D} \times N_{1D}$ antenna array and a point transmitter in the far-field, the channel seen by the array is a mixture of 2D sinusoids, each corresponding to a propagation path, and is given by

$$h_{m,n} = \sum_{k=1}^{K} g_k e^{j\left(\omega_{x,k}m + \omega_{z,k}n\right)}, \ g_k \in \mathbb{C}, \ 1 \leq m,n, \leq N_{1D},$$

where g_k is the propagation gain of the k-th path, $\boldsymbol{\omega}_k = (\omega_{x,k}, \omega_{z,k})$ are the spatial frequencies corresponding to the k-th path (w.l.o.g. we have assumed that the square array is placed in the x-z plane, with its sides aligned to the coordinate axes) and $h_{m,n}$ refers the channel seen by the m,n-th antenna element,. The spatial frequencies of the k-th path are given by $\omega_{x,k} = 2\pi(d/\lambda)\sin\theta_k\cos\phi_k$ and $\omega_{z,k} = 2\pi(d/\lambda)\sin\theta_k\sin\phi_k$, where d denotes the array spacing, λ the carrier wavelength and (θ_k, ϕ_k) the inclination and azimuthal angles of the k-th path relative to $x - z$ plane. We vectorize the 2D sinusoid $\left[e^{j(\omega_x m + \omega_z n)}, \ 0 \leq m, n \leq N_{1D} - 1\right]$ and denote the resulting N_{1D}^2 long vector by $\mathbf{x}(N_{1D}, \boldsymbol{\omega})$, where $\boldsymbol{\omega} = (\omega_x, \omega_z)$ is the frequency of the 2D sinusoid. Vectorizing $[h_{m,n}, \ 1 \leq m, n \leq N_{1D}]$ in an identical manner gives us

$$\mathbf{h} = \sum_{k=1}^{K} g_k \mathbf{x}(N_{1D}, \boldsymbol{\omega}_k).$$

Now, consider a base station transmitter with a regularly spaced 2D array of size $N_{t,1D} \times N_{t,1D}$, and a mobile receiver with a regular 2D antenna array of size $N_{r,1D} \times N_{r,1D}$. Let $\mathbf{H}$ denote the corresponding $N_{t,1D}^2 \times N_{r,1D}^2$ channel matrix. Denoting by $\mathbf{h}_i$ the ith row of this matrix, $\mathbf{h}_i^T$ is the response of the receive antenna array to the ith transmit antenna. Denoting $\mathbf{x}(N_{t,1D}, \boldsymbol{\omega})$ by $\mathbf{x}_t(\boldsymbol{\omega})$ and $\mathbf{x}(N_{r,1D}, \boldsymbol{\omega})$ by $\mathbf{x}_r(\boldsymbol{\omega})$, under the far-field

assumption, it can be shown that

$$\mathbf{H} = \sum_{k=1}^{K} g_k \mathbf{x}_t \left(\boldsymbol{\omega}_k^t \right) \mathbf{x}_r^T \left(\boldsymbol{\omega}_k^r \right), \ g_k \in \mathbb{C}. \tag{4.1}$$

Since we know the array geometries (in this case, a regularly spaced 2D array), an estimate of the $N_{t,1D}^2 \times N_{r,1D}^2$ MIMO channel matrix $\mathbf{H}$ can be efficiently arrived at by estimating the spatial frequencies and the associated gains: $\{(g_k, \boldsymbol{\omega}_k^t, \boldsymbol{\omega}_k^r), \ k = 1, \ldots, K\}$. Such a parametric approach is far more efficient that direct estimation of individual entries of $\mathbf{H}$, and enables us to drastically reduce the number of measurements required.

4.3 Compressive channel estimation

We now describe our compressive approach for spatial channel estimation, which consists of a channel sounding strategy and an estimation algorithm which allows a base station to estimate the propagation gains $\{|g_k|\}$ and the spatial frequencies $\{\boldsymbol{\omega}_k^t\}$ in parallel for *all* mobiles in the picocell.

4.3.1 Channel sounding

The basestation sounds the channel using M compressive beacons. Each beacon is a known signal sent using a different set of transmit weights. The weights are chosen uniformly and independently at random from a small set of coarse phase shifts (for e.g. from the set $\{\pm 1, \pm j\}$, where $j = \sqrt{-1}$). We may therefore view each beacon as being transmitted from a different "virtual antenna" with a quasi-omnidirectional pattern. Each of the M transmit beacons are repeated L times by the basestation (see Figure 1.4). For each of these M transmit beacons, a mobile receiver employs L "virtual antennas" to measure the channel response, using receive weights chosen uniformly at

random from $\{\pm 1, \pm j\}$. Let $y(m, n)$ denote the response at the (m, n)th receive element due to a given transmit beacon. Letting $b(m, n, l) \in \{\pm 1, \pm j\}$ denote the weight for (m, n)th receive element for the lth virtual receive antenna, the response seen by the lth virtual receive antenna is given by

$$r(l) = \sum_{1 \leq m, n \leq N_{r,1D}} b(m, n, l) \times y(m, n), \ 1 \leq l \leq L. \tag{4.2}$$

These measurements are used to construct the $M \times L$ Multiple Input Multiple Output (MIMO) "virtual channel" matrix $\mathbf{V}$ between the virtual transmit and receive antennas. Note that we do not require that the base station know the receive weights used by the mobile, or that the mobile know the transmit weights used by the base station.

Denoting the vectorized version of weights of the i-th virtual transmit antenna by $\mathbf{a}_i$ (a vector of shape $N_{t,1D}^2 \times 1$) and that of the j-th virtual receive antenna by $\mathbf{b}_j$ ($N_{r,1D}^2 \times 1$), the i, j-th element of $\mathbf{V}$ (the channel between the (i, j)-th virtual transmit-receive pair) is given by $v_{i,j} = \mathbf{a}_i^T \mathbf{H} \mathbf{b}_j$. Letting $\boldsymbol{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_M]^T$ and $\boldsymbol{B} = [\mathbf{b}_1 \ \dots \ \mathbf{b}_L]^T$, it is easy to see that $\mathbf{V} = \boldsymbol{A}\mathbf{H}\boldsymbol{B}^T$. Using (4.1), we have that

$$\mathbf{V} = \sum_{k=1}^{K} g_k \left(\boldsymbol{A}\mathbf{x}_t \left(\boldsymbol{\omega}_k^t \right) \right) \left(\boldsymbol{B}\mathbf{x}_r \left(\boldsymbol{\omega}_k^r \right) \right)^T. \tag{4.3}$$

The channel measurements are perturbed by i.i.d measurement noise, and are given by

$$y_{i,j} = \sqrt{P_e} v_{i,j} + z_{i,j}, \ z_{i,j} \sim \mathcal{CN}(0, \sigma^2),$$

where P_e is the per-element transmit power. Letting $\mathbf{Y}$ and $\mathbf{Z}$ denote $M \times L$ matrices with their i, j-th entries given by $y_{i,j}$ and $z_{i,j}$ respectively, the "measured virtual channel"

is given by

$$\mathbf{Y} = \sqrt{P_e}\mathbf{V} + \mathbf{Z} = \sqrt{P_e}\sum_{k=1}^{K} g_k \left(\boldsymbol{A}\mathbf{x}_t\left(\boldsymbol{\omega}_k^t\right)\right)\left(\boldsymbol{B}\mathbf{x}_r\left(\boldsymbol{\omega}_k^r\right)\right)^T + \mathbf{Z}. \tag{4.4}$$

4.3.2   Feedback strategies

Our goal is to track the mm-wave spatial channel as seen from the basestation, described by the parameters $P = \{(|g_k|, \boldsymbol{\omega}_k^t), k = 1, \ldots, K\}$. To this end, every mobile in the picocell needs to feed back a portion of the measured virtual channel $\mathbf{Y}$ to the basestation. From (4.4), we see that the information regarding the spatial channel as seen from the basestation, given by P, is available in the *column space* of $\mathbf{Y}$. Building on this observation, we propose two feedback strategies:

 (i) The receiver feeds back the entire matrix $\mathbf{Y}$.

 (ii) A more elaborate strategy involves feeding back $Q \leq L$ strongest left singular vectors of $\mathbf{Y}$ scaled by their corresponding singular values. i.e., if $\mathbf{Y} = \sum_{l=1}^{L} \sigma_l \mathbf{u}_l \mathbf{v}_l^H$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_L \geq 0$, the receivers feed back $\mathbf{D} \equiv [\sigma_1 \mathbf{u}_1 \ \cdots \ \sigma_Q \mathbf{u}_Q]$. This strategy identifies the Q-dimensional *subspace* of the column space of $\mathbf{Y}$ with maximum energy.

4.4   Estimation Algorithm

We now present an algorithm to estimate the parameters $\{(|g_k|, \boldsymbol{\omega}_k^t), k = 1, \ldots, K\}$ characterizing the mm-wave channel as seen from the basestation. The same algorithm applies for both forms of feedback discussed in Section 4.3.2: the entire measured virtual MIMO matrix $\mathbf{Y}$ or the dominant weighted left singular vectors $\mathbf{D}$.

The lth column of $\mathbf{Y}$ is given by

$$\mathbf{y}_l = \sum_{k=1}^{K} h_{k,l} \boldsymbol{A}\mathbf{x}_t \left(\boldsymbol{\omega}_k^t\right) + \mathbf{z}_l, \ l = 1, \ldots, L \tag{4.5}$$

where $\mathbf{z}_l \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbb{I}_M\right)$ denotes the l-th column of $\mathbf{Z}$ and $h_{k,l} = \sqrt{P_e} g_k \mathbf{b}_l^T \mathbf{x}_r \left(\boldsymbol{\omega}_k^r\right)$. We assume the weight sequence $\{\mathbf{b}_l\}$ used to construct the receive virtual antennas at the receive antenna array is not available at the transmitter, and hence cannot jointly estimate $\boldsymbol{\omega}_k^t$ and $\boldsymbol{\omega}_k^r$. However, since $\{\mathbf{b}_l, \ l = 1, \ldots, L\}$ were picked in an i.i.d manner, we have that $\{h_{k,l}, \ l = 1, \ldots, L\}$ are i.i.d realizations of a random variable with $\mathbb{E}|h_{k,l}|^2 = P_e \mathbb{E}\left|g_k \mathbf{b}_l^T \mathbf{x}_r\left(\boldsymbol{\omega}_k^r\right)\right|^2 = P_e N_{r,1D}^2 |g_k|^2$. This allows us to estimate $|g_k|^2$ as follows:

$$P_e |\hat{g}_k|^2 = \left(1/\left(LN_{r,1D}^2\right)\right) \sum_{l=1}^{L} \left|\hat{h}_{k,l}\right|^2.$$

From here on, in Section 4.4, we use the notation $\boldsymbol{\omega}_k$ to refer to $\boldsymbol{\omega}_k^t$ and $\mathbf{x}(\boldsymbol{\omega}_k)$ to refer to $\boldsymbol{A}\mathbf{x}_t(\boldsymbol{\omega}_k^t)$. Thus, the measurements can be written as

$$\mathbf{y}_l = \sum_{k=1}^{K} h_{k,l} \mathbf{x}\left(\boldsymbol{\omega}_k\right) + \mathbf{z}_l, \ l = 1, \ldots, L. \tag{4.6}$$

We now provide an algorithm to estimate $\{\boldsymbol{\omega}_k, \{h_{k,l}\}\}$.

4.4.1   Single path

We first present an algorithm for estimating a single path $K = 1$, which forms the basis for our sequential estimation algorithm for $K > 1$. Omitting the path index k in (4.6), we have

$$\mathbf{y}_l = h_l \mathbf{x}\left(\boldsymbol{\omega}\right) + \mathbf{z}_l, \ l = 1, \ldots, L.$$

Given that $\{\mathbf{z}_l\}$ are independent realizations of $\mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbb{I}_M\right)$, the maximum likelihood (ML) estimates of $\boldsymbol{\omega}, \{h_l\}$ are given by:

$$\hat{\boldsymbol{\omega}}, \left\{\hat{h}_l\right\} = \arg \min_{\boldsymbol{\omega}, \{h_l\}} \sum_{l=1}^{L} \left\|\mathbf{y}_l - h_l \mathbf{x}\left(\boldsymbol{\omega}\right)\right\|^2 . \tag{4.7}$$

For any $\boldsymbol{\omega}$, the optimal h_l-s are given by least-squares estimates:

$$h_l^\star\left(\boldsymbol{\omega}\right) = \left\langle \mathbf{x}\left(\boldsymbol{\omega}\right), \mathbf{y}_l\right\rangle / \left\|\mathbf{x}\left(\boldsymbol{\omega}\right)\right\|^2 \tag{4.8}$$

where $\langle x, y\rangle$ denotes $x^H y$. Plugging into (4.7), the ML estimate of $\boldsymbol{\omega}$ is given by:

$$\hat{\boldsymbol{\omega}} = \arg \max_{\boldsymbol{\omega}} \ \frac{1}{\left\|\mathbf{x}\left(\boldsymbol{\omega}\right)\right\|^2} \sum_{l=1}^{L} \left|\left\langle \mathbf{x}\left(\boldsymbol{\omega}\right), \mathbf{y}_l\right\rangle\right|^2 \tag{4.9}$$

and the ML estimate of h_l is given by $h_l^\star\left(\hat{\boldsymbol{\omega}}\right)$. We employ a two-step algorithm to arrive the ML estimates: a "detection" phase followed by a "refinement" phase.

Detection phase: Using M two-dimensional FFT computations, we precompute $\mathbf{x}(\boldsymbol{\omega})$ for frequencies of the form $\boldsymbol{\omega} \in \Phi \equiv \left\{(2\pi i/T, 2\pi j/T), \ 0 \leq i, j \leq T-1\right\}, \ T = RN_{1D,t},$ where R is the oversampling factor. We pick the frequency $\hat{\boldsymbol{\omega}}$ from Φ which maximizes (4.9). The corresponding gains are given by $\hat{h}_l = h_l^\star\left(\hat{\boldsymbol{\omega}}\right)$. We remove the contribution of the newly detected sinusoid from the measured channel response and this residual measurement is given by

$$\boldsymbol{r}_l = \mathbf{y}_l - \hat{h}_l \mathbf{x}\left(\hat{\boldsymbol{\omega}}\right) . \tag{4.10}$$

(This residue is used for sequential detection for $K > 1$, as discussed shortly.)

Refinement phase: Our estimate from the detection phase is restricted to the discrete set Φ and consequently we do not expect $\hat{\boldsymbol{\omega}}$ to be equal to the ML estimate given by (4.9) (where the maximization is over $[-2\pi d/\lambda, 2\pi d/\lambda]^2$ with d being the spacing between

transmitter antennas). However, if we make the grid fine enough, the best estimate of $\boldsymbol{\omega}$ in Φ is expected to be close enough to the optimal solution to allow refinement via local optimization. In order to do this, we first *fix* the gain estimates $\{\hat{h}_l\}$ and refine only the estimate of the spatial frequency $\boldsymbol{\omega}$ by seeking the minimizer of the ML cost function

$$C(\boldsymbol{\omega}) = \sum_{l=1}^{L} \left\| \mathbf{y}_l - \hat{h}_l \mathbf{x}(\boldsymbol{\omega}) \right\|^2$$

in the neighborhood of the current estimate $\hat{\boldsymbol{\omega}}$ using the Newton method. This involves evaluating the gradient vector $\mathcal{G}(\boldsymbol{\omega})$ and the Hessian matrix $\mathcal{H}(\boldsymbol{\omega})$ of $C(\boldsymbol{\omega})$ at the current estimate $\hat{\boldsymbol{\omega}}$. The corresponding expressions are given by:

$$\mathcal{G}_i(\boldsymbol{\omega}) = \frac{\partial C(\boldsymbol{\omega})}{\partial \omega_i} = -2 \sum_{l=1}^{L} \Re \left\{ \left\langle \boldsymbol{r}_l, \hat{h}_l \frac{\partial \mathbf{x}(\boldsymbol{\omega})}{\partial \omega_i} \right\rangle \right\},$$

$$\mathcal{H}_{ij}(\boldsymbol{\omega}) = \frac{\partial^2 C(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} = -2 \sum_{l=1}^{L} \Re \left\{ \left\langle \boldsymbol{r}_l, \hat{h}_l \frac{\partial^2 \mathbf{x}(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} \right\rangle - \right.$$
$$\left. \left| \hat{h}_l \right|^2 \left\langle \frac{\partial \mathbf{x}(\boldsymbol{\omega})}{\partial \omega_i}, \frac{\partial \mathbf{x}(\boldsymbol{\omega})}{\partial \omega_j} \right\rangle \right\}, \ 1 \le i, j \le 2$$

where $\boldsymbol{\omega} = [\omega_1, \omega_2]$. $\{\boldsymbol{r}_l\}$-s are the residual measurements given by (4.10). The Newton update for $\hat{\boldsymbol{\omega}}$ is

$$\hat{\boldsymbol{\omega}} \leftarrow \hat{\boldsymbol{\omega}} - \mathcal{H}^{-1}(\hat{\boldsymbol{\omega}}) \mathcal{G}(\hat{\boldsymbol{\omega}}). \tag{4.11}$$

We follow this up by updating our estimates $\left\{ \hat{h}_l \right\}$ by plugging the new value of $\hat{\boldsymbol{\omega}}$ in (4.8), i.e.,

$$\hat{h}_l \leftarrow h_l^\star(\hat{\boldsymbol{\omega}}) = \langle \mathbf{x}(\hat{\boldsymbol{\omega}}), \mathbf{y}_l \rangle / \|\mathbf{x}(\hat{\boldsymbol{\omega}})\|^2 \tag{4.12}$$

and modifying the residues ($\{\boldsymbol{r}_l\}$) accordingly using (4.10). The algorithm alternates between the updates in (4.11) and (4.12) for a few iterations.

4.4.2   Multiple paths

We now build on the preceding single path algorithm for the general setting of $K \geq 1$. Suppose that our current estimate of the sinusoids/paths is given by $\mathcal{P}_K = \left\{ \left(\hat{\boldsymbol{\omega}}_k, \{\hat{h}_{k,l}\} \right), \ k = 1, \ldots, K \right\}$. The residual measurements corresponding to a set of estimated parameters $\mathcal{P}$ is given by:

$$\mathbf{v}_l(\mathcal{P}) = \mathbf{y}_l - \sum\nolimits_{\boldsymbol{\omega}_k, \{h_{k,l}\} \in \mathcal{P}} h_{k,l} \mathbf{x}(\boldsymbol{\omega}_k).$$

Detect a new path: Assuming that the measurements $\mathbf{y}_l$ are given by the current residue $\mathbf{v}_l(\mathcal{P}_q)$ (corresponding to the q detected paths), we use the single path algorithm in Section 4.4.1 to detect and refine a new sinusoid $(\hat{\boldsymbol{\omega}}_{q+1}, \{\hat{h}_{q+1,l}\})$. Let $\mathcal{P}_{q+1}$ denote the new set of estimated parameters $\mathcal{P}_q \cup \{(\hat{\boldsymbol{\omega}}_{q+1}, \{\hat{h}_{q+1,l}\})\}$.

Refine existing paths: Once we add this new path, we refine the parameters of *all* $q+1$ sinusoids in $\mathcal{P}_{q+1}$ one by one. Consider the parameters $(\hat{\boldsymbol{\omega}}_k, \{\hat{h}_{k,l}\})$ of the k-th sinusoid. We use the refinement algorithm in Section 4.4.1 to refine $(\hat{\boldsymbol{\omega}}_k, \{\hat{h}_{k,l}\})$ by assuming that the measurements $\mathbf{y}_l$ are given by the residual measurements after excluding the sinusoid of interest. i.e, $\mathbf{v}_l(\mathcal{P}_{q+1} \setminus \{(\hat{\boldsymbol{\omega}}_k, \{\hat{h}_{k,l}\})\})$. Sinusoids are refined in a round robin manner, and the process is repeated for a few rounds: $1 \to 2 \to \cdots \to (q+1) \to 1 \to \cdots \to (q+1)$.

Stopping criterion: The algorithm terminates when the residues $\{\mathbf{v}_l(\mathcal{P}_q)\}$ after refinement satisfy:

$$\max_{\boldsymbol{\omega} \in \mathrm{DFT}} \ \frac{1}{\|\mathbf{x}(\boldsymbol{\omega})\|^2} \sum_{l=1}^{L} |\langle \mathbf{x}(\boldsymbol{\omega}), \mathbf{v}_l(\mathcal{P}_q) \rangle|^2 \leq \tau, \tag{4.13}$$

with the threshold τ given by $\sigma^2 \gamma^{-1} \left(L, \Gamma(L)(1 - P_{\mathrm{fa}})^{1/N_{1D,t}^2} \right)$ where $\gamma^{-1}(S, y)$ is the inverse of incomplete gamma function $\gamma(S, x)$ with respect to integral limit x and DFT $= \{(2\pi i/N_{1D,t}, 2\pi j/N_{1D,t}) \ : \ 0 \leq i, j \leq N_{1D,t} - 1\}$. This choice of τ corresponds to a Constant False Alarm Rate (CFAR) test (the false alarm rate being P_{fa}) for whether the

residual measurements $\{\mathbf{v}_l(\mathcal{P}_q)\}$ can be explained by another path. In arriving at this expression for τ we make two assumptions: (i) channel sounding is not compressive: i.e, $\boldsymbol{A} = \mathbb{I}_{N_{1D,t}^2}$ and (ii) new paths correspond to specific beam orientations given by $\boldsymbol{\omega} \in \text{DFT}$. However, our simulation results show that, this is a good approximation for our case with compressive measurements and continuous values of beam directions. We refer the reader to Appendix A for details on the stopping criterion.

4.4.3   Tracking

We sound the channel often enough so that between any two successive channel estimation cycles, the *geometry* of the mm-wave channel, given by the spatial frequencies $\{\boldsymbol{\omega}_k\}$ of the paths, does not change "significantly," even if the path gains $\{g_k\}$ do. This ensures that angle of departure estimates from the prior sounding round do not become stale over the course of the communication phase during which they are needed for beamforming purposes. For example, if we do not wish to tolerate a beamforming loss of 3dB or more, then our estimate from the previous round $\hat{\boldsymbol{\omega}}$ should be close enough to the current $\boldsymbol{\omega}$ so that

$$|\langle \mathbf{x}(\hat{\boldsymbol{\omega}}), \mathbf{x}(\boldsymbol{\omega}) \rangle|^2 / \|\mathbf{x}(\boldsymbol{\omega})\|^2 > 0.5$$

over the entire communication phase. This condition is met if $\|\boldsymbol{\omega} - \hat{\boldsymbol{\omega}}\|_\infty < 0.5 \times (2\pi/N_{t,1D})$. Therefore, the estimates of spatial frequencies from the previous sounding round are good approximations of their current true value (within a DFT spacing of $\frac{2\pi}{N_{t,1D}}$). We exploit this by using $\{\hat{\boldsymbol{\omega}}_k, \; k = 1, \ldots, K\}$ from the prior round to initialize our algorithm (as opposed to using the empty set $\{\}$). We do this by constructing the matrix $\mathbf{X} = [\mathbf{x}(\hat{\boldsymbol{\omega}}_1) \; \ldots \; \mathbf{x}(\hat{\boldsymbol{\omega}}_K)]$ and setting $\hat{h}_{k,l}$ to be the (k,l)-th entries of $(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{Y}$, where $\mathbf{Y} = [\mathbf{y}_1 \; \ldots \; \mathbf{y}_L]$. We refine all parameters in $\mathcal{P}_K = \{(\hat{\boldsymbol{\omega}}_k, \{\hat{h}_{k,l}\}), \; k = 1, \ldots, K\}$ using the refinement algorithm in Section 4.4.2 before proceeding to seek for new paths

using the algorithm in Section 4.4.2.

Deleting weak paths: Paths estimated in prior rounds may not be viable at the current time instant (e.g, because of blockage). Therefore, we need means to remove such stale paths. We use the stopping criterion (4.13) as a means to delete weak paths. If deleting the path under question (say k) and optimizing other parameters $\mathcal{P}_{k,\mathrm{opt}} \leftarrow \mathrm{Refine}\left(\mathcal{P} \setminus (\hat{\boldsymbol{\omega}}_k, \{\hat{h}_{k,l}\})\right)$ using our refinement algorithm results in residual measurements $\{\mathbf{v}_l(\mathcal{P}_{k,\mathrm{opt}})\}$ that can be explained by noise (i.e., satisfies (4.13)), we delete the path permanently. Otherwise we keep the path.

4.5 Protocol Parameter Choices

In this section, we give a principled approach to choosing parameters of the compressive channel estimation protocol, namely the number of unique transmit beacons M, the number of receive measurement weights L and the minimum effective Signal to Noise Ratio (SNR) needed for channel estimation, which we use to choose the sounding bandwidth W_s. We then turn to the question of how frequently the channel has to be sounded. In Section 4.7, we take two scenarios and apply this recipe to arrive at system level parameters for the protocol.

4.5.1 Number of compressive transmit beacons

Classical compressive sensing aims to reconstruct signals which are sparse in a discrete basis, based on a small number of projections. In order for reconstruction to be successful, these projections must preserve the underlying geometry [45]. In [34], these ideas are extended to compressive estimation, in which a small number of projections are used to estimate continuous-valued parameters.

We now translate these concepts to our present context. Our goal is to estimate the

spatial frequencies $\{\boldsymbol{\omega}_k^t\}$ from measurements of the form

$$\mathbf{y}_l = \sum_{k=1}^{K} h_{k,l} \boldsymbol{A} \mathbf{x}_t \left(\boldsymbol{\omega}_k^t\right) + \mathbf{z}_l, \ l = 1, \ldots, L.$$

The algorithm in Section 4.4 aims to estimate parameters $\{h_{k,l}, \boldsymbol{\omega}_k\}$ by minimizing the ML cost function:

$$\sum_{l=1}^{L} \left\| \mathbf{y}_l - \sum_{k=1}^{K} \hat{h}_{k,l} \boldsymbol{A} \mathbf{x}_t \left(\hat{\boldsymbol{\omega}}_k\right) \right\|^2$$
$$= \sum_{l=1}^{L} \left\| \boldsymbol{A} \times \sum_{k=1}^{K} \left(h_{k,l} \mathbf{x}_t \left(\boldsymbol{\omega}_k^t\right) - \hat{h}_{k,l} \mathbf{x}_t \left(\hat{\boldsymbol{\omega}}_k^t\right)\right) + \mathbf{z}_l \right\|^2,$$

where $\left\{\hat{h}_{k,l}, \hat{\boldsymbol{\omega}}_k^t\right\}$ refer to our estimates of $\{h_{k,l}, \boldsymbol{\omega}_k^t\}$. If the compressive measurement matrix $\boldsymbol{A}$ ensures that

$$\left\| \boldsymbol{A} \times \sum_{k=1}^{K} \left(h_{k,l} \mathbf{x}_t \left(\boldsymbol{\omega}_k^t\right) - \hat{h}_{k,l} \mathbf{x}_t \left(\hat{\boldsymbol{\omega}}_k^t\right)\right) \right\|^2 \tag{4.14}$$
$$\approx M \left\| \sum_{k=1}^{K} \left(h_{k,l} \mathbf{x}_t \left(\boldsymbol{\omega}_k^t\right) - \hat{h}_{k,l} \mathbf{x}_t \left(\hat{\boldsymbol{\omega}}_k^t\right)\right) \right\|^2, \ \forall h_{k,l}, \hat{h}_{k,l},$$

for relevant $(\{\boldsymbol{\omega}_k\}, \{\hat{\boldsymbol{\omega}}_k\})$-pairs, the cost structure of the estimation problem is roughly preserved. Therefore, estimation using compressive measurements is similar to estimation with all $N_{t,1D}^2$ measurements (except for a reduction in effective SNR, given by $M/N_{t,1D}^2$) [34].

We now state an isometry property relevant for our purpose. For some fixed ϵ, the matrix $\Phi \in \mathbb{C}^{m \times p}$ is said to enjoy the s-isometry property for the basis B (of size $p \times n$) if there exists a constant $C > 0$ such that

$$C(1 - \epsilon) \leq \|\Phi B \mathbf{u}\|^2 / \|B \mathbf{u}\|^2 \leq C(1 + \epsilon),$$

for all s-sparse $\mathbf{u}$ in $\mathbb{C}^n$. It can be shown using the celebrated Johnson-Lindenstrauss Lemma (JL Lemma) [46] (with arguments similar to those in [45]) that if $m = O(s\epsilon^{-2} \log n)$, a randomly picked Φ (with elements drawn from proper distributions, e.g., Uniform$\{\pm 1, \pm j\}$ [34, 46]) satisfies this isometry property with high probability.

In order to provide a concrete definition of geometry preservation in our setting, we consider a discretized version of the problem, when spatial frequencies are restricted to an oversampled DFT grid G of size $R = O(N_{t,1D}^2)$. The condition (4.14) reduces to a $2K$-*isometry property* of the measurement matrix $\boldsymbol{A}$ relative to the basis $\mathbf{X}$, where $\mathbf{X}$ is the $N_{t,1D}^2 \times R$ matrix with its columns given by $\{\mathbf{x}_t(\boldsymbol{\omega}) : \boldsymbol{\omega} \in G\}$. We can now apply the preceding isometry property to conclude that, when the number of unique transmitter beacons scales as $M = O\left(K\epsilon^{-2} \log R\right) = O\left(K\epsilon^{-2} \log N_{t,1D}\right)$, then the $2K$-pairwise isometry criterion w.r.t the basis $\mathbf{X}$ is met by the randomly picked sounding matrix $\boldsymbol{A}$ (it can be shown that $C = M$ for our choice of scale), thereby ensuring that the geometry of the spatial frequency estimation problem is preserved.

While the preceding calculations give order estimates for the number of measurements M required, in order to provide numerical values for our protocol design, we employ simulations. We consider the example of the 32×32 transmitter array and plot the maximum and minimum values of $(1/M) \left\| \boldsymbol{A}\mathbf{X}\mathbf{u} \right\|^2 / \left\| \mathbf{X}\mathbf{u} \right\|^2$ from 5×10^6 random realizations of a $2K = 8$-sparse $\mathbf{u}$ (which represent different realizations of $K = 4$ paths and their corresponding potential estimates) in Figure 4.1 using the 64-times oversampled DFT grid as the choice of basis $\mathbf{X}$. We see that this ratio is within $[-5, 3]$ dB when $M \geq 30$. This illustrates that for estimating $K = 4$ paths using a 32×32 array, measuring the response corresponding to $M = 30$ random beacons approximates the effect of measuring all $32 \times 32 = 1024$ antenna elements individually.
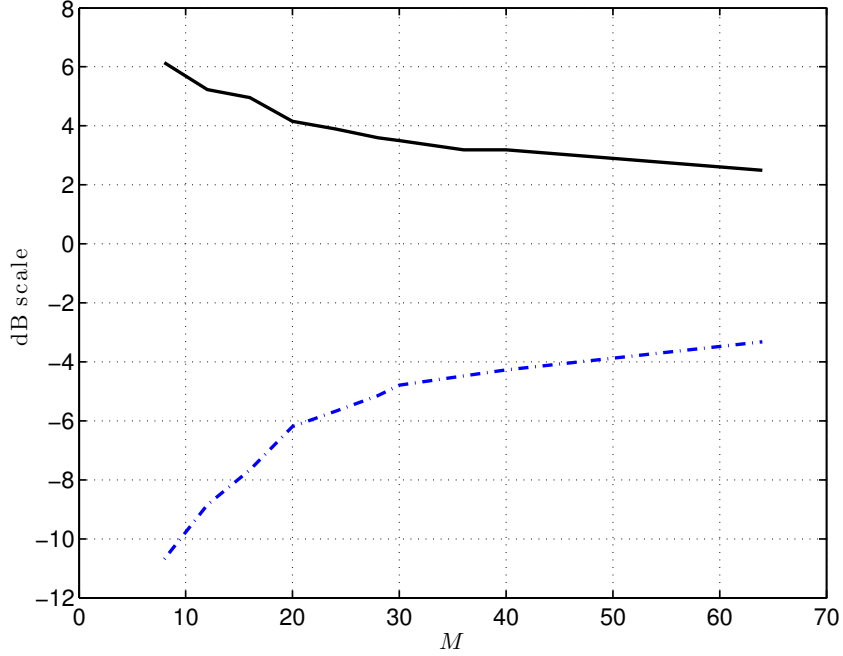
Figure 4.1: Maximum and minimum values of $\left\| \boldsymbol{A}\mathbf{X}\mathbf{u} \right\|^2 \big/ \left(M \left\| \mathbf{X}\mathbf{u} \right\|^2 \right)$ for different values of M, the number of transmitter beacons, across 5×10^6 random realizations of 8-sparse $\mathbf{u}$. The basis $\mathbf{X}$ corresponds to the responses for a 32×32 array evaluated uniformly over a $R = 64 \times 32^2$-sized grid

4.5.2   Number of compressive receive measurements

While we do not track $\{\boldsymbol{\omega}_k^r, \ k = 1, \ldots, K\}$, the spatial frequencies at the receiver, we need to ensure that the set of measurements made at the receiver have sufficient information to estimate transmitter spatial frequencies. Suppose that $\left\| \boldsymbol{B}\mathbf{x}_r \left(\boldsymbol{\omega}_k^r \right) \right\| \approx 0$, it follows from $h_{k,l} = g_k \sqrt{P_e} \mathbf{b}_l^T \mathbf{x}_r \left(\boldsymbol{\omega}_k^r \right)$ that all L measurements $\{\mathbf{y}_l, 1 \leq l \leq L\}$ will have very small contributions from the k-th path. i.e., $\left| h_{k,l} \right| \approx 0, 1 \leq l \leq L$. To see this observe that:

$$\sum_{l=1}^{L} \left| h_{k,l} \right|^2 = P_e \left| g_k \right|^2 \sum_{l=1}^{L} \left| \mathbf{b}_l^T \mathbf{x}_r \left(\boldsymbol{\omega}_k^r \right) \right|^2 = P_e \left| g_k \right|^2 \left\| \boldsymbol{B}\mathbf{x}_r \left(\boldsymbol{\omega}_k^r \right) \right\|^2 . \tag{4.15}$$

66
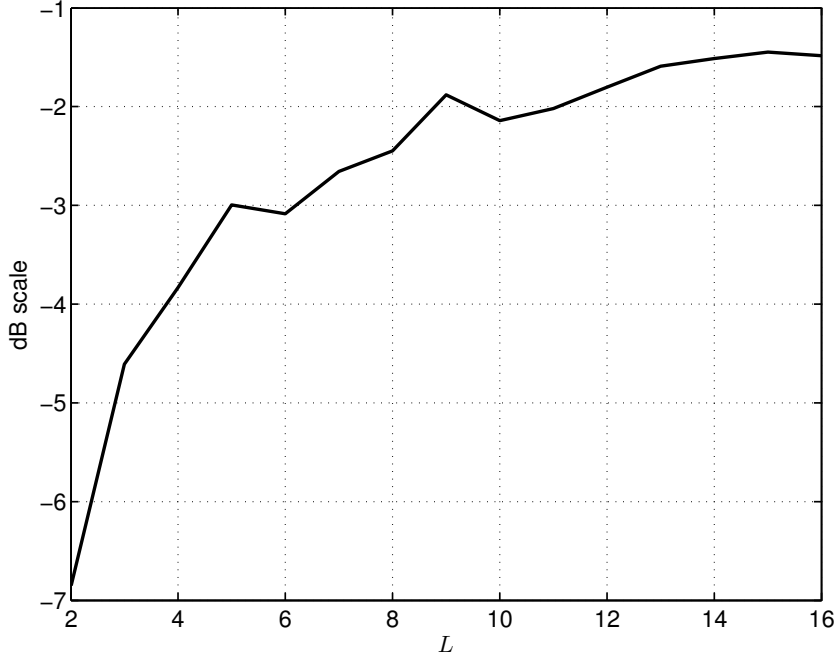
Figure 4.2: Maximum SNR degradation $\min_{\boldsymbol{\omega}} \|\boldsymbol{B}\mathbf{x}_r(\boldsymbol{\omega})\|^2 \big/ (L \|\mathbf{x}_r(\boldsymbol{\omega})\|^2)$ for the most favorable realization (from 10^4 runs) of an $L \times N^2_{r,1D}$ matrix $\boldsymbol{B}$ with $N_{r,1D} = 4$

Again, restricting the receive spatial frequencies to an oversampled DFT grid G of size $R = O(N^2_{r,1D})$, it can be shown that for $L = O\left(\log R\right) = O\left(\log N_{r,1D}\right)$, $\|\boldsymbol{B}\mathbf{x}_r\left(\boldsymbol{\omega}\right)\|^2 \approx L \|\mathbf{x}_r\left(\boldsymbol{\omega}\right)\|^2 = LN^2_{r,1D}$ w.h.p. This ensures that

$$\sum_{l=1}^{L} |h_{k,l}|^2 \approx P_e LN^2_{r,1D} |g_k|^2 \;\; \text{w.h.p,}$$

thereby capturing power along the k-th path. We perform computations for the maximum power lost across spatial frequencies when using a 4×4 array and plot the results in Figure 4.2. This shows that around 5 carefully chosen projections (we pick the best measurement matrix from 10^4 random instances) suffice to ensure that SNR degradation (relative to the nominal value of L) is no greater than 3dB for a 4×4 receive array.

4.5.3   SNR for successful estimation
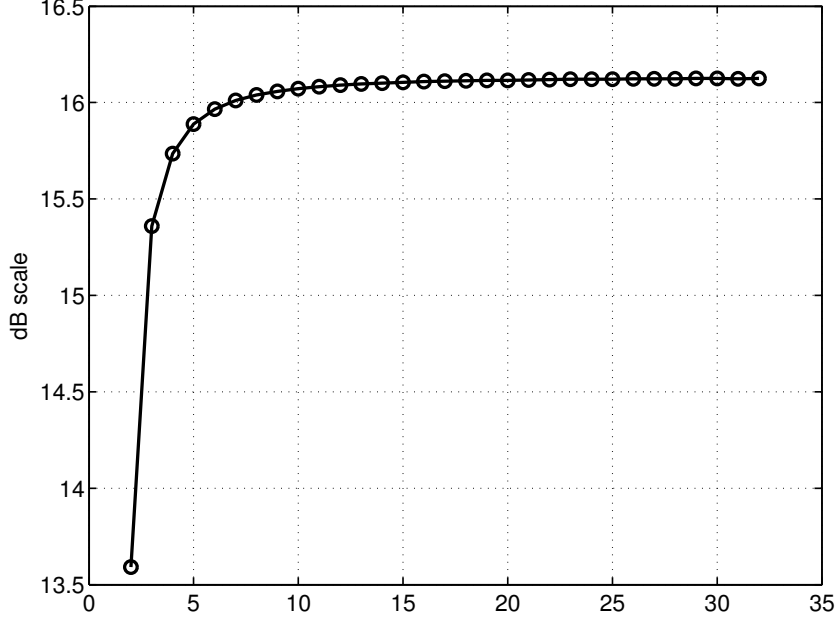
Figure 4.3: ZZB threshold SNR $\texttt{SNR}_{\text{th}}$ for estimating the frequency of a $N_{t,1D} \times N_{t,1D}$ sinusoid as a function of $N_{t,1D}$

The preceding criteria delineate the regime in which the *geometry* of the estimation problem is preserved approximately. We now turn to another factor which affects estimation performance, namely the SNR. Consider measurements of the form

$$y_{m,n} = e^{j(\omega_1 m + \omega_2 n + \phi)} + z_{m,n}, \ 0 \le m, n \le N_{t,1D} - 1, \tag{4.16}$$

where $z_{m,n}$ are i.i.d. $\mathcal{CN}(0, \sigma^2)$ and spatial frequencies ω_1, ω_2 and phase ϕ are parameters to be estimated. The Cramér Rao Bound[47] (CRB) for estimating ω_1 from measurements (4.16) is given by $C(\sigma^2) = 6 / \left(\texttt{SNR} \left(N_{t,1D}^2 - 1 \right) \right)$, where $\texttt{SNR} = \|\mathbf{x}_t(\boldsymbol{\omega})\|^2 / \sigma^2 = N_{t,1D}^2 / \sigma^2$ (same expression holds for ω_2). Assuming an uniform prior over $[0, 2\pi)^3$ for the parameters $(\omega_1, \omega_2, \phi)$, the Ziv Zakai Bound (ZZB) with periodic distortion[48] for estimating

ω_1 evaluates to:

$$Z(\text{SNR}) = \int_0^\pi Q\left(\sqrt{\text{SNR}\left(1 - \left|\frac{\sin\left(N_{t,1D}h/2\right)}{N_{t,1D}\sin\left(h/2\right)}\right|\right)}\right) h\mathrm{d}h.$$

An indicator of the SNR needed for successful estimation is the convergence of the ZZB to the CRB[34]. We use the SNR beyond which the ZZB is within 0.1dB of the CRB as a measure of this convergence. We plot this ZZB *threshold* SNR for different values of $N_{t,1D}$ in Figure 4.3. e.g, $\text{SNR}_{\text{th}} = 16.04\text{dB}$ for an 8×8 array and $\text{SNR}_{\text{th}} = 16.13\text{dB}$ for a 32×32 array.

The total energy E_{tot} corresponding to the k-th path collected across the ML measurements $\{\mathbf{y}_l, 1 \leq l \leq L\}$ is given by:

$$\text{E}_{\text{tot}} = \left\|\boldsymbol{A}\mathbf{x}_t\left(\boldsymbol{\omega}_k^t\right)\right\|^2 \times \sum_{l=1}^L |h_{k,l}|^2.$$

Using (4.15) in the above, we have that

$$\text{E}_{\text{tot}} = \left\|\boldsymbol{A}\mathbf{x}_t\left(\boldsymbol{\omega}_k^t\right)\right\|^2 \left\|\boldsymbol{B}\mathbf{x}_r\left(\boldsymbol{\omega}_k^r\right)\right\|^2 P_e |g_k|^2$$
$$\approx MLN_{t,1D}^2 N_{r,1D}^2 P_e |g_k|^2$$
$$= MLPN_{r,1D}^2 |g_k|^2,$$

where $P = N_{t,1D}^2 P_e$ is the total transmit power supplied to the $N_{t,1D} \times N_{t,1D}$ antenna array. The above approximation holds when M and L satisfy the preceding geometry preservation criteria in Sections 4.5.1 and 4.5.2 respectively. The effective SNR of the i-th path is given by $\text{SNR}_{\text{eff}} = \text{E}_{\text{tot}}/\sigma^2$. It is important to note that the per-measurement noise variance σ^2 is given by $\sigma^2 = N_{r,1D}^2\sigma_e^2$, where σ_e^2 is the noise variance per antenna element. Assuming no interference (which we account for in Section 4.6.4), $\sigma_e^2 = \text{N}_0 W_s$

with W_s denoting the sounding bandwidth and N_0 the thermal noise floor. The reason for the scale factor $N_{r,1D}^2$ in the expression for σ^2 is the following: Our measurement process consists of multiplying the received signal at each antenna (of which there are $N_{r,1D}^2$) by phasors $b(m,n,l) \in \{\pm 1, \pm j\}$ and *adding* the resultant signal (as per (4.2)). Since thermal noise seen by the $N_{r,1D}^2$ isolated receive antennas are independent random variables, we have that $\sigma^2 = N_{r,1D}^2 \times (\mathrm{N}_0 W_s)$. Therefore, the effective SNR of the i-th sinusoid is given by:

$$\mathtt{SNR}_{\mathrm{eff}} = MLP \left| g_k \right|^2 / \mathrm{N}_0 W_s \ .$$

This must exceed the ZZB threshold $\mathtt{SNR}_{\mathrm{th}}$ for successful estimation. Noting that ML/W_s is the time taken for channel sounding, the ZZB threshold $\mathtt{SNR}_{\mathrm{th}}$ gives us the means to evaluate the minimum overhead in time to estimate the channel for a given path gain $\left| g_k \right|^2$:

$$\text{Time taken} = ML/W_s \geq \mathtt{SNR}_{\mathrm{th}} \, \mathrm{N}_0 \big/ P \left| g_k \right|^2 \tag{4.17}$$

The size of the picocell gives us a lower bound on $\left| g_k \right|^2$ and we later use this to guide us in choosing the sounding bandwidth W_s using (4.17).

4.5.4   Sounding rate

We round off the discussion on choice of protocol parameters by giving a rule of thumb for the rate/frequency f_B at which the spatial channel $\{\boldsymbol{\omega}_k\}$ needs to be *reestimated*. We use the estimated spatial frequency $\hat{\boldsymbol{\omega}}$ for beamforming purposes in the time period between two channel sounding rounds (communication phase sandwiched between consecutive sounding phases; see Figure 4.5). Following the discussion in Section 4.4.3, we have that if $\|\boldsymbol{\omega}(t) - \hat{\boldsymbol{\omega}}\|_\infty < \pi/N_{t,1D}$ throughout the communication phase, where $\boldsymbol{\omega}(t)$ denotes the true spatial frequency and $\hat{\boldsymbol{\omega}}$ the estimate from the prior sounding

round, then the *loss* in SNR, given by $\|\mathbf{x}_t(\boldsymbol{\omega}(t))\|^2/\langle\mathbf{x}_t(\hat{\boldsymbol{\omega}}), \mathbf{x}_t(\boldsymbol{\omega}(t))\rangle$, is smaller than 3dB. If we assume that the closest user to the basestation array is at a distance R meters and that the maximum speed of a user in the picocell is given by $v_{\max}$ meters per second, then the maximum change (in terms of the ℓ_∞-norm) in spatial frequency $\Delta\omega$ between consecutive sounding phases, spaced $1/f_B$ apart, is given by $2\pi d v_{\max}/f_B R \lambda$. The worst-case geometry which achieves this bound is when the user is at a distance R along the bore-sight of the array and heading in a direction aligned with the one of the array axes. For this worst-case geometry (plotted in Figure 4.4), we have that:

$$\Delta\omega \leq (2\pi d/\lambda)\sin\Delta\theta \approx \frac{2\pi d v_{\max}}{f_B \lambda R}.$$

Assuming that the estimate $\hat{\boldsymbol{\omega}}$ from the previous sounding phase is accurate, if we ensure that $2\pi d v_{\max}/f_B R \lambda \leq \pi/N_{t,1D}$, we have that the beamforming losses in the intervening period are smaller than 3dB. This tells us that channel needs to be sounded often enough so that

$$f_B \geq 2d v_{\max} N_{t,1D}/R\lambda . \tag{4.18}$$

In the following discussions, we use the preceding in conjunction with (4.17) to determine the overhead incurred in estimating the channel using the compressive architecture proposed herein.

4.6   System Design

We now discuss some key aspects of downlink system design related to our compressive architecture. We start by choosing basestation transmit power based on rules set by regulatory authorities, and then filling in the other details of the protocol according to the prescriptions laid out in Section 4.5. Fixing the mobile array to be 4×4 ($N_{r,1D} = 4$),
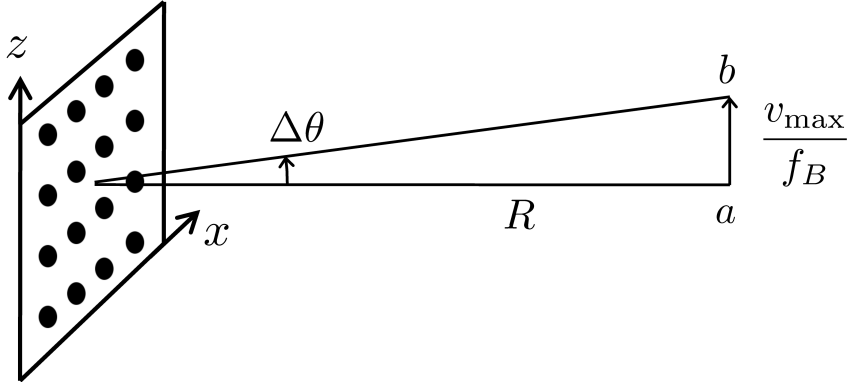
Figure 4.4: Geometry corresponding to maximum change in ω_z: The user moves along the z axis at a speed of $v_{\max}$ in the time interval $1/f_B$ between two consecutive channel sounding rounds

we consider two different choices for the base station array size: 8×8 ($N_{t,1D} = 8$) and 32×32 ($N_{t,1D} = 32$). All arrays are $d = \lambda/2$-spaced. The total available bandwidth for communication and sounding is 2GHz.

4.6.1   Transmit power

We fix the effective isotropically radiated power (EIRP) to 40dBm, consistent with Federal Communications Commission (FCC) regulations for 60 GHz unlicensed transmission. Accounting for transmit beamforming gain using an $N_{t,1D} \times N_{t,1D}$ array, the total transmit power

$$P = 40 - 20 \log N_{t,1D} \text{ dBm} \tag{4.19}$$

which evaluates to 22dBm and 10dBm for $N_{t,1D} = 8$ and $N_{t,1D} = 32$ respectively. Assuming that this power is evenly split among the $N_{t,1D}^2$ transmit elements, the power per transmit element is given by $P_e = P - 20 \log N_{t,1D} \text{ dBm} = 40 - 20 \log N_{t,1D} \text{ dBm}$, which evaluates to 4dBm and -20dBm, respectively. Assuming that we design each element to operate at a fixed power, this is also the power per element used in the beaconing phase, even though the latter does not get the benefit of transmit beamforming.

4.6.2   Communication range

In order to ensure that the SNR for compressive estimation is adequate over a picocell, we first determine the picocell size using a nominal communication link budget, and then calculate the overhead required for successful estimation at that range. Standard link budget calculations, assuming oxygen absorption of 16 dB/km, an EIRP of 40dBm and a 4×4 receive array providing directivity gains of 12 dBi, can be used to show that we can attain a per-symbol SNR of 6dB at a link margin of 10dB for a symbol rate of 2GHz at a range of 200m. For omnidirectional free space propagation, the power gain in dB as a function of range r is given by

$$G_{dB}(r) = -\mu r + 20 \log_{10} \frac{\lambda}{4\pi r} \quad \text{(dB)}$$

where $\mu = 0.016\text{dB/m}$ to account for oxygen absorption. Note that μ can be increased in order to account for rain. However, since our purpose is to ensure that channel estimation is successful whenever communication is successful, the contribution due to $G_{dB}(r)$ cancels out, as we show shortly. Thus, while the particular value of μ determines picocell size, we shall see that it does not affect the overhead for channel estimation.

The SNR per symbol is given by

$$\text{SNR}_{\text{c}}(\text{dB}) = \text{EIRP (dBm)} + G_{dB}(r) + 20 \log_{10} N_{r,1D} \tag{4.20}$$
$$- 10 \log_{10} (\text{N}_0 W_c) - L_{margin}(comm)$$

where $L_{margin}(comm)$ is the link margin (dB) for communication. Note that $10 \log_{10} N_0 = -174 + NF$ dBm over a bandwidth of 1 Hz, where NF denotes the receiver noise figure in dB. Plugging in $W_c = 2\text{GHz}$, $N_{r,1D} = 4$, and $NF = 6\text{dB}$, we obtain a per symbol SNR of 7 dB at a range of $r = 100$ meters.

4.6.3   Channel sounding protocol

Our channel sounding protocol is specified by four parameters: (i) bandwidth used by each basestation when sounding the channel, which we denote by W_s (ii) number of transmit beacons (or virtual transmit antennas) M (iii) number of receive measurements per transmit beacon (or virtual receive antennas) L and (iv) sounding rate f_B which determines how often the channel is sounded. The parameters M, L and W_s together determine the effective sounding SNR. This must exceed the ZZB threshold SNR for successful channel estimation. This gives rise to the condition in 4.17. Imposing an estimation link margin $L_{margin}(est)$ (dB) and going to the dB domain, we have

$$10 \log_{10} \left(ML/W_s \right) \geq \texttt{SNR}_{\text{th}} + L_{margin}(est) \tag{4.21}$$
$$+ 10 \log_{10} N_0 - P - G_{dB}(r)$$

Adding (4.21) and (4.20) and simplifying, we obtain

$$10 \log_{10} \left(ML/W_s \right) \geq \texttt{SNR}_{\text{th}} - \texttt{SNR}_{\text{c}} + L_{margin}(est)$$
$$- L_{margin}(comm) + 20 \log_{10} N_{t,1D}$$
$$+ 20 \log_{10} N_{r,1D} - 10 \log_{10} W_c$$

The key take-away is that ML/W_s must be large enough to compensate for the fact that we do not have the benefit of beamforming during the sounding phase. Notice that the range r (i.e., the dependence on picocell size) has cancelled out. Setting $L_{margin}(est) =$ 16dB (we use a higher link margin for channel sounding to account for power losses due

to randomness of $\boldsymbol{A}$ and $\boldsymbol{B}$), we obtain

$$\text{Time taken} = \frac{ML}{W_s} \geq \begin{cases} 16.34 \ \mu s & N_{t,1D} = 8 \\ 0.2669 \ \text{ms} & N_{t,1D} = 32. \end{cases} \tag{4.22}$$

We choose the number of transmitter beacons for the 8×8 and 32×32 transmitter arrays based on the geometry preservation criterion for the transmitter's spatial channel estimation problem discussed in Section 4.5.1. We use $M = 24$ for $N_{t,1D} = 8$ and $M = 30$ for $N_{t,1D} = 32$ by numerically evaluating the worst-case distortion of pairwise distances relevant for the channel estimation problem (in Figure 4.1 we plot the worst-case distortion as a function of L for a random instance of $\boldsymbol{A}$ and $N_{t,1D} = 32$). Using the receive energy preservation criterion given in Section 4.5.2, we choose the number of receive weights for the 4×4 receive array as $L = 6$. Using these values for M and L in (4.22), we obtain that the channel sounding bandwidth must satisfy

$$W_s \leq \begin{cases} 8.8124 \ \text{MHz} & N_{t,1D} = 8 \\ 674.34 \ \text{KHz} & N_{t,1D} = 32. \end{cases}$$

Our specification of the channel sounding protocol will be complete when we give f_B, the rate at which we sound the channel (see Figure 4.5) which must satisfy (4.18). Assuming that the closest user is at a distance of $R = 20$m and that the maximum speed of a user in the picocell $v_{\max}$ is 45 miles per hour (20 m/s), we have that: $f_B \geq 8$ Hz for $N_{t,1D} = 8$ and $f_B \geq 32$ Hz for $N_{t,1D} = 32$. Choosing the minimum value for f_B, we have that the overhead for our channel sounding protocol is $MLf_B/W_s = 0.0131\%$ for $N_{t,1D} = 8$ and 0.8542% for $N_{t,1D} = 32$.

$$\xleftarrow{\hspace{1em}} \frac{ML}{W_s} \xrightarrow{\hspace{1em}}$$

| Sounding phase | Communication phase | Sounding phase | |

$$\xleftarrow{\hspace{6em}} 1/f_B \xrightarrow{\hspace{6em}}$$
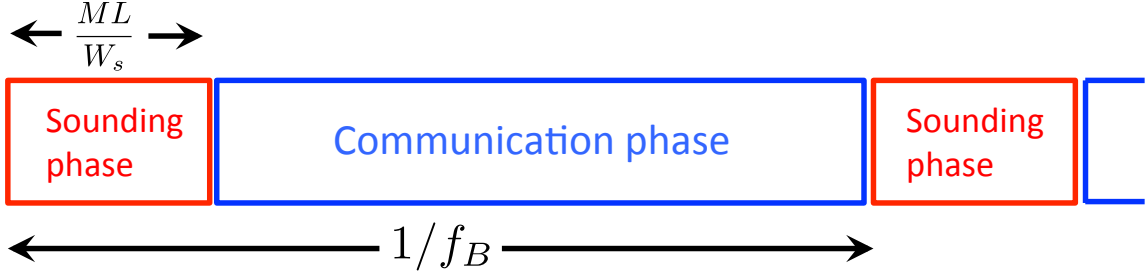
Figure 4.5: Channel sounding and communication phases of the proposed system

4.6.4 Reuse analysis for channel sounding

We investigate how a sequence of basestations employed in an urban canyon environment can share resources when estimating the spatial channel to users in their respective cells. The envisioned mm-wave system involves alternating between channel estimation and communication phases as shown in Figure 4.5. We assume that channel sounding rounds across basestations are aligned in time. We now characterize how the 2GHz spectrum is to be shared in space so as to limit the effect of interference from neighboring picocells on channel estimation performance. Such interference management is essential in the sounding phase; unlike the highly directive beams used in the communication phase, compressive sounding beacons are essentially omnidirectional. To see this, consider the average transmit power along any direction $\boldsymbol{\omega}$. This is given by $P_e \left\| \boldsymbol{A} \mathbf{x}_t(\boldsymbol{\omega}) \right\|^2 / L \approx P_e \left\| \mathbf{x}_t(\boldsymbol{\omega}) \right\|^2 = P_e N_{t,1D}^2 = P$, the total transmit power. The approximation $\left\| \boldsymbol{A} \mathbf{x}_t(\boldsymbol{\omega}) \right\|^2 / L \approx \left\| \mathbf{x}_t(\boldsymbol{\omega}) \right\|^2$ holds when the number of beacons L is large enough. Therefore, the average energy *per-measurement* received by an antenna at a distance r from a transmitter sending compressive beacons is given by $PG(r)$, where $G(r) = 10^{G_{dB}(r)/10} = \lambda^2 / (16\pi^2 r^2) \, e^{-\nu r}$ $(\nu = (\mu/10) \ln 10)$ is the omnidirectional power gain at range r. We assume that basestations are deployed regularly as shown in Figure 4.6 and that the inter-basestation separation (along the street) is given by S. Suppose that the reuse factor is R (i.e, every Rth basestation uses the same slice of the frequency spectrum to estimate downlink spatial channels). We assume that for narrow urban

canyons, the distance between a user and all interfering basestations (those that are allocated the same sounding BW) are well approximated by $\{kR_fS,\ k \in \mathbb{Z} \setminus \{0\}\}$. Thus, the interference power seen by a single antenna is given by

$$I = 2 \times \sum_{k=1}^{\infty} \sum_{\text{paths}} PG(kR_fS) = 8P \sum_{k=1}^{\infty} G(kR_fS),$$

where we have assumed that there are 4 viable paths between the interfering basestation and user, each introducing the same amount of interference as the LoS path. This is a pessimistic assumption, since NLOS paths are attenuated by larger path lengths and reflection losses. Plugging in the expression for $G(r)$, we have that

$$I = \left(P\lambda^2 / 2\pi^2 R_f^2 S^2 \right) \sum_{k=1}^{\infty} e^{-\nu R_f Sk} / k^2$$

$$= \left(P\lambda^2 / 2\pi^2 R_f^2 S^2 \right) \text{Li}_2 \left(e^{-\nu R_f S} \right),$$

where $\text{Li}_2(z) = \sum_{k=1}^{\infty} z^k / k^2$ is the dilogarithm function. The interference seen per antenna adds to thermal noise to give an effective per-element noise level of $\sigma_e^2 = \text{N}_0 W_s + I$. Assuming a worst-case geometry for the user of interest (distance of S from the basestation) and proceeding as in Section 4.5.3, we see that effective Signal to Interference and Noise Ratio $\texttt{SINR}_{\text{eff}}$ is given by

$$\texttt{SINR}_{\text{eff}} = MLPG(S) / \sigma_e^2 .$$

This can be rewritten as

$$1/\texttt{SINR}_{\text{eff}} = 1/\texttt{SNR}_{\text{eff}} + 1/\texttt{SIR}_{\text{eff}} ,$$

where $\text{SNR}_{\text{eff}} = MLPG(S)/\text{N}_0 W_s$ and the Signal to Interference Ratio $\text{SIR}_{\text{eff}} = MLPG(S)/I = MLR_f^2 e^{-\mu S}/8\text{Li}_2\left(e^{-\mu R_f S}\right)$. We need to ensure that SINR_{eff} exceeds the ZZB SNR threshold for successful estimation. We choose the reuse factor R_f so that we are in the noise-limited regime by setting

$$1/\text{SIR}_{\text{eff}} < 0.1 \times (1/\text{SNR}_{\text{th}}) \approx -10 - 16 \text{ dB}.$$

Assuming that protocol parameters are chosen so that SNR_{eff} exceeds SNR_{th}, we can ignore interference in SINR_{eff} calculations when

$$\text{SIR}_{\text{eff}} > \text{SNR}_{\text{th}} + 10 \approx 26\text{dB for } N_{t,1D} = 8, 32.$$

In Figure 4.7, we plot achievable effective SIRs as a function of frequency reuse factor R_f for two example systems in Section 4.6.3: i.e, 8×8 and 32×32 arrays with total number of measurements given by $ML = 24 \times 6$ and $ML = 30 \times 6$ respectively. As the picocell size S grows, exponential attenuation due to oxygen absorption (the $e^{-\nu S}$ term in the expression for SIR_{eff}) helps in attenuating interference and improving SIR for same reuse factor R_f. To illustrate this we plot SIR as a function of R_f for three cell sizes $S = 50, 100, 200\text{m}$ in Figure 4.7. We observe that, in order to ensure $\text{SIR}_{\text{eff}} > 26\text{dB}$, a reuse factor of $R_f = 4$ is needed for $S = 50\text{m}$, while $R_f = 3$ suffices for $S = 200\text{m}$. Plugging in the per-basestation sounding bandwidth W_s calculations in Section 4.6.3, we see that the overall *system-level* channel sounding bandwidth $W_s \times R_f$ for a $S = 50\text{m}$ picocell is as small as $8.8124\text{MHz} \times 4 = 35.2\text{MHz}$ and $674.34\text{KHz} \times 4 = 2.7\text{MHz}$ for $N_{t,1D} = 8, 32$, respectively, which is dwarfed by the total available bandwidth (2GHz).
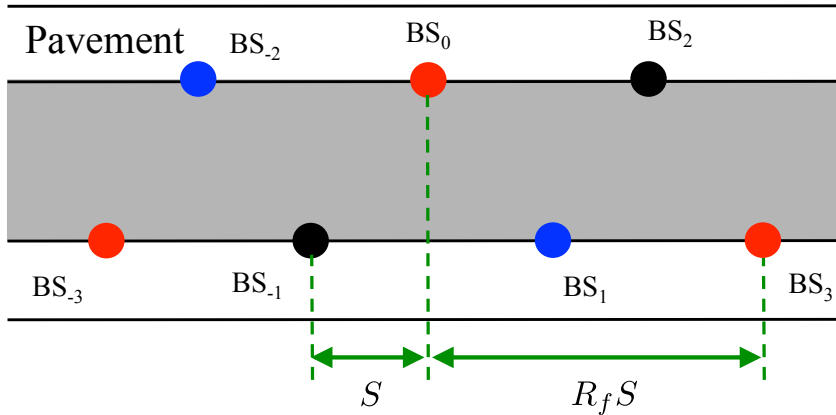
Figure 4.6: Reuse of frequency resources for reuse factor $R_f = 3$

4.7   Simulation Results

We perform simulations for the two example systems considered in Section 4.6 (8×8 and 32×32 transmit arrays). We report results for the algorithm proposed in Section 4.4 and two feedback strategies: (i) 'full': users feedback the measured virtual channel matrix $\mathbf{Y}$ ($M \times L$ matrix; $L = 6$ for both systems) and (ii) 'svd': users feedback the 2 dominant left singular vectors of $\mathbf{Y}$, scaled by their corresponding singular values ($M \times 2$ matrix; one-third feedback overhead).

We consider 6 mobile users moving in the urban canyon at speeds of 20, 3, 15, 1.5, 2.1 and 10 meters per second (covering both vehicular and pedestrian settings). The height at which each mobile device is held is in the $1.3 - 1.4$m range. The basestation is mounted on a lamppost on the pavement (7 meters from a canyon wall), at a height of 6 meters. The basestation antenna array is tilted by about $7.5°$ in both the azimuth and elevation directions so that the boresight of the array points towards middle of the corresponding cell. This helps in more accurate spatial frequency estimation: since a change in direction near the boresight of the array results in larger changes in spatial frequencies than far away from the boresight, resolving paths is easier when the array points towards a direction in which we are likely to see more paths. We do not model
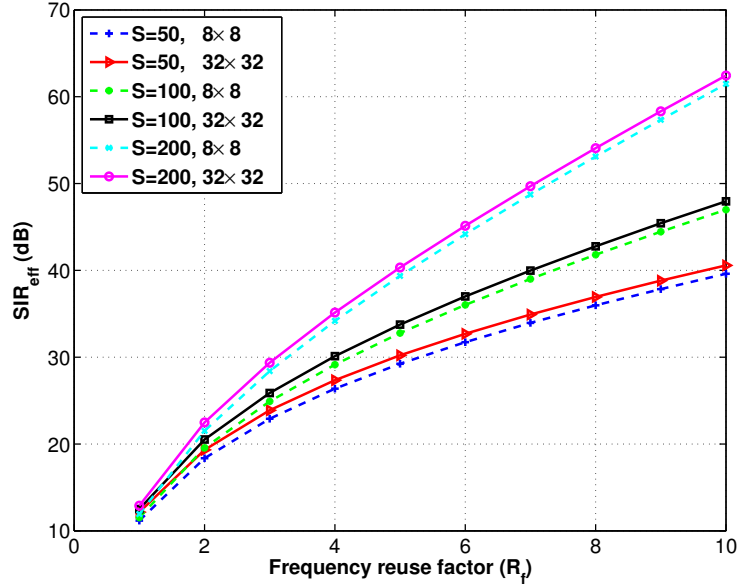
Figure 4.7: Effective Signal to Interference Ratio $\texttt{SIR}_{\text{eff}}$ for $ML = 24 \times 6$ (8×8 scenario) and $ML = 30 \times 6$ (32×32 scenario; dashed) as a function of reuse factor R_f

blockage in these simulations, assuming that the LoS path and the three first order reflections are all available. Our goal is to estimate and track the $K = 4$ paths to all 6 users.

Estimation error: Let $T = \{\boldsymbol{\omega}_m \ : \ m = 1, \ldots, K\}$ denote the true spatial frequencies and $P = \{\hat{\boldsymbol{\omega}}_n \ : \ n = 1, \ldots, \hat{K}\}$ denote the set of estimated spatial frequencies. When the base station uses one of the estimates in P, say $\hat{\boldsymbol{\omega}}$, to form a beam, we do not realize the full $20 \log N_{t,1D}$ dB beamforming gain. A measure of the sub-optimality is the estimation error $\|\boldsymbol{\omega} - \hat{\boldsymbol{\omega}}\|_2$, which we normalize by the DFT spacing of $2\pi/N_{t,1D}$ to define the following error metric:

$$\Delta\omega(m) = \min_n \|\boldsymbol{\omega}_m - \hat{\boldsymbol{\omega}}_n\|_2 \Big/ \left(2\pi/N_{t,1D}\right) . \tag{4.23}$$

When no true spatial frequency exists near an estimate $\hat{\boldsymbol{\omega}}$, i.e, when $\hat{\boldsymbol{\omega}}$ is a "phantom estimate", we will quickly be able to discard it when we beamform in the direction of $\hat{\boldsymbol{\omega}}$
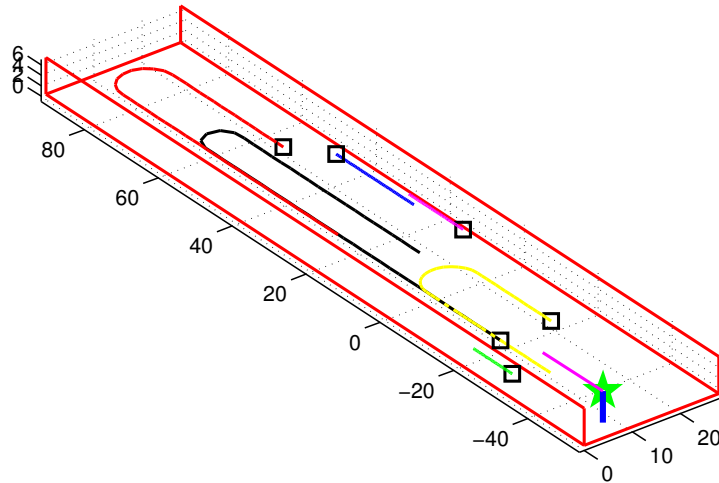
Figure 4.8: Six users in the urban canyon moving over the duration of the 7 second simulation interval. Their positions at time $t = 0$ is marked using a $\square$-symbol

and find that the mobile does not receive power commensurate to what it expects with the $20 \log N_{t,1D}$ beamforming gain.

We plot the Complementary Cumulative Distribution Function (CCDF) of estimation errors (4.23) for the two systems ($N_{t,1D} = 8, 32$) in Figure 4.9 and the Probability Distribution Function (PDF) of the number of paths estimated $\hat{K}$ (correct value is $K = 4$) in Figure 4.10. We have set $P_{fa} = 10^{-3}$ to arrive at τ used in the stopping criterion (4.13). From Figure 4.9 and 4.10, we see that feedback of dominant singular vectors is an efficient feedback strategy which performs just as well as feeding back the entire matrix $\mathbf{Y}$, while using only a third of uplink resources.

Next, in order to evaluate the effect of errors in spatial frequency estimation on beamforming performance, we simulate a simple scenario in which the transmitter beamforms toward the strongest estimated path. Figure 4.11 shows the CDF of the achievable beamforming gain for an 8×8 array. While ideal beamforming requires adjustment of both gains and phases, suboptimal approaches for RF beamsteering with severely quantized phase-only control (four phases) have been studied in our earlier conference paper [31]. We see from Figure 4.11 that if ideal beamforming were performed with our estimates,
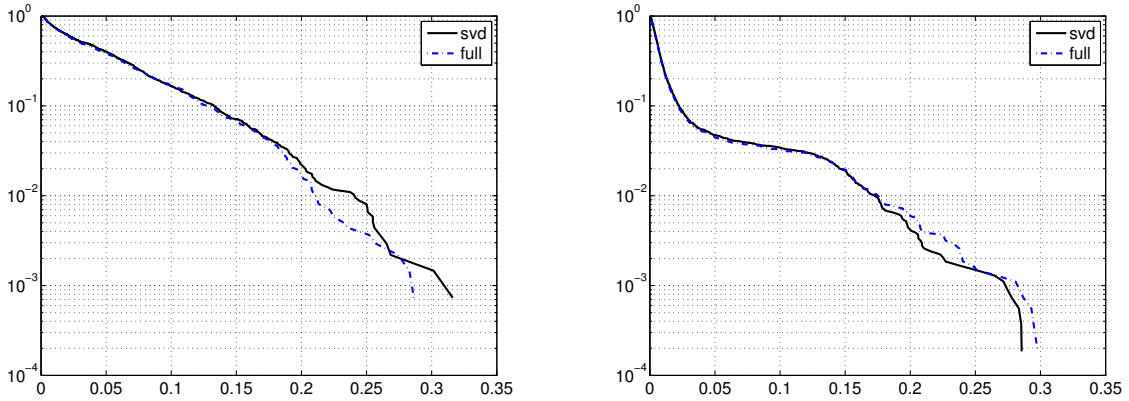
Figure 4.9: CCDF of frequency estimation errors with new stopping criterion for $N_{t,1D} = 8$ (left) and $N_{t,1D} = 32$ (right). Two feedback strategies considered: (i) $M \times 6$ matrix $\mathbf{Y}$ ('full') and (ii) top two dominant singular vectors (one-third overhead).

then the SNR loss is less than 0.3 dB. If four-phase control is used based on our estimates, then the SNR loss is less than 1 dB. The results for 32×32 arrays are entirely similar, and are therefore not plotted here.

Thus far, we have not said anything about channel frequency selectivity. Our proposed algorithm uses a small segment of the band to estimate the spatial channel, and the problem of channel dispersion is not addressed. However, we note that beamforming using a large array should reduce the effect of undesired paths, which simplifies the task of equalization. Figure 4.12 shows the channel impulse responses for the 32×32 and 8×8 antenna arrays for a typical snapshot, when the transmitter beamforms towards the strongest estimated path. In our simulated setting, the LoS and ground reflection are close to each other in terms of both delays and angles of departure. We see that 8×8 array fails to resolve them, with both paths falling into the antenna's main lobe, while the 32×32 antenna array, which has smaller beamwidth ($4°$ half power beamwidth), attenuates the undesired tap down to one-ninth of the desired path. Of course, it is possible to utilize the channel estimates far more intelligently, potentially with nulls directed both at strong
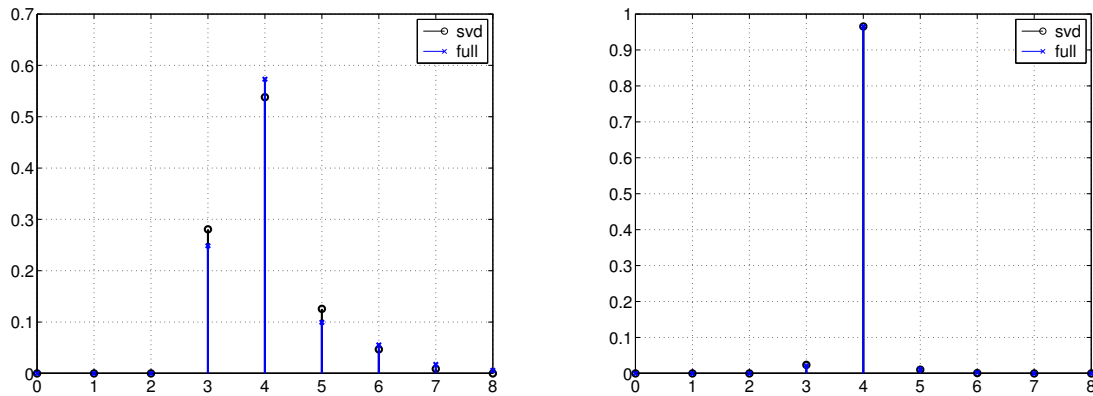
Figure 4.10: PDF of # of estimated paths $\hat{K}$ ($K = 4$) with new stopping criterion for $N_{t,1D} = 8$ (left) and $N_{t,1D} = 32$ (right). Two feedback strategies considered: (i) $M \times 6$ matrix $\mathbf{Y}$ ('full') and (ii) top two dominant singular vectors (one-third overhead).

undesired paths for the mobile of interest, and at the dominant paths for other nearby mobiles. The latter can be particularly useful for combating intra-cell interference when a base station face has multiple antenna arrays, each communicating with a different mobile.
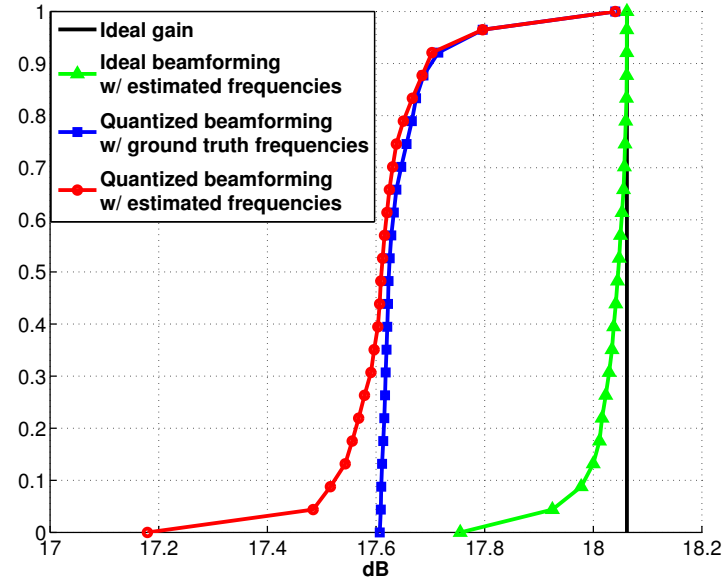
Figure 4.11: CDF of beamforming gain achieved by an 8×8 antenna array for ideal and quantized beamforming techniques
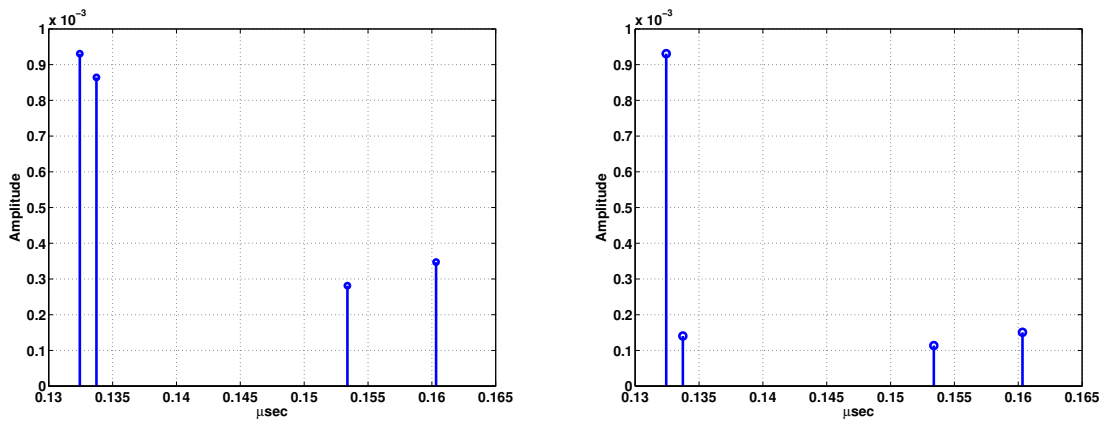

Figure 4.12: Channel impulse response with quantized beamforming towards estimated strongest path for the 8×8 (left) and 32×32 (right) scenarios

Chapter 5

Conclusions and Future Work

In the preceding chapters we have established that 60GHz outdoor picocells are indeed a feasible approach for delivering the orders of magnitude increase required in network capacity. Moreover, the mm-wave picocellular architecture proposed here could be exploited for other purposes i.e., *vehicular sensing and communication* as well as *localization*. However, there are significant design challenges in realizing our proposed architecture. We end by discussing a few of these open issues and other applications that could benefit from it:

Picocellular networks: Our findings lead us to identify the need for new, novel architectures and system designs such as picoclouds, opening up a rich new research area on 60GHz outdoor network design and experimentation. The compressive approach allows each base station to build up an inventory of viable paths to nearby mobiles, but there is a huge design space to be explored on how base stations coordinate using this information to alleviate the effects of blockage, to maintain the connectivity for *mobile* users and to handle the channel dispersion. Furthermore, frequent switching between basestations, together with rate adaptation switching between different paths, requires careful coordination to support applications such as TCP. Such considerations,

along with the large per-user rate, call for careful design that accounts for backhaul and storage constraints.

Compressive channel estimation and tracking: We have shown that it is possible to super-resolve mm-wave spatial channels with a relatively small number of compressive measurements, in a manner that is compatible with coarse phase-only control and RF beamforming. This allows scaling to a very large number of antenna elements without relying on channel reciprocity. While our discussion of system design issues such as link budget and inter-cell beacon interference is tailored to outdoor 60 GHz picocellular networks, the basic approach is broadly applicable (e.g., to other bands, and to indoor environments).

We later conducted comprehensive experimental validation of our compressive approach. In [1] we proposed the *noncoherent* framework which enables immediate deployment of compressive tracking with commodity IEEE 802.11ad hardware, since it requires the same information (i.e., RSS) as conventional scan-based techniques. Using this framework, we have provided the first experimental demonstration of compressive mm-wave path estimation, showing that it provides accurate estimates at significantly lower overhead than standard scanning.

Some natural next steps are as follows. First, it is of interest to extend the noncoherent framework for estimating multiple paths, and to understand it more deeply at a theoretical level. Preliminary results combining compressive estimation with phase retrieval are promising [49]. A second avenue is to pursue the custom hardware and protocol changes required for obtaining the performance advantage provided by coherent compressive measurements, and evaluating experimentally whether the required phase coherence across beacons, and the gains promised by theoretical studies such as [50], can be attained despite the higher phase noise at mm-wave frequencies.

A hardware enhancement to pure RF beamforming as considered here (where a single RF chain serves all antenna elements) is to employ hybrid analog-digital beamforming, with a number of RF chains smaller than the number of antenna elements. We may term this an *array of subarrays,* with RF-level control for subarrays, and digital processing of subarray outputs. Such an approach is used in [30, 51] for spatial channel estimation. Our work has shown, however, that a single RF beamformed array suffices for this purpose. In general hybrid/fully digital beamforming is not always ideally suited for practical 5G applications due to increased cost, energy consumption and complicated integration in mobile devices. However, arrays of subarrays are certainly required for more advanced functionalities such as multiuser MIMO [52], spatial multiplexing [53, 54, 55, 56, 57], and spatial diversity [55, 58, 59]. Moreover, *beam squint*[1] is a well-known problem for analog beamforming architectures using phase offsets. This is a serious drawback considering current 5G plans to make use of large bandwidths in the mm-wave band and could encourage migration to hybrid or fully digital architectures where digital control of the RF chain enables optimization of the phases according to the frequency over a large band [60]. Hence integrating the compressive approach proposed in this work within an array of subarrays architecture is an interesting area for future work.

Vehicular communications: mm-wave band are strong candidates for future V2V (Vehicle-to-Vehicle) and V2X (Vehicle-to-everything) links where current technologies (i.e., 4G cellular or DSRC (Dedicated Short-Range Communications)) has proven to be insufficient [61]. In vehicular scenarios, contextual awareness is a key factor. mm-wave band offers the potential to expands each vehicles own-sensing with realtime information retrieved from nearby vehicles. Such use of V2V mm-wave links to enable connected intelligence can effectively help to address both blind area and bad weather problems

---

[1]The phase relation between different antenna elements is calculated with a certain carrier frequency in mind. Now, if the antenna elements are fed with a signal of a slightly different frequency the actual angle of the main lobe shifts by a certain angle.

inherent to LiDAR and other sensing equipments [62]. In this context there are plenty of open issues in realizing the new paradigm of V2X for sensing and cellular communication.

Localization: mm-wave systems offer huge potential for radio-based localization due to their ultrawide bandwidths as well as directional antennas that enables accurate identification of multipath components in temporal and angular domains [63, 64, 65]. Localization accuracy can be improved by providing *a priori* information on spatial characteristics of urban environment, also termed REM (Radio Enviromental Mapping). One possible future direction could be realizing REM via mm-wave localization which can later be exploited to improve the localization accuracy.

Appendix A

Stopping Criterion

In this section we derive the CFAR-based stopping criterion (4.13). In essence, we wish to determine when the residual measurements, after subtracting out the contribution due to the paths estimated thus far, can be explained well enough by noise, up to a nominal false alarm probability of P_{fa}.

Under the noise-only hypothesis (i.e., assuming all existing paths have been detected and subtracted), we have

$$\mathbf{v}_l(\mathcal{P}_q) \approx \mathbf{z}_l \quad \mathbf{z}_l \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbb{I}_M\right) \quad l = 1, \dots, L$$

Recall that we use (4.9), maximized over an oversampled DFT grid, to detect a new frequency. For our stopping criterion (4.13), we consider the maximum over a DFT grid (this simplifies analysis and design), and stop if this is lower than a threshold, as follows:

$$\max_{\boldsymbol{\omega} \in DFT} \frac{1}{||\mathbf{x}(\boldsymbol{\omega})||^2} \sum_{l=1}^{L} |\langle \mathbf{x}(\boldsymbol{\omega}), \mathbf{v}_l(\mathcal{P}_q)\rangle|^2 \leq \tau \tag{A.1}$$

The false alarm rate is the probability that the threshold is exceeded under the noise-only

hypothesis:

$$P_{fa} = \Pr\left\{\max_{\boldsymbol{\omega} \in DFT} \frac{1}{||\mathbf{x}(\boldsymbol{\omega})||^2} \sum_{l=1}^{L} |\langle \mathbf{x}(\boldsymbol{\omega}), \mathbf{z}_l\rangle|^2 > \tau\right\} \tag{A.2}$$

The noise projections are i.i.d. $\mathcal{CN}(0,1)$ across ω (due to the orthogonality of the DFT basis) and i.i.d. across measurement time indices l. Thus,

$$\nu_{l,\omega} := \frac{\langle \mathbf{x}(\boldsymbol{\omega}), \mathbf{z}_l\rangle}{\sigma||\mathbf{x}(\boldsymbol{\omega})||} \sim \mathcal{CN}(0,1)$$

and the sum of the squared magnitudes across projections is a Gamma random variable:

$$\mu_{\omega} := \sum_{l=1}^{L} |\nu_{l,\omega}|^2 \sim Gamma\,(L,1)$$

We therefore obtain

$$P_{fa} = P\left\{\max_{\boldsymbol{\omega} \in DFT} \mu_{\omega} > \frac{\tau}{\sigma^2}\right\} = 1 - P\left\{\mu_{\omega_0} < \frac{\tau}{\sigma^2}\right\}^{N_{1D,t}^2}$$

$$= 1 - \left(\gamma(L, \tau/\sigma^2)/\Gamma(L)\right)^{N_{1D,t}^2}$$

where $\gamma(L,x) = \int_0^x t^{L-1}e^{-t}\,dt$ is the incomplete gamma function, and $\Gamma(L) = \gamma(L,\infty)$ is the gamma function.

For a given P_{fa}, the threshold τ is now given by

$$\tau = \sigma^2\gamma^{-1}\left(L, \Gamma(L)(1 - P_{fa})^{1/N_{1D,t}^2}\right) \tag{A.3}$$

where $\gamma^{-1}(L,y)$ denotes the inverse of the incomplete gamma function with respect to the integral limit x.

Appendix B

Complexity Analysis

We go over the computations that are needed to maintain the estimate of one user's channel over time and point out portions of the computation phase that can be reused across users.

The first step involves computing responses to each of the $R^2 N$ spatial frequency grids ($N = N_{1D,t}^2$ as the number of transmitter antennas):

$$\mathcal{X} = \{\mathbf{x}(\boldsymbol{\omega}) = \boldsymbol{A}\mathbf{x}_t(\boldsymbol{\omega}) : \boldsymbol{\omega} \in \Phi\}.$$

This can be efficiently computed using the Fast Fourier Transform in $O(MNR^2 \log(NR))$ time and can be shared across users. Furthermore, the set of responses $\mathcal{X}$ can be stored at the basestation and reused across sounding rounds. Thus, the cost of computing $\mathcal{X}$ is amortized over users in the network and across sounding rounds.

The next step involves finding the strongest path using (4.9) and the corresponding gains. This operation takes $O(MLNR^2)$ time. We then proceed to the refinement stage which takes $O(M(L+N))$ time ($O(MN)$ to compute $\partial\mathbf{x}(\boldsymbol{\omega})/\partial\omega_i$ and $O(ML)$ for gradient and Hessian computation).

When multiple paths are present, each path is refined again for every subsequently detected path. Therefore, the overall complexity when a new user joins the network is $O(K^2) \times$ refinement complexity $+ K \times$ detection complexity $= O(MK^2(L + N) + KMLNR^2)$. In arriving at this scaling, we use the fact that our algorithm is residue centric and therefore we keep track of $\{\mathbf{v}_l(\mathcal{P}_q)\}$. As a result we can compute $\{\mathbf{v}_l(\mathcal{P}_q \setminus (\hat{\boldsymbol{\omega}}_k, \{\hat{h}_{k,l}\}))\}$ in $O(LM)$ time as opposed to the $O(qLM)$ time that would have been needed if this computation were to be done directly from $\{\mathbf{y}_l\}$.

When we track the channel using estimates from prior rounds, we compute the gains $\{\hat{h}_{k,l}\}$ of the existing spatial frequency estimates from prior round given by $\{\hat{\boldsymbol{\omega}}_k\}$ as $(X^H X)^{-1} X^H Y$ and this takes $O(KM(L + N))$ time. We then refine these estimates using our refinement algorithm in $O(KM(L + N))$ time.

Paths that were viable in previous sounding rounds may be blocked in the current sounding round. To prune out such paths, we test whether the residual measurements after the deletion of the under question path satisfy our stopping criterion. This check takes $O(KM(L + N) + MLN)$ time for each path and path pruning complexity scales as $O(K^2 M(L + N) + KMLN)$. Checking whether a new path exists (stopping criterion) also costs $O(MLN)$ time. Therefore, the overall tracking complexity scales as $O(KM(KL + KN + LN)) = O(KMN(L + K))$ (since $N > K$) time.

Bibliography

[1] M. E. Rasekh, Z. Marzi, Y. Zhu, U. Madhow, and H. Zheng, *Noncoherent mmwave path tracking*, in *Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications*, pp. 13–18, ACM, 2017.

[2] http://www.techjournal.org/2011/09/
mobile-broadband-useage-is-set-to-explode-infographic.

[3] http://money.cnn.com/2012/02/24/technology/spectrum_crunch_solutions/
index.htm.

[4] D. Costello and J. Forney, G.D., *Channel coding: The road to channel capacity*, *Proc. of the IEEE* **95** (2007), no. 6 1150–1177.

[5] http://www.dailywireless.org/2011/05/06/
att-moves-to-hotzones-and-picocells.

[6] A. Balasubramanian, R. Mahajan, and A. Venkataramani, *Augmenting mobile 3G using WiFi*, in *Proc. of MobiSys*, 2010.

[7] D. Ramasamy, R. Ganti, and U. Madhow, *On the capacity of picocellular networks*, in *Proc. of ISIT*, 2013.

[8] Y. Zhu, Z. Zhang, Z. Marzi, C. Nelson, U. Madhow, B. Y. Zhao, and H. Zheng, *Demystifying 60ghz outdoor picocells*, in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, MobiCom '14, (New York, NY, USA), pp. 5–16, ACM, 2014.

[9] S. Bayhan, G. Gür, and A. Zubow, *The future is unlicensed: Coexistence in the unlicensed spectrum for 5g, arXiv preprint arXiv:1801.04964* (2018).

[10] I. Qualcomm Technologies, *Making 5g nr a reality, white paper* (2016).

[11] H. Haas and N. Serafimovski, *Does 5g have a bright future? light based communications for wireless solutions*, .

[12] "IEEE 802.11 Task Group AD."
http://www.ieee802.org/11/Reports/tgad_update.htm.

[13] M. Rasekh, A. Shishegar, and F. Farzaneh, *A study of the effect of diffraction and rough surface scatter modeling on ray tracing results in an urban environment at 60 ghz*, in *Millimeter-Wave and Terahertz Technologies (MMWaTT), 2009 First Conference on*, pp. 27–31, Dec, 2009.

[14] K. Maeda *et. al.*, *Urban pedestrian mobility for mobile wireless network simulation*, *Ad Hoc Networks* **7** (2009), no. 1 153–170.

[15] R. C. Hansen, *Phased array antennas.* John Wiley & Sons, 2nd ed., 2009.

[16] H. Zhang, S. Venkateswaran, and U. Madhow, *Channel modeling and mimo capacity for outdoor millimeter wave links*, in *Proc. of WCNC*, 2010.

[17] D. Ramasamy, R. Ganti, and U. Madhow, *On the capacity of picocellular networks*, in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pp. 241–245, July, 2013.

[18] E. Visotsky and U. Madhow, *Optimum beamforming using transmit antenna arrays*, in *Vehicular Technology Conference, 1999 IEEE 49th*, vol. 1, pp. 851–856, IEEE, 1999.

[19] A. Wiesel, Y. C. Eldar, and S. Shamai, *Linear precoding via conic optimization for fixed mimo receivers*, *IEEE Transactions on Signal Processing* **54** (2006), no. 1 161–176.

[20] S. Ulukus and R. D. Yates, *Adaptive power control and mmse interference suppression*, *Wireless Networks* **4** (1998), no. 6 489–496.

[21] F. Rashid-Farrokhi, K. R. Liu, and L. Tassiulas, *Transmit beamforming and power control for cellular wireless systems*, *IEEE Journal on Selected Areas in Communications* **16** (1998), no. 8 1437–1450.

[22] F. Rashid-Farrokhi, L. Tassiulas, and K. R. Liu, *Joint optimal power control and beamforming in wireless networks using antenna arrays*, *IEEE transactions on communications* **46** (1998), no. 10 1313–1324.

[23] B. Zhuang, D. Guo, and M. L. Honig, *Traffic driven resource allocation in heterogenous wireless networks*, *arXiv preprint arXiv:1405.6636* (2014).

[24] B. Zhuang, D. Guo, E. Wei, and M. Honig, *Large-scale spectrum allocation for cellular networks via sparse optimization*, *IEEE Transactions on Signal Processing* (2018).

[25] A. Valdes-Garcia, S. Nicolson, J.-W. Lai, A. Natarajan, P.-Y. Chen, S. Reynolds, J.-H. Zhan, D. Kam, D. Liu, and B. Floyd, *A fully integrated 16-element phased-array transmitter in sige BiCMOS for 60-ghz communications*, *Solid-State Circuits, IEEE Journal of* **45** (Dec, 2010) 2757–2773.

[26] E. Cohen, M. Ruberto, M. Cohen, O. Degani, S. Ravid, and D. Ritter, *A CMOS bidirectional 32-element phased-array transceiver at 60ghz with ltcc antenna*, in *Radio Frequency Integrated Circuits Symposium (RFIC), 2012 IEEE*, pp. 439–442, June, 2012.

[27] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, *Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems*, *Selected Areas in Communications, IEEE Journal on* **27** (October, 2009) 1390–1399.

[28] S. Hur, T. Kim, D. Love, J. Krogmeier, T. Thomas, and A. Ghosh, *Multilevel millimeter wave beamforming for wireless backhaul*, in *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pp. 253–257, Dec, 2011.

[29] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, *Millimeter wave beamforming for wireless backhaul and access in small cell networks*, *Communications, IEEE Transactions on* **61** (2013), no. 10 4391–4403.

[30] A. Alkhateeb, O. El Ayach, G. Leus, and R. Heath, *Channel estimation and hybrid precoding for millimeter wave cellular systems*, *Selected Topics in Signal Processing, IEEE Journal of* (Oct, 2014).

[31] D. Ramasamy, S. Venkateswaran, and U. Madhow, *Compressive adaptation of large steerable arrays*, in *Information Theory and Applications Workshop (ITA), 2012*, pp. 234 –239, feb., 2012.

[32] D. Ramasamy, S. Venkateswaran, and U. Madhow, *Compressive tracking with 1000-element arrays: A framework for multi-gbps mm wave cellular downlinks*, in *Proc. of Allerton*, 2012.

[33] D. Ramasamy, S. Venkateswaran, and U. Madhow, *Compressive estimation in awgn: general observations and a case study*, in *Proc. 2012 Asilomar Conf. on Signals, Systems and Computers (Asilomar 2012)*, November, 2012. to appear.

[34] D. Ramasamy, S. Venkateswaran, and U. Madhow, *Compressive parameter estimation in AWGN*, *Signal Processing, IEEE Transactions on* (April, 2014).

[35] S. Singh, R. Mudumbai, and U. Madhow, *Interference analysis for highly directional 60-ghz mesh networks: The case for rethinking medium access control*, *Networking, IEEE/ACM Transactions on* **19** (Oct, 2011) 1513–1527.

[36] M. Park and P. Gopalakrishnan, *Analysis on spatial reuse, interference, and mac layer interference mitigation schemes in 60 ghz wireless networks*, in *Ultra-Wideband, 2009. ICUWB 2009. IEEE International Conference on*, pp. 1–5, Sept, 2009.

[37] A. Maltsev, R. Maslennikov, A. Maltsev, A. Khoryaev, and M. Shilov, *Performance analysis of spatial reuse mode in millimeter-wave wpan systems with multiple links*, in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, pp. 1–4, Sept, 2008.

[38] T. Bai and R. W. H. Jr., *Coverage and rate analysis for millimeter wave cellular networks*, *CoRR* **abs/1402.6430** (2014).

[39] B. Langen, G. Lober, and W. Herzig, *Reflection and transmission behaviour of building materials at 60 ghz*, in *Personal, Indoor and Mobile Radio Communications, 1994. Wireless Networks - Catching the Mobile Future., 5th IEEE International Symposium on*, pp. 505–509 vol.2, Sep, 1994.

[40] Z. Marzi, U. Madhow, and H. Zheng, *Interference analysis for mm-wave picocells*, in *Global Communications Conference (GLOBECOM), 2015 IEEE*, pp. 1–6, IEEE, 2015.

[41] M. Sauter, *3G, 4G and beyond: Bringing networks, devices and the web together*. John Wiley & Sons, 2013.

[42] Wikipedia contributors, "Demographics of new york city — Wikipedia, the free encyclopedia." `https://en.wikipedia.org/w/index.php?title=Demographics_of_New_York_City&oldid=858444014`, 2018. [Online; accessed 11-September-2018].

[43] A. Ghosh, T. Thomas, M. Cudak, R. Ratasuk, P. Moorut, F. Vook, T. Rappaport, G. MacCartney, S. Sun, and S. Nie, *Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks*, *Selected Areas in Communications, IEEE Journal on* (June, 2014).

[44] D. Berraki, S. Armour, and A. Nix, *Application of compressive sensing in sparse spatial channel recovery for beamforming in mmwave outdoor systems*, in *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, April, 2014.

[45] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, *A simple proof of the restricted isometry property for random matrices*, *Constructive Approximation* **28** (2008), no. 3 253–263.

[46] D. Achlioptas, *Database-friendly random projections*, in *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 274–281, ACM, 2001.

[47] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. 2007.

[48] S. Basu and Y. Bresler, *A global lower bound on parameter estimation error with periodic distortion functions*, Information Theory, IEEE Transactions on (2000).

[49] M. E. Rasekh and U. Madhow, *Noncoherent compressive channel estimation for mm-wave massive mimo*, arXiv preprint arXiv:1801.06608 (2018).

[50] Z. Marzi, D. Ramasamy, and U. Madhow, *Compressive channel estimation and tracking for large arrays in mm-wave picocells*, IEEE Journal of Selected Topics in Signal Processing **10** (2016), no. 3 514–527.

[51] A. Alkhateeb, O. El Ayach, G. Leus, and R. Heath, *Single-sided adaptive estimation of multi-path millimeter wave channels*, in *Signal Processing Advances in Wireless Communications (SPAWC), 2014 IEEE 15th International Workshop on*, June, 2014.

[52] T. Bogale and L. B. Le, *Beamforming for multiuser massive mimo systems: Digital versus hybrid analog-digital*, in *Global Communications Conference (GLOBECOM), 2014 IEEE*, pp. 4066–4071, Dec, 2014.

[53] C. Sheldon, M. Seo, E. Torkildson, M. Rodwell, and U. Madhow, *Four-channel spatial multiplexing over a millimeter-wave line-of-sight link*, in *Microwave Symposium Digest, 2009. MTT'09. IEEE MTT-S International*, pp. 389–392, IEEE, 2009.

[54] G. Wang and G. Ascheid, *Joint pre/post-processing design for large millimeter wave hybrid spatial processing systems*, in *European Wireless 2014; 20th European Wireless Conference; Proceedings of*, May, 2014.

[55] J. Singh and S. Ramakrishna, *On the feasibility of beamforming in millimeter wave communication systems with multiple antenna arrays*, in *Global Communications Conference (GLOBECOM), 2014 IEEE*, Dec, 2014.

[56] O. Ayach, R. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, *Low complexity precoding for large millimeter wave mimo systems*, in *Communications (ICC), 2012 IEEE International Conference on*, June, 2012.

[57] O. El Ayach, R. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, *The capacity optimality of beam steering in large millimeter wave mimo systems*, in *Signal Processing Advances in Wireless Communications (SPAWC), 2012 IEEE 13th International Workshop on*, June, 2012.

[58] H. Zhang and U. Madhow, *Statistical modeling of fading and diversity for outdoor 60 ghz channels*, in *Proceedings of the 2010 ACM international workshop on mmWave communications: from circuits to networks*.

[59] H. Zhang, S. Venkateswaran, and U. Madhow, *Channel modeling and mimo capacity for outdoor millimeter wave links*, in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*.

[60] P. White and G. L. Reil, *Millimeter-wave beamforming: Antenna array design choices & characterization white paper*, .

[61] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, *Millimeter-wave vehicular communication to support massive automotive sensing*, *IEEE Communications Magazine* **54** (2016), no. 12 160–167.

[62] C. Perfecto, J. Del Ser, M. Bennis, and M. N. Bilbao, *Beyond wysiwyg: Sharing contextual sensing data through mmwave v2v communications*, in *Networks and Communications (EuCNC), 2017 European Conference on*, pp. 1–6, IEEE, 2017.

[63] M. Ruble and İ. Güvenç, *Wireless localization for mmwave networks in urban environments*, *EURASIP Journal on Advances in Signal Processing* **2018** (2018), no. 1 35.

[64] F. Lemic, J. Martin, C. Yarp, D. Chan, V. Handziski, R. Brodersen, G. Fettweis, A. Wolisz, and J. Wawrzynek, *Localization as a feature of mmwave communication*, in *Wireless Communications and Mobile Computing Conference (IWCMC), 2016 International*, pp. 1033–1038, IEEE, 2016.

[65] R. Mendrzik, H. Wymeersch, and G. Bauch, *Joint localization and mapping through millimeter wave mimo in 5g systems-extended version*, *arXiv preprint arXiv:1804.04417* (2018).