MODELING AND MITIGATING BEAM SQUINT IN MILLIMETER WAVE WIRELESS COMMUNICATION

A Dissertation

Submitted to the Graduate School of the University of Notre Dame in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy by Mingming Cai

J. Nicholas Laneman, Director

Graduate Program in Electrical Engineering Notre Dame, Indiana March 2018
MODELING AND MITIGATING BEAM SQUINT IN MILLIMETER WAVE WIRELESS COMMUNICATION

Abstract

by

Mingming Cai

There has been increasing demand for accessible radio spectrum with the rapid development of mobile wireless devices and applications. For example, a GHz of spectrum is needed for fifth-generation (5G) cellular communication, but the available spectrum below 6 GHz cannot meet such requirements. Fortunately, spectrum at higher frequencies, in particular, millimeter-wave (mmWave) bands, can be utilized through phased-array analog beamforming to provide access to large amounts of spectrum. However, the gain provided by a phased array is frequency dependent in the wideband system, an effect called beam squint. We examine the nature of beam squint and develop convenient models with either a uniform linear array (ULA) or a uniform planar array (UPA). Analysis shows that beam squint decreases channel capacity, and therefore, path selection should take beam squint into consideration. Current channel estimation algorithms assume no beam squint, and channel estimation error is increased by the beam squint, further decreasing the channel capacity.

Three problems involving phased-array beamforming are studied to incorporate and compensate for beam squint. First, we show that carrier aggregation can be used to improve system throughput to a point. We study the optimal beam alignment to maximize channel capacity, and demonstrate that, with sufficient band separation, focusing on only one band outperforms carrier aggregation. Approximations are de-
veloped for a system with two symmetric bands to determine the critical values of system parameters like band separation, and angle of arrival beyond which it is preferable not to aggregate. Second, beamforming codebooks are designed to compensate for one-sided beam squint by imposing a channel capacity constraint. Analysis and numerical examples suggest that a denser codebook is required compared to the case without beam squint, and the codebook size increases as bandwidth or the number of antennas in the array increases and diverges as either of these parameters exceeds certain limits. Third, to decouple the transmitter and receiver arrays with two-sided beam squint, and to extend conventional codebook design algorithms, codebooks with a minimum array gain constraint for all frequencies and angles of arrival or departure are also designed to compensate for the beam squint. Again, either the bandwidth or the number of antennas in the array is limited by the effects of beam squint if the other one is fixed.
To the memory of my mom and grandpa
CONTENTS

FIGURES ................................................................. vi

TABLES ................................................................. ix

ACKNOWLEDGMENTS ................................................ x

SYMBOLS ................................................................. xi

CHAPTER 1: INTRODUCTION .............................................. 1
  1.1 Overview ......................................................... 1
  1.2 Contributions ................................................... 4
  1.3 Outline ........................................................... 5

CHAPTER 2: BACKGROUND ............................................... 6
  2.1 Architectures for MmWave Beamforming ......................... 6
    2.1.1 Analog Beamforming ....................................... 6
    2.1.2 Hybrid Beamforming ....................................... 7
    2.1.3 Digital Beamforming ....................................... 8
    2.1.4 Phase and Beam Quantization .............................. 9
  2.2 Beam Squint ...................................................... 9
    2.2.1 System Model of the Receiver Array ...................... 11
    2.2.2 Beam Squint Reduction and Elimination .................. 15
  2.3 Channel Models ................................................ 17
    2.3.1 Passband Propagation Model .............................. 18
    2.3.2 Passband Array Model ..................................... 19
    2.3.3 Passband Beamforming Model ............................. 20
    2.3.4 Discrete-Time Equivalent Channel Model ................ 22
  2.4 Related Topics ................................................ 24
    2.4.1 Initial Access and Channel Estimation .................. 24
    2.4.2 Carrier Aggregation ...................................... 26
    2.4.3 Beamforming Codebook Design ............................ 27
  2.5 Summary ........................................................ 28
CHAPTER 3: MODELING OF BEAM SQUINT

3.1 ULA with Beam Squint
3.2 UPA with Beam Squint
3.3 Discrete-Time Equivalent Channel Model with Beam Squint
3.4 Summary

CHAPTER 4: EFFECTS OF BEAM SQUINT

4.1 Effect on Channel Capacity
4.2 Effect on Path Selection
4.3 Effect on Channel Estimation
   4.3.1 Channel Estimation with MLE
   4.3.2 Channel Estimation of ULA
4.4 Summary

CHAPTER 5: CARRIER AGGREGATION WITH BEAM SQUINT

5.1 System Setup
5.2 Carrier Aggregation with One-Sided Beam Squint
   5.2.1 Numerical Optimization
   5.2.2 Carrier Aggregation of Two Symmetric Bands
   5.2.3 To Aggregate or Not to Aggregate
      5.2.3.1 Increasing Band Separation
      5.2.3.2 Approximating the Critical Band Separation
      5.2.3.3 Approximating the Critical AoA
      5.2.3.4 Criterion for Whether or Not to Aggregate
      5.2.3.5 Numerical Results
5.3 Carrier Aggregation with Two-Sided Beam Squint
   5.3.1 Numerical Optimization
   5.3.2 Carrier Aggregation of Two Symmetric Bands
   5.3.3 To Aggregate or Not to Aggregate
      5.3.3.1 Increasing Band Separation
      5.3.3.2 Criterion for Whether or Not to Aggregate
5.4 Summary

CHAPTER 6: BEAMFORMING CODEBOOK WITH CHANNEL CAPACITY CONSTRAINT

6.1 Definition of a Beamforming Codebook
6.2 Properties of the Smallest Codebook
6.3 Codebook Design
   6.3.1 Odd-Number Codebook
   6.3.2 Even-Number Codebook
   6.3.3 Codebook Design Algorithm
6.4 Constraint of the Fractional Bandwidth b
6.5 Numerical Results
FIGURES

2.1 Analog beamforming architecture. ........................................ 7
2.2 Hybrid beamforming architecture. ................................. 8
2.3 Digital beamforming architecture. ................................. 9
2.4 Structure of a uniform linear array (ULA) with analog beamforming using phase shifters. ................................. 10
2.5 Example of an angle response for a ULA. .......................... 13
2.6 Example of fine beam angle responses for different frequencies in a wideband system for ULA. ................................. 15
2.7 Reduction of Beam Squint with the Phase Improvement Scheme [32]. ................................. 16
2.8 Illustration of nested channel models. ............................. 17
3.1 Example of array gain for a fine beam. ............................ 31
3.2 Structure of a uniform planar array (UPA) with analog beamforming using phase shifters. ................................. 33
3.3 Examples of frequency-domain channel gains with and without beam squint for a ULA. ................................. 39
4.1 Characteristics of $|g(x)|^2$ from (3.7), its derivative, and its second-order derivative within the main lobe. ................................. 44
4.2 Capacities with and without beam squint as a function of bandwidth. ................................. 47
4.3 The maximum and minimum relative capacity loss across the 3 dB beamwidth in a fine beam as a function of beam focus angle. ................................. 48
4.4 Maximum relative capacity loss among all fine beams as a function of fractional bandwidth ................................. 49
4.5 Example scenario of path selection with consideration of beam squint. ................................. 51
4.6 Channel capacities from path 1 and path 2 as function of AoA and AoD. ................................. 52
4.7 Example of MLE channel estimation with beam squint in the receiver. ................................. 56
4.8 MSE in a fine beam as a function of AoA for fixed beam focus angle. ................................. 58
4.9 The maximum MSE across 3 dB beamwidth in a fine beam as a function of beam focus angle. ................................. 59
4.10 The minimum MSE across 3 dB beamwidth in a fine beam as a function of beam focus angle.

4.11 The maximum MSE across 3 dB beamwidth and beam focus angle range.

5.1 Example of aggregated channel capacities as functions of the band separation.

5.2 Sub-optimal beam focus angles for two-band carrier aggregation obtained from numerical optimization and Approximation 1 as functions of band separation.

5.3 Maximum channel capacity for two-band carrier aggregation obtained from numerical optimization and Approximation 1 as functions of band separation.

5.4 Critical band separations for two-band carrier aggregation of numerical optimization and Approximation 2 as functions of SNR.

5.5 Example of channel capacities for two-band carrier aggregation with two-sided beam squint as functions of the band separation.

5.6 Example of channel capacities for two-band carrier aggregation with two-sided beam squint as functions of the band separation.

6.1 Upper bound of the fractional bandwidth as a function of the number of antennas.

6.2 Capacity improvement ratio in a fine beam as a function of beam focus angle.

6.3 Maximum capacity improvement ratio as a function of fractional bandwidth.

6.4 Minimum codebook size in a ULA as a function of the number of antennas.

7.1 Beamforming codebook design procedure with array gain constraint.

7.2 Example of beamforming codebook design for the minimum codebook size subject to array gain constraint.

7.3 Generalized beam edge as a function of the codebook size for various fractional bandwidths.

7.4 Optimized codebook array gain as a function of the codebook size for various fractional bandwidths.

7.5 Optimized codebook array gain grows as codebook size increases.

7.6 Minimum codebook size as a function of fractional bandwidth.

7.7 Maximum relative approximation error as a function of the number of antennas.

7.8 Minimum codebook size as a function of the number of antennas.
7.9 Optimized codebook array gain as a function of the number of antennas \[126\]

A.1 Relation between odd codebook size and even codebook size. . . . . . \[139\]
**TABLES**

1.1 COMPARISON BETWEEN SPECTRUM ACCESS BELOW 6 GHZ AND AT MMWAVE  

7.1 SUMMARY OF CODEBOOK PARAMETERS
ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude to my adviser and mentor, Dr. J. Nicholas Laneman. His support, encouragement, guidance, and patience have been indispensable part of this dissertation as well as my progress during my graduate study. Without his long-time words and deeds to help me improve my communication skills, I could not have won the People’s Choice Award in the Three Minute Thesis (3MT) competition of the graduate school in March 2017. I will always remember the whole night he stayed with me to improve algorithms and program for DARPA Spectrum Challenge.

I would also like to thank Dr. Bertrand Hochwald. I have benefited significantly from his vision and experiences. He gave me many good suggestions not only on the radio design and implementation, but also on my career.

My thanks also go to Ding Nie and Kang Gao. They are very smart and we have been working closely to make steady progress in our projects. I would like to thank Neil Dodson, as well. He substantially helped on the setup of lab environments. Zhanwei Sun, Nikolaus Kleber, Christine Landaw, Cindy Fuja, Jun Chen, Sahand Golnarian, Hamed Pezeshki, Abbas Termos, Chao Luo, Zhenhua Gong, Zhen Tong, Glenn Bradford, and Ebrahim MolavianJazi also provided me with good advice in my graduate study. I would like to thank them all.

Last, but not least, I would like to express my gratitude to all my friends and my family for their support.
SYMBOLS

\[ \beta \] phase shift
\[ \theta \] angle of arrival or departure in uniform linear array (ULA),
or azimuth angle in uniform planar array (UPA)
\[ \psi \] virtual angle of \( \theta \)
\[ \phi \] elevation angle
\[ \xi \] ratio of subcarrier frequency to carrier frequency
\[ \lambda \] wavelength
\[ \sigma^2/2 \] spectral density of stationary white Gaussian noise
\[ \Omega \] angle of departure (AoD) or angle of arrival (AoA) in a UPA
\[ \Omega_T \] AoD in a UPA
\[ \Omega_R \] AoA in a UPA
\[ \Psi \] virtual AoD or AoA in a UPA
\[ \Theta \] channel impulse response matrix for signal between Tx and Rx RF chains
\[ \mathbf{a} \] response vector of a ULA
\[ b \] fractional bandwidth
\[ c \] speed of light
\[ d \] distance between two adjacent antennas in a uniform array
\[ f \] frequency
\[ g \] array gain of ULA
\[ g_T \] transmitter array gain of ULA
\[ g_R \] receiver array gain of ULA
\[ h \] path gain
\textbf{r} received signal vector
\textbf{s} transmitted signal vector
\textbf{w} beamforming vector, i.e., weight vector of phase shifters
\textbf{w}_T \quad \text{transmitter beamforming vector, i.e., weight vector of phase shifters}
\textbf{w}_R \quad \text{receiver beamforming vector, i.e., weight vector of phase shifters}
\textbf{z} Gaussian noise vector
\textbf{B} bandwidth
\textbf{C} channel capacity
\textbf{C} codebook
\textbf{D} relative capacity loss
\textbf{G} array gain of UPA
\textbf{G}_T \quad \text{transmitter array gain of UPA}
\textbf{G}_R \quad \text{receiver array gain of UPA}
\textbf{H} frequency-domain multipath channel matrix between Tx\&Rx phase shifters
\textbf{\mathcal{H}} \quad \text{frequency-domain channel gain vector between Tx\&Rx RF chains}
\textbf{I} capacity improvement
\textbf{L} discrete delay spread
\textbf{M} codebook size
\textbf{N} number of antennas in a ULA
\textbf{N}_a \quad \text{number of antennas in a UPA}
\textbf{N}_f \quad \text{number of subcarriers in an OFDM symbol}
\textbf{N}_h \quad \text{number of antennas in horizontal direction of a UPA}
\textbf{N}_l \quad \text{number of paths in the channel model}
\textbf{N}_p \quad \text{number of pilots in an OFDM symbol}
\textbf{N}_v \quad \text{number of antennas in vertical direction of a UPA}
CHAPTER 1

INTRODUCTION

1.1 Overview

Applications of wireless technologies and the number of wireless devices have been growing significantly in the past decade. For example, the global mobile data traffic grew 63 percent in 2016, and it is expected to grow at a compound annual growth rate of 47 percent over the next five years [13]. There are increasing demands for radio spectrum to meet the requirements of higher data rates for an increasing number of mobile devices. As noted in a report from the President’s Council of Advisors on Science and Technology (PCAST), access to spectrum will be an increasingly important foundation for economic growth and technological leadership in the coming years [42]. However, the current process of long-term, static frequency allocation below 6 GHz cannot meet these demands for mobile spectrum.

Two major directions are being explored for enhancing spectrum access, namely, reusing currently under-utilized spectrum through dynamic spectrum access (DSA), or spectrum sharing, and enabling access to spectrum at higher frequencies, e.g., millimeter-wave (mmWave) bands [7, 8, 18, 19, 38]. Although both directions require progress on interesting technical issues, this dissertation focuses on enabling spectrum access in mmWave bands based upon, on the one hand, its potential for wide bandwidths and less congested airwaves and, on the other hand, DSA’s ongoing regulatory and market uncertainties despite over a decade of significant research effort.
TABLE 1.1

COMPARISON BETWEEN SPECTRUM ACCESS BELOW 6 GHz AND AT MMWAVE

<table>
<thead>
<tr>
<th></th>
<th>Spectrum Below 6 GHz</th>
<th>mmWave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrum Available</td>
<td>Hundreds of MHz</td>
<td>Tens of GHz</td>
</tr>
<tr>
<td>Signal Attenuation</td>
<td>Relatively small</td>
<td>Relatively large</td>
</tr>
<tr>
<td>Range / Coverage</td>
<td>Relatively large</td>
<td>Relatively small</td>
</tr>
<tr>
<td>Radio Cost</td>
<td>Relatively low</td>
<td>Relatively high</td>
</tr>
<tr>
<td>Regulatory Timelines</td>
<td>Relatively long</td>
<td>Relatively short</td>
</tr>
</tbody>
</table>

Table 1.1 summarizes the advantages and disadvantages of each of the two directions.

There are several key features of mmWave bands that make them very attractive as the next frontier in wireless technology development. First, the mmWave band ranges from 30 GHz to 300 GHz, with spectrum opportunities on the order of tens of GHz wide. In a loose sense, the 20-30 GHz band is also considered part of the mmWave band, further increasing the amount of available spectrum. By contrast, the total amount of spectrum available for sharing below 6 GHz is limited, on the order of hundreds of MHz. Second, signals at mmWave frequencies experience higher path loss and atmospheric absorption, typically 20 dB or more attention, than those below 6 GHz. Therefore, mmWave systems often require higher antenna directional-ity and have comparatively smaller range or coverage than systems operating below 6 GHz. Third, relatively sparse deployments currently in mmWave bands leave much of the spectrum open to new entrants and simpler interference avoidance mechanisms, whereas the more congested spectrum below 6 GHz presents many technical
challenges for deployment of spectrum sharing such as spectrum sustainability and hidden nodes [23, 51, 59]. Fourth, commercially viable radio hardware for mmWave access is less mature than radio hardware operating below 6 GHz, and the device cost is usually higher. As development of device technologies for mmWave accelerates, the cost gap should close rapidly.

A promising method to compensate for the higher attenuation in mmWave bands is beamforming [26]. To achieve high directional gain, either a large physical aperture or a phased-array antenna is employed [25, 27]. The energy from the aperture or multiple antennas is focused on one direction or a small set of directions. The cost of a large physical aperture is relatively high, especially in terms of installation and maintenance. Fortunately, the short wavelengths of mmWave frequencies in principle allows for integration of a large number of antennas into a small phased array, which would be suitable for application in commercial mobile devices. In mmWave beamforming, a phased array with a large number of antennas can compensate for the higher attenuation. In this dissertation, we consider practical implementations of mmWave analog beamforming with one radio frequency (RF) chain and a phased array employing phase shifters.

Traditionally, in a communication system with analog beamforming, the set of phase shifter values in a phased array are designed at a specific frequency, usually the carrier frequency, but applied to all frequencies within the transmission bandwidth due to practical hardware constraints. Phase shifters are relatively good approximations to the ideal time shifters for narrowband transmission; however, this approximation breaks down for wideband transmission if the angle of arrival (AoA) or angle of departure (AoD) is far from the broadside because the required phase shifts are frequency-dependent. The net result is that beams for frequencies other than the carrier “squint” as a function of frequency in a wide signal bandwidth [34]. This phenomenon is called beam squint [21]. As we will see, beam squint translates
into array gains for a given AoA or AoD that vary with frequency, and becomes more pronounced if either the number of antennas in the array or the bandwidth increases.

1.2 Contributions

This dissertation models beam squint, illustrates its effects on channel capacity and estimation, and develops approaches to mitigate it. Tradeoffs are provided among system parameters such as beam focus angle, bandwidth, number of antennas, and beamforming codebook size. The main contributions of the dissertation are summarized as follows.

- **Modeling beam squint.** We model beam squint in a convenient way for a uniform linear array (ULA) and a uniform planar array (UPA). Throughout this dissertation, beam squint is incorporated into the discrete-time equivalent channel model whereas most current literature ignores beam squint in the channel model.

- **Effects of beam squint.** The array gain varies over frequency due to beam squint, and this variation reduces channel capacity compared to the case without beam squint. Our analysis and numerical results show that the reduction in channel capacity increases with the number of antennas in the array, the bandwidth, or the beam focus angle. With beam squint affecting multiple paths for different AoD / AoA, a path with higher maximum array gain may actually provide a lower throughput. Therefore, we propose using channel capacity as the performance metric for path selection. Finally, current channel estimation algorithms do not consider beam squint, and we illustrate that beam squint increases channel estimation error.

- **Carrier aggregation with beam squint.** In mmWave bands, multiple non-contiguous bands or carriers could be aggregated to provide more spectrum and therefore potentially higher channel capacity. We study carrier aggregation for mmWave considering beam squint, and we determine beam focus angles to maximize the total channel capacity. Specifically, we analyze the carrier aggregation problem for the case of two symmetric bands and provide design criteria based upon, among others, the band separation for determining if the beam should focus on the center of the two bands or if the beam should focus on only one of the bands.

- **Beamforming codebooks with channel capacity constraint.** In switched beamforming, the transmitter and / or the receiver select a set of phase shifter values from pre-designed collection of such sets, the collection being called a
beamforming codebook. We introduce the problem of beamforming codebook
design with beam squint subject to a constraint on the minimum channel ca-
pacity, and focus on the case of one-sided beam squint since this metric couples
the transmitter and receiver design problems otherwise. The effects of beam
squint increase the required codebook size, and substantially so as the number
of antennas or the bandwidth increases.

• **Beamforming codebooks with array gain constraint.** To be consistent
with the majority of beamforming codebook research, we also design beam-
forming codebooks to compensate for beam squint subject to a minimum array
gain for a desired angle range and for all frequencies in the wideband system.
As we will see, beam squint with this design criterion fundamentally limits the
bandwidth or the number of antennas of the array if the other one is fixed.

1.3 Outline

An outline of the remainder of this dissertation is as follows. Chapter 2 provides
a detailed background on mmWave beamforming. Specifically, current architectures
for mmWave beamforming, channel models, channel estimation, and path selection
techniques are summarized along with numerous references. We also qualitatively
illustrate beam squint, and we summarize expensive hardware approaches to reduce
or eliminate its effects. Chapter 3 develops models for beam squint for a ULA and
a UPA, resulting in a baseband-equivalent models that incorporate beam squint.
Chapter 4 illustrates the effects of beam squint on wireless communication systems,
including effects on channel capacity, path selection, and channel estimation. Chap-
ter 5 studies carrier aggregation with beam squint at mmWave. Chapter 6 develops
a beamforming codebook design algorithm with channel capacity constraint. Chap-
ter 7 designs beamforming codebooks with a minimum array gain constraint. Finally,
Chapter 8 summarizes conclusions of the dissertation as well as directions for future
research.
CHAPTER 2
BACKGROUND

In this chapter, we review existing work on mmWave beamforming, including architectures for mmWave beamforming, beam squint, channel models, initial access, channel estimation, carrier aggregation, and beamforming codebook design.

2.1 Architectures for MmWave Beamforming

To compensate for high path loss and to achieve high directional gain, beamforming is employed in mmWave bands [25, 27]. Fortunately, the short wavelength of mmWave allows integrating a large number of antennas into a small phased array, which is suitable for the application of mobile devices.

There are three basic architectures for mmWave beamforming, including analog beamforming, hybrid beamforming and digital beamforming.

2.1.1 Analog Beamforming

Figure 2.1 illustrates the architecture of analog beamforming. Analog beamforming is implemented by a phased array with only RF chain driven by a digital-to-analog converter (DAC) in the transmitter or an analog-to-digital converter (ADC) in the receiver. A transmitter RF chain consists of frequency up converter, power amplifier and so on; a receiver RF chain consists of low-noise amplifier, frequency down converter and so on [5, 53].

The antenna weights in the phased array are constrained to be phase shifts that can be controlled digitally. The phases of the phase shifters are typically quantized
to limited resolution, and there is no ability to adjust the relative amplitudes of the signal fed into the antennas in the transmitter. The resulting transmit signal is constructive in some directions and destructive in other directions, forming a beam. The phases of the phase shifters can be dynamically adjusted based on specific strategies to steer the beam. Similar characteristics apply to the receiver.

2.1.2 Hybrid Beamforming

The architecture of the hybrid beamforming is shown in Figure 2.2. On the transmitter side, there could be multiple data streams as the inputs to the Baseband Precoding block. Multiple DACs and RF chains are used and each of their outputs is combined to all or some of the antennas through phases shifters or switches in the RF Precoding Block [36]. On the receiver side, each signal from the antenna is connected to multiple phase shifters or switches, each for one RF chain and corresponding ADC. The signals from all or some of the antennas are combined to each RF chain separately. Similarly, there could be multiple data streams from the outputs of the Baseband Decoding Block. An alternative way to implement hybrid beamforming is using a large physical aperture [25, 27]. However, the cost of large physical apert-
ture is relatively high, especially in terms of installation and maintenance. Hybrid beamforming can support multiple users and multiple data streams; however, the implementation complexity and cost are higher than those of analog beamforming.

2.1.3 Digital Beamforming

The architecture of digital beamforming is illustrated in Figure 2.3. Each antenna is connected to its own RF chain and either a DAC or an ADC, making digital beamforming more flexible than analog beamforming and hybrid beamforming in terms of signal processing. However, there are three disadvantages of digital beamforming for mmWave. First, the array for mmWave needs to be packed into a small area, precluding a complete RF chain for each antenna. Second, the electronic components in each RF chain have large power consumption. Third, there could be Gigabits of data for each RF chain to process in a single second, and such high total data rate from all RF chains is a challenge for current baseband signal processing hardware.

Analog beamforming is more appropriate for mobile mmWave applications be-
cause of its low cost and low complexity. Consequently, in this dissertation, we focus on analog beamforming.

2.1.4 Phase and Beam Quantization

Based on whether or not the beam can be continuously adjusted in space, beamforming can be divided into two categories: continuous beamforming and switched beamforming. As implied by the name, beams in continuous beamforming can be adjusted continuously in space. In contrast, beams in switched beamforming can only be adjusted to focus on a finite set of angles. To cover a certain range of AoA/AoD in space, a beamforming codebook is usually used in switched beamforming \[50, 55\]. A beamforming codebook consists of multiple beams, with each beam determined by a set of beamforming phases. Each beam is a codeword of the codebook. The transmitter or receiver can only use one beam in its codebook at a time.

2.2 Beam Squint

We consider a ULA with \(N\) identical and isotropic antennas as shown in Figure 2.4.
Figure 2.4. Structure of a uniform linear array (ULA) with analog beamforming using phase shifters. There are $N$ antennas, labeled as $1, 2, 3, \ldots, N$. The distance between adjacent antennas is $d$, and $\theta$ denotes either the angle-of-arrival (AoA) for reception or the angle-of-departure (AoD) for transmission. The additional distance traveled by the electromagnetic wavefront from the first antenna to the $n$th antenna is denoted $\Delta d_n$. 
The $N$ antennas are labeled as $1, 2, 3, \ldots, N$. Each antenna is connected to a phase shifter. The distances between two adjacent antenna elements are the same and all denoted as $d$. For simplicity, we also assume the phases of the phase shifters are continuous without quantization. The AoA or AoD $\theta$ is the angle of the signal relative to the broadside of the array, increasing counterclockwise, and $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. This structure could represent the phased array of either a transmitter or a receiver, and their system analysis is similar. We focus on the receiver array for example.

2.2.1 System Model of the Receiver Array

Suppose the signal arriving at the array is $s(t)$ as shown in Figure 2.4. The signal received by the $n$th antenna is denoted as $y_n(t)$. Then the received signal vector $y(t)$ before the phase shifters is

$$y(t) = [y_1(t), y_2(t), \ldots, y_N(t)]^T,$$

where superscript $(\cdot)^T$ indicates transpose of a vector. As shown in Figure 2.4, the additional distance traveled by the signal from the first antenna to the $n$th antenna is denoted $\Delta d_n$, where

$$\Delta d_n = (n - 1) d \sin \theta, \quad n = 1, 2, 3, \ldots, N,$$

and $\Delta d_1 = 0$. Then $y(t)$ can be denoted in terms of $s(t)$ by

$$y(t) = \left[ s \left( t - \frac{\Delta d_1}{c} \right), s \left( t - \frac{\Delta d_2}{c} \right), \ldots, s \left( t - \frac{\Delta d_n}{c} \right), \ldots, s \left( t - \frac{\Delta d_N}{c} \right) \right]^T$$

$$= \left[ s(t), \ldots, s \left( t - \frac{(n - 1) d \sin \theta}{c} \right), \ldots, s \left( t - \frac{(N - 1) d \sin \theta}{c} \right) \right]^T,$$
where $c$ is the speed of light. The frequency-domain received signal vector $y(f, \theta)$ is then
\[
y(f, \theta) = [s(f), \ldots, s(f)e^{j2\pi c^{-1}f(n-1)d\sin \theta}, \ldots, s(f)e^{j2\pi c^{-1}f(N-1)d\sin \theta}]^T
\]
where $y(f, \theta)$ is a function of AoA $\theta$, and $\theta \in [-\pi/2, \pi/2]$. Define the ULA response vector as
\[
a(\theta, f) = [1, e^{j2\pi c^{-1}fd\sin \theta}, \ldots, e^{j2\pi c^{-1}f(n-1)d\sin \theta}, \ldots, e^{j2\pi c^{-1}f(N-1)d\sin \theta}]^T,
\]
where $\theta \in [-\pi/2, \pi/2]$ [40]. $a(f, \theta)$ is frequency dependent. Then
\[
y(f, \theta) = s(f)a(\theta, f).
\]
The optimal combiner of the signal $y(f, \theta)$ should be a filter matched to $a(f, \theta)$ defined as
\[
h_{MF}(f, \theta) = a^*(\theta, f)
\]
\[
= [1, e^{-j2\pi c^{-1}fd\sin \theta}, \ldots, e^{-j2\pi c^{-1}f(n-1)d\sin \theta}, \ldots, e^{-j2\pi c^{-1}f(N-1)d\sin \theta}]^T,
\]
where superscript $(\cdot)^*$ denotes complex conjugate. The elements in the matched filter $h_{MF}(f, \theta)$ are true-time-delay devices [33] [46]. However, the cost of true-time-delay devices is high, and there are also some implementation issues at mmWave band, which will be discussed in Section 2.2.2 In practice, phase shifters are used to approximate the optimal matched filter $h_{MF}(f, \theta)$. A phase shifter is usually modeled as constant phase shift for all frequency within
its designed range \[1, 21, 34\]. The \(n\)th phase shifter is connected to the \(n\)th antenna with phase shift denoted by \(\beta_n\). Define the corresponding beamforming vector as

\[
w = [e^{j\beta_1}, \ldots, e^{j\beta_n}, \ldots, e^{j\beta_N}]^T.
\] (2.8)

We note that \(w\) is frequency independent.

Following [4], the array gain of the phased-array receiver with AoA \(\theta\) and beamforming vector \(w\) is

\[
g(w, \theta, f) = \frac{1}{\sqrt{N}}w^H a(\theta, f) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{j[2\pi c^{-1}f(n-1)d\sin\theta - \beta_n]}.
\] (2.9)

where superscript \((\cdot)^H\) denotes Hermitian transpose. Each beamforming vector \(w\) forms a beam in space. Among all the possible beams, we are interested in the ones that have the highest array gain, because the main purpose of beamforming at mmWave is to compensate for the high signal attenuation at mmWave bands.

The set of phase shifters in the phased array is designed for a specific frequency, usually the carrier frequency, denoted as \(f_c\). Define the beam focus angle, denoted as \(\theta_F\), as the AoA/AoD with the highest array gain for the carrier frequency. To achieve the highest array gain for the carrier frequency and beam focus angle \(\theta_F\), the phase shifters should follow

\[
\beta_n(\theta_F) = 2\pi c^{-1}f_c(n-1)d\sin\theta_F, \quad n = 1, 2, \ldots, N.
\] (2.10)

Note that (2.10) corresponds to the matched filter (2.7) evaluated at \(\theta = \theta_F\) and \(f = f_c\). We call the beam with phase shifts described in (2.10) a fine beam, and we focus on the analysis of fine beams throughout this dissertation.

Based on (2.9) and (2.10), the array gain for a signal with frequency \(f\) and AoA
Figure 2.5. Example of an angle response for a ULA with beam focus angle $\theta_F = 0$. The number of antennas in the array $N = 16$. The antenna spacing is half of the wavelength corresponding to the carrier frequency $f_c = 73$ GHz.

The array gain using beam focus angle $\theta_F$ is

$$g (\theta_F, \theta, f) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{j[2\pi c^{-1}(n-1)d(f\sin\theta-f_c\sin\theta_F)]}.$$  \hspace{1cm} (2.11)

Figure 2.5 shows an example of a fine beam angle response $g(0, \theta, f_c)$ at the carrier frequency for a ULA with 16 antennas. From the figure, the array only exhibits high gain for a small angle range. The fine beam acts as spatial filter so that the array gains for the other angles are relatively small.

In general, the array response (2.11) varies with frequency across a wide bandwidth, which is the effect called beam squint [21, 34]. This results from the constant phase shift in the frequency domain of the fine beam corresponding to (2.10) instead of the linear phase shift in the ideal matched filter (2.7). Figure 2.6 illustrates beam squint by example, where the fine beam angle responses $g(\pi/6, \theta, f_{\min})$, $g(\pi/6, \theta, f_c)$, and $g(\pi/6, \theta, f_{\max})$ differ for signals with the same AoA/AoD $\theta$. In particular, the
array gains for $\theta = \pi/6$ are smaller at both $f_{\min}$ and $f_{\max}$ compared to the gain at the carrier frequency $f_c$.

### 2.2.2 Beam Squint Reduction and Elimination

Several hardware-based approaches have been developed to reduce or eliminate beam squint. In [32], a phase improvement scheme shown in Figure 2.7 is proposed to reduce the effect of beam squint. In this scheme, banks of bandpass filters at mmWave are used to separate signals into smaller subbands; additional phase shifters are added to each subband to reduce the beam squint of each subband. Simulations show that the beam squint can be reduced. However, the signal components at the boundary of two subbands can be significantly distorted, and the bandpass filters increase cost.

To eliminate beam squint, true time-delay (TTD) devices can be applied [33, 46]. There are two basic methods to implement TTD: optical and electronic.
• In optical methods, RF signals are first modulated into optical signals, then the optical signals are delayed by long optical fibers to introduce time delay \[35, 46\], and finally the delayed optical signals are demodulated into RF signals again. To obtain different delays, multiple optical fibers are combined based on combination algorithms, such as binary fiber-optic delay line \[20\]. One disadvantage of optical TTD is its poor RF performance of the modulator and detector, such as 40 dB insertion loss \[31\]. Another issue is its large size and high cost, preventing mobile applications.

• Electronic methods include coaxial cable, Micro-electro-mechanical Systems (MEMS), and monolithic microwave integrated circuits (MMICs) \[33\]. The weight of the TTD with coaxial cable is too large for mobile applications. MEMS TTD uses MEMS switches to combine the fabricated delay line on chip. The insertion loss is relatively small, such as 4 dB at 30 GHz \[29\]. MEMS is a promising technology for implementing TDD in mobile applications. However, it can fail with high power signals, a recent paper shows that a single MEMS TDD device has the size of tens of square millimeters, such as 4 mm \(\times\) 9 mm \[30\], and the insertion loss for MMICs can be as much as 25 dB at 20 GHz \[56\].

In summary, the approaches outlined above are unappealing for mobile wireless communication due to combinations of high implementation cost, significant insertion loss, large size, or excessive power consumption. In this dissertation, we focus on system-level awareness of the effects of beam squint and developing algorithms to compensate for its negative effects.
2.3 Channel Models

As shown in Figure 2.8, we consider four nested channel models for analog beamforming based on different encapsulated components. These channel models include three continuous-time models, namely, the passband propagation model, the passband array model, and the passband beamforming model, and one discrete-time equivalent channel model at baseband.

Suppose there are $N_T$ and $N_R$ antennas in the transmitter and receiver, respectively. The AoD is denoted as $\theta_T$, and the AoA is denoted as $\theta_R$. Throughout this section, we consider a ULA for illustration. The models can be easily extended to a UPA.
2.3.1 Passband Propagation Model

Propagation characteristics of mmWave signals are unique due to their relatively small wavelength compared to the objects in the environment \[25\]. Suppose there are \(N_l\) signal paths, labeled as 1, 2, \ldots, \(N_l\). The received signal of path \(l\) can be expressed as

\[
v_l(t) = g_l u_l(t - \tau_l), \quad l = 1, 2, \ldots, N_l,
\]

(2.12)

where \(u_l(t)\) is the sent signal of path \(l\), \(g_l\) is the channel gain of path \(l\), and \(\tau_l\) is the time delay of path \(l\). Suppose the power of \(u_l(t)\) is \(P_t\), and the power of \(v_l(t)\) is \(P_r\).

Based on Friis’ Law \[14\], if the signal is transmitted in free space,

\[
P_r = P_t G_t G_r \left(\frac{c}{4\pi f d_p}\right)^2,
\]

(2.13)

where \(G_t\) is transmitter antenna gain, \(G_r\) is receiver antenna gain, \(c\) is the speed of light, \(d_p\) is the distance between the transmitter and receiver antennas, and \(f\) is signal frequency. Therefore, the magnitude of \(g_l\) can be expressed as

\[
|g_l| = \frac{\sqrt{G_t G_r c}}{4\pi f d_p}.
\]

(2.14)

The Friis’ Law in dB scale can be expressed as

\[
10 \log_{10} \left(\frac{P_r}{P_t}\right) \text{[dB]} = 10 \log_{10} (P_t) \text{[dB]} + 10 \log_{10} \left(\frac{G_t G_r c^2}{16\pi^2 f^2}\right) - 20 \log_{10} (d_p).
\]

(2.15)

Friis’ Law indicates that the signal attenuation is proportional to the square of the frequency. Therefore, signals at mmWave experience much higher loss than signals below 6 GHz. This attenuation is one reason why beamforming is required at mmWave.
In practical environments, objects can attenuate, block or reflect signals. A significant amount of work has been performed to develop statistical models at mmWave for the distribution of path loss, especially on short-range links [48, 60]. From [25], the most commonly studied statistical model of path loss can be expressed as

\[
PL(d_p) [\text{dB}] = \alpha + 10 \gamma \log_{10}(d_p) + \kappa, \quad \kappa \sim \mathcal{N}(0, \sigma_0^2),
\]

where \(d_p\) is the distance between the transmitter and receiver antennas, \(\alpha\) is a constant, and \(\gamma\) is a log-normal random variable, and \(\sigma_0^2\) is variance of the path loss. Note that

\[
20 \log_{10}|g_t| = -PL(d_p) [\text{dB}].
\]

Friis’ model (2.15) is a special case of (2.16) with \(\gamma = 2\) almost surely.

2.3.2 Passband Array Model

As shown in Figure 2.8, the passband array model describes the relation between the transmitted signal vector \(\mathbf{x}(t)\) after phase shifters and the received signal vector \(\mathbf{y}(t)\) before phase shifters, where

\[
\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_{N_T}(t)]^T,
\]

\[
\mathbf{y}(t) = [y_1(t), y_2(t), \ldots, y_{N_R}(t)]^T.
\]

From (2.5), the ULA response vectors for the transmitter and receiver are [40]

\[
a_T(\theta_T, f) = \left[1, e^{j2\pi f d \sin \theta_T}, \ldots, e^{j2\pi f (N_T-1)d \sin \theta_T}\right]^T, \quad (2.20)
\]

\[
a_R(\theta_R, f) = \left[1, e^{j2\pi f d \sin \theta_R}, \ldots, e^{j2\pi f (N_R-1)d \sin \theta_R}\right]^T, \quad (2.21)
\]
respectively, where $\theta_T, \theta_R \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The arrays are designed based on the carrier frequency. Typically, the antenna spacing is $d = \lambda_c/2$, where $\lambda_c$ is the wavelength of the carrier frequency, i.e.,

$$\lambda_c = \frac{c}{f_c},$$

(2.22)

$f_c$ is the carrier frequency, and $c$ is the speed of light.

Doppler shift is the change of perceived signal frequency if the signal source moves towards or away from an observer. We assume the channel is slow fading so that the Doppler shifts of all paths are small and can be ignored [52]. Suppose path $l$ has AoD $\theta_{T,l}$, AoA $\theta_{R,l}$, time delay $\tau_l$, and complex path gain $g_l$. According to [49], the frequency response of the passband array model is

$$H(f) = \sum_{l=1}^{N_l} g_l a_R(\theta_{R,l}, f) a_T^H(\theta_{T,l}, f) e^{-j2\pi\tau_l f},$$

(2.23)

where $(\cdot)^H$ denotes conjugate transpose. With this frequency-response definition, we have $y(f) = H(f)x(f) + z(f)$, which corresponds in the time domain to

$$y(t) = \sum_{l=1}^{N_l} g_l a_R(\theta_{R,l}) a_T^H(\theta_{T,l}) x(t - \tau_l) + z(t),$$

(2.24)

where $z(t)$ is a noise vector that captures the effects of thermal noise and other interference.

### 2.3.3 Passband Beamforming Model

Phase shifters are part of the passband beamforming model. The effect of the phase shifters can be modeled by beamforming vector. The transmitter and receiver
beamforming vectors are

\[
\mathbf{w}_T = \left[ e^{j\beta_{T,1}}, \ldots, e^{j\beta_{T,n}}, \ldots, e^{j\beta_{T,N_T}} \right]^T,
\]

(2.25)

\[
\mathbf{w}_R = \left[ e^{j\beta_{R,1}}, \ldots, e^{j\beta_{R,n}}, \ldots, e^{j\beta_{R,N_R}} \right]^T,
\]

(2.26)

respectively, where \( \beta_{T,n} \) is the phase of the \( n \)th phase shifter connected to the \( n \)th antenna in the transmitter array, and \( \beta_{R,n} \) is the phase of \( n \)th phase shifter connected to the \( n \)th antenna in the receiver array. Again, the phase shifts \( \beta_{T,n}, n = 1, 2, \ldots, N_T \) and \( \beta_{R,n}, n = 1, 2, \ldots, N_R \) remain the same for all frequencies in the wideband of study due to hardware constraints [46].

Suppose \( s_C(t) \) is the transmitted RF signal before the phased array and \( r_C(t) \) is the received RF signal right after phased array as shown in Figure 2.8. Then

\[
r_C(f) = H_C(f) s_C(f) + z(f),
\]

(2.27)

where \( z(f) \) captures the effects of receiver thermal noise. From [4], the passband beamforming model can be expressed as

\[
H_C(f) = \frac{1}{\sqrt{N_R N_T}} \mathbf{w}_R^H \mathbf{H}(f) \mathbf{w}_T
\]

\[
= \frac{1}{\sqrt{N_R N_T}} \sum_{l=1}^{N_t} g_l \mathbf{w}_R^H \mathbf{a}_R(\theta_{R,l}, f) \mathbf{a}_T^H(\theta_{T,l}, f) \mathbf{w}_T e^{-j2\pi f l},
\]

(2.28)

where the array response vectors \( \mathbf{a}_T(\theta_T, f) \) and \( \mathbf{a}_R(\theta_R, f) \) are, in general, frequency-dependent. However, in current widely used channel models [25], the array response vectors are approximated to be constant across frequency. Specifically, the constant approximation often corresponds to evaluating the true array response only at the
carrier frequency $f_c$. For this approximation, we define

$$a_T'(\theta_T) = a_T(\theta_T, f_c)$$

$$= \left[ 1, e^{j2\pi c^{-1}f_c d \sin \theta_T}, \ldots, e^{j2\pi c^{-1}f_c d \sin (N_T-1)d \sin \theta_T} \right]^T, \quad (2.29)$$

$$a_R'(\theta_R) = a_R(\theta_R, f_c)$$

$$= \left[ 1, e^{j2\pi c^{-1}f_c d \sin \theta_R}, \ldots, e^{j2\pi c^{-1}f_c d \sin (N_R-1)d \sin \theta_R} \right]^T, \quad (2.30)$$

so that the approximate passband beamforming model is

$$H'_C(f) = \frac{1}{\sqrt{N_R N_T}} \sum_{l=1}^{N_l} g_l w_R^H a_R'(\theta_{R,l}) a_T'^H(\theta_{T,l}) w_T e^{-j2\pi \tau_l f}. \quad (2.31)$$

We stress that, as we will see, beam squint becomes apparent in the true model (2.23), whereas it is hidden in the approximate model (2.31).

2.3.4 Discrete-Time Equivalent Channel Model

To obtain a discrete-time equivalent channel model, we apply orthogonal frequency-division multiplexing (OFDM) [12] with bandwidth $B$ and $N_f$ subcarriers. Label all the subcarriers as $0, 1, \ldots, N_f - 1$ with increasing frequency. The signals are down converted to baseband such that subcarrier $N_f/2$ is the DC tone. Suppose the baseband transmitted signal vector is $s \in \mathbb{C}^{N_f \times 1}$, and the baseband received signal vector is $r \in \mathbb{C}^{N_f \times 1}$.

$$s = [s_0, s_1, \ldots, s_{N_f-1}]^T, \quad (2.32)$$

$$r = [r_0, r_1, \ldots, r_{N_f-1}]^T. \quad (2.33)$$
The discrete-time equivalent channel vector is denoted as
\[
\mathbf{H} = [\mathbf{H}(0), \mathbf{H}(1), \ldots, \mathbf{H}(N_f - 1)]^T,
\] (2.34)
where \( \mathbf{H} \in \mathbb{C}^{N_f \times 1} \), and \( \mathbf{H}(n_f) \) is the channel gain of subcarrier \( n_f \). The received noises for all subcarriers are independent and identically distributed (i.i.d) zero-mean Gaussian noise with variance \( \sigma^2 \), and the noise vector is denoted as \( \mathbf{z} \). We have
\[
\mathbf{r} = \mathbf{H} \odot \mathbf{s} + \mathbf{z},
\] (2.35)
where \( \odot \) stands for Hadamard product, i.e., element-wise product.

The beamforming vectors for subcarrier \( n_f \) are
\[
\mathbf{a}_T(\theta_T, n_f) = \left[ 1, e^{j2\pi c^{-1}f_{nf}d\sin\theta_T}, \ldots, e^{j2\pi c^{-1}f_{nf}nd\sin\theta_T}, \ldots, e^{j2\pi c^{-1}f_{nf}(N_T - 1)d\sin\theta_T} \right]^T,
\] (2.36)
\[
\mathbf{a}_R(\theta_R, n_f) = \left[ 1, e^{j2\pi c^{-1}f_{nf}d\sin\theta_R}, \ldots, e^{j2\pi c^{-1}f_{nf}nd\sin\theta_R}, \ldots, e^{j2\pi c^{-1}f_{nf}(N_R - 1)d\sin\theta_R} \right]^T,
\] (2.37)
where \( f_{nf} \) is the frequency of subcarrier \( n_f \) in the passband. Based on (2.28), the discrete-time equivalent channel model for subcarrier \( n_f \) can be expressed as
\[
\mathbf{H}(n_f) = \frac{1}{\sqrt{N_RN_T}} \sum_{l=1}^{N_i} g_l \mathbf{w}_R^H \mathbf{a}_R(\theta_{R,l}, n_f) \mathbf{a}_T^H(\theta_{T,l}, n_f) \mathbf{w}_T e^{-j2\pi \chi_l \frac{n_f + N_f/2}{N_f}}.
\] (2.38)
where \( N_f \) is size of the OFDM symbol, \( \chi_l \) is the delay of path \( l \) in the number of discrete-time samples. With the approximation of the array response vectors in (2.29) and (2.30), the commonly used approximate discrete-time equivalent channel model
\[ H'(n_f) = \frac{1}{\sqrt{N_R N_T}} \sum_{l=1}^{N_l} g_l w_R^H a'_R(\theta_{R,l}) a'_T H(\theta_{T,l}) w_T e^{-j2\pi \chi l \frac{n_f + N_f/2}{N_f}}. \] (2.39)

\[ H'(n_f) \] in (2.39) can be reorganized as

\[ H'(n_f) = \sum_{l=1}^{N_l} g_t \left( \frac{1}{\sqrt{N_R}} w_R^H a'_R(\theta_{R,l}) \right) \left( \frac{1}{\sqrt{N_T}} w_T^H a'_T(\theta_{T,l}) \right)^H e^{-j2\pi \chi l \frac{n_f + N_f/2}{N_f}} \]

\[ = \sum_{l=1}^{N_l} g_t g'_R(w_R, \theta_{R,l}) g'_T(w_T, \theta_{T,l}) e^{-j2\pi \chi l \frac{n_f + N_f/2}{N_f}}, \] (2.40)

where

\[ g'_T(w_T, \theta_{T,l}) = \frac{1}{\sqrt{N_T}} w_T^H a'_T(\theta_{T,l}), \] (2.41)

\[ g'_R(w_R, \theta_{R,l}) = \frac{1}{\sqrt{N_R}} w_R^H a'_R(\theta_{R,l}). \] (2.42)

\[ g'_T(w_T, \theta_{T,l}) \] and \[ g'_R(w_R, \theta_{R,l}) \] are the transmitter and receiver array gains, respectively. Again, beam squint is not considered in \[ g'_T(w_T, \theta_{T,l}) \] and \[ g'_R(w_R, \theta_{R,l}) \].

2.4 Related Topics

In this section, we review topics on initial access, channel estimation, carrier aggregation and beamforming codebook design, providing the background information for our discussion in Chapter 4 to Chapter 7.

2.4.1 Initial Access and Channel Estimation

When the mmWave system starts up, the transmitter and receiver do not know on which directions their beams should focus. A process called initial access is required for the transmitter and receiver arrays to find the appropriate beam focus angles.
Selecting transmitter and receiver beams is equivalent to selecting a signal path, and ideally one of the strongest signal paths, for transmission. Usually, path gain (2.42) is used as the performance metric, because beam squint is typically ignored.

There are four basic types of initial access schemes: exhaustive search, hierarchical search, compressive sensing and context information (CI)-based search [22]. In these initial access schemes, switched beamforming is assumed and therefore, a beamforming codebook is used. Continuous beamforming can be approximated with switched beamforming by increasing the number of beams in the codebook, i.e., codebook size. Details on the four initial access schemes are as follows.

- **Exhaustive Search** is a brute-force search method in which all combinations of the transmitter and receiver beams are tested [28]. Suppose there are $M_T$ beams in the transmitter codebook and $M_R$ beams in the receiver codebook. There will be $M_T \times M_R$ tests.

- **Hierarchical Search** consists of a multi-stage scan of the angular space [16]. Each stage has a corresponding codebook. From the first stage to the last stage, the beamwidth of the beams in the codebook decreases. One beam in the codebook of previous stage covers the angle range of several beams in the next stage. In one stage, the best transmitter and receiver beams are found through exhaustive search. In the next stage, we only search through the beams that are covered by the beam in the previous stage. Usually, fine beams are used in the last stage.

- **Compressive Sensing (CS)** techniques enable sparse channel recovery [24]. MmWave propagation tends to exhibit one line-of-sight (LOS) and a few non-LOS paths [43], and CS uses this sparse nature of the mmWave propagation to recover the channel with a small number of observations.

- **CI-Based Search** is a location-assisted method [10]. Transmitter and / or receiver location information, such as GPS coordinates, is used to reduce the scope of the beam search.

In exhaustive search and hierarchical search, there is a tradeoff between search time and average array gain of the codebook. The higher the average array gain for the codebook, the larger codebook size that is needed, and therefore, the longer the search time required.
A significant step in any of the initial access schemes is to estimate the channel to determine which beam is preferred, and for this, a channel estimation algorithm is required. Usually, wideband signal is used in mmWave beamforming, and therefore, the channel estimation algorithm must be compatible with the wideband signal.

In mmWave system, OFDM could be used to provide higher spectral efficiency [45]. There are several pilot-aided channel estimation algorithms for OFDM, such as maximum likelihood estimator (MLE), minimum mean squared error (MMSE) estimator, and interpolation of pilots [17, 37]. Among the three, MMSE estimator has least channel estimation error, and interpolation provides the highest channel estimation error. MMSE estimator requires channel statistics, and MLE can be implemented in practice without channel statistics. We note that that current initial access schemes assume approximated channel model as shown in (2.31).

2.4.2 Carrier Aggregation

In LTE, carrier aggregation is one of the technologies for increasing throughput by aggregating multiple component carriers or bands [45]. The bands could be contiguous or non-contiguous, providing flexibility in utilization of the spectrum [61, 54].

For mmWave bands, the bandwidth of each band could be much larger than the LTE band, such as 1 GHz to several GHz. Multiple non-contiguous bands could also be aggregated to further increase channel capacity [9]. To the best of our knowledge, current study of carrier aggregation in mmWave beamforming does not consider beam squint. However, beam squint causes array gain’s variation over frequency, making carrier aggregation in mmWave bands different from that at the microwave bands. Therefore, carrier aggregation in mmWave bands with beam squint should be explored to determine its application values.
2.4.3 Beamforming Codebook Design

Beamforming has space selectivity, that is, the array gain is high around the beam focus angle, but small in other angle range. In switched beamforming, a codebook consists of multiple beam, and each beam is a codeword. Again, there is a tradeoff between the average array gain and the time or system overhead to find the optimal beam within a codebook.

Most current beamforming codebook designs assume that the beams in the codebook have the same angle response for all frequencies of study [3, 11, 57], and therefore, beam squint is ignored. These codebook designs focus on two directions, minimizing the effects of side lobes for reduced interference and maximizing the beamwidth while maintaining reasonable array gain.

In practice, the array gains of different frequencies or subcarriers for a specific angle are not the same in the wideband system due to beam squint. Current study of the effect of beam squint on phased-array communication is merely on codebook design for linear arrays, and there are few publications discussing the effect of beam squint on codebook design [32, 55]. The authors in [55] mention that there is gain drop due to beam squint in their codebook design. The authors in [32] analyze the effect of beam squint on a codebook for IEEE 802.15.3c, and propose to reduce the effect of beam squint by applying extra phase shifts to different subbands using additional phase shifters and bandpass filters. Again, the added hardware increases the system complexity as well as cost and power consumption. It is therefore important to design beamforming codebooks considering beam squint to meet certain criteria without additional hardware.
2.5 Summary

In this chapter, we summarized the architectures for analog beamforming, hybrid beamforming and digital beamforming, and motivated our focus on analog mmWave beamforming throughout this dissertation. A system model for a ULA employing analog beamforming was discussed, and we explained how beam squint results from using phase shifters instead of true-time-delay devices. Given that hardware-based methods to eliminate or reduce beam squint have higher complexity and cost, we proposed developing system designs and algorithms to work around the effects of beam squint. For this development in the remainder of the thesis, we summarized passband and baseband channel models that exhibit beam squint along with their simplification to ignore beam squint as is common in the literature. Finally, we reviewed background information on initial access, channel estimation, carrier aggregation, and beamforming codebook design to set the stage for our discussion in Chapters 4 to 7.
CHAPTER 3
MODELING OF BEAM SQUINT

In this chapter, we model beam squint in both a uniform linear array (ULA) and a uniform planar array (UPA), and we derive corresponding discrete-time equivalent channel models.

3.1 ULA with Beam Squint

We consider a ULA with \( N \) identical and isotropic antennas as shown in Figure 2.4. AoA/AoD \( \theta \) is the angle of the signal propagation path relative to the broadside of the array, increasing counterclockwise. The analysis of the effect of beam squint is the same for a transmitter or a receiver array. Without loss of generality, we focus on the case of a receive array.

Traditionally, the beamforming vector \( \mathbf{w} \) is designed for the carrier frequency. To achieve the highest array gain for a beam focus angle \( \theta_F \), the phase shifts in the beamforming vector \( \mathbf{w} \) follow (2.10). The beam with phase shifts described in (2.10) is a fine beam, and we focus on the analysis of fine beams throughout the dissertation.

For a wideband system, a subcarrier’s absolute frequency \( f \) can be expressed as

\[
f = \xi f_c,
\]

where \( f_c \) is the carrier frequency, and \( \xi \) is the ratio of the subcarrier frequency to the carrier frequency. Note that in this dissertation, the concept of subcarrier is not limited to an OFDM system. Suppose the baseband bandwidth of the signals of
interest is \( B \). Then \( f \in \left[ f_c - \frac{B}{2}, f_c + \frac{B}{2} \right] \). Define the fractional bandwidth as

\[
b := \frac{B}{f_c},
\]

(3.2)

Then \( \xi \in \left[ 1 - \frac{b}{2}, 1 + \frac{b}{2} \right] \). As a specific example, \( \xi \) varies from 0.983 to 1.017 in a system with 2.5 GHz bandwidth at 73 GHz carrier frequency.

The array gain in (2.11) can be transformed into

\[
g(\theta_F, \theta, \xi) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{j[2\pi c^{-1} f_c (n-1) d(\xi \sin \theta - \sin \theta_F)]}
\]

\[
= \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{j[2\pi \lambda_c^{-1} (n-1) d(\xi \sin \theta - \sin \theta_F)]},
\]

(3.3)

where \( \lambda_c \) is wavelength of the carrier frequency.

Let \( \theta' \) be the equivalent AoA for the subcarrier with coefficient \( \xi \). Comparing (3.3) to (2.11), \( \theta' \) can be expressed as

\[
\theta' = \arcsin (\xi \sin \theta).
\]

(3.4)

The variation of beam patterns for different frequencies in the wideband system can therefore be interpreted as the variation of AoAs in the same angle response for different subcarriers. In other words, frequency variation and AoA variation due to beam squint are two sides of the same coin.

Typically, the beamforming array is designed such that the antenna spacing satisfies

\[
d = \frac{\lambda_c}{2}.
\]

(3.5)

Note that through this dissertation, \( d = \lambda_c/2 \) is the default value unless specifically stated.
From (3.3),

\[
g(\theta_F, \theta, \xi) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{j[\pi(n-1)(\xi \sin \theta - \sin \theta_F)]} = \frac{\sin \left(\frac{N\pi}{2} (\xi \sin \theta - \sin \theta_F)\right)}{\sqrt{N \sin \left(\frac{\pi}{2} (\xi \sin \theta - \sin \theta_F)\right)}} e^{j\frac{(N-1)\pi}{2} (\xi \sin \theta - \sin \theta_F)},
\]

so that \(g(\theta_F, \theta, \xi)\) has linear phase of \(\xi \sin \theta - \sin \theta_F\). If we define the single-argument function

\[
g(x) = \frac{\sin \left(\frac{N\pi x}{2}\right)}{\sqrt{N \sin \left(\frac{\pi x}{2}\right)}} e^{j\frac{(N-1)\pi x}{2}},
\]

the array gain for frequency with \(\xi\) in (3.6) simplifies to \(g(\xi \sin \theta - \sin \theta_F)\).

Figure 3.1 illustrates an example plot of \(|g(x)|\). As can be seen, the side lobes have much smaller gains, and our study focuses on the main lobe. The range of the main lobe in \(g(x)\) is \(x \in \left[-\frac{2}{N}, \frac{2}{N}\right]\). The support of the main lobe in \(x\) is therefore \(\frac{4}{N}\).

In this dissertation, we focus on the magnitude (squared) of the array gain. Define
\( g^{-1}(y), y \in \mathbb{R}^+ \) as the inverse function of \(|g(x)|\) for \( x \in [0, \frac{2}{N}] \), that is, given \( y \) as the magnitude of the array gain, \( g^{-1}(y) \) outputs the corresponding angle in the positive part of the main lobe. \( -g^{-1}(y) \) can be used for the angle in the negative part of the main lobe.

Now let

\[
\psi = \sin \theta
\]

(3.8)
to convert the angle \( \theta \) space to the \( \psi \) space. Correspondingly, the AoA becomes \( \psi = \sin \theta \), the beam focus angle becomes \( \psi_F = \sin \theta_F \), and \( \psi_F(w) = \sin \theta_F(w) \). The array gain for subcarrier \( \xi \), AoA \( \psi \), and beam focus angle \( \psi_F \) can then be expressed as

\[
g(\xi \psi - \psi_F) = \frac{\sin \left( \frac{N\pi}{2} (\xi \psi - \psi_F) \right)}{\sqrt{N} \sin \left( \frac{\pi}{2} (\xi \psi - \psi_F) \right)} e^{j \frac{(N-1)\pi}{2} (\xi \psi - \psi_F)},
\]

(3.9)
where \( \psi, \psi_F \in [-1,1] \). The gain for subcarrier \( \xi \) at AoA \( \psi \) is equivalent to the gain for the carrier frequency at AoA \( \psi' = \xi \psi \).

The maximum array gain is

\[
g_m = \max_{\psi \in [-1,1]} |g(\psi - \psi_F)| = \sqrt{N}
\]

(3.10)
achieved by \( \psi = \psi_F \).

For simplicity, we study the effect of beam squint in \( \psi \) space. From (3.9) and Figure 3.7, the effect of beam squint increases as the AoA diverges from the beam focus angle, and it also tends to increase as the beam focus angle increases.
Figure 3.2. Structure of a uniform planar array (UPA) with analog beamforming using phase shifters. Each dot indicates one antenna, and each antenna is connected to one phase shifter. There are \( N_a = N_h \times N_v \) antennas, where \( N_h \) is the number of antennas in the horizontal direction and \( N_v \) is the number of antennas in the vertical direction. The distance between adjacent antennas is \( d \). The AoA / AoD has two components: \( \theta \) denotes the azimuth angle defined between the \( y \) axis and the signal’s projection on the \( x-y \) plane, and \( \phi \) denotes the elevation angle defined between the signal and the \( x-y \) plane. Note that \( \theta, \phi \in [-\pi/2, \pi/2] \).
3.2 UPA with Beam Squint

We consider a UPA as shown in Figure 3.2. The array is assumed to have \( N_a = N_h \times N_v \) antennas, where \( N_h \) is the number of antennas in the horizontal direction and \( N_v \) is the number of antennas in the vertical direction. One antenna can be located with coordinate \((m, n)\), where \( m = 1, 2, \ldots, N_v \) and \( n = 1, 2, \ldots, N_h \). \( \theta \) and \( \phi \) are used for the azimuth angle and elevation angle, respectively. The AoA/AoD consists of two components and is denoted

\[
\Omega = [\theta, \phi]^T,
\] (3.11)

Correspondingly, the AoD and AoA can be denoted as \( \Omega_T = [\theta_T, \phi_T]^T \), and \( \Omega_R = [\theta_R, \phi_R]^T \), respectively. The beam focus angle of the array is \( \Omega_F = [\theta_F, \phi_F]^T \).

The array gain of a UPA for a signal with wavelength \( \lambda \), beam focus angle \( \Omega_F \) and AoA/AoD \( \Omega \) is

\[
G(\Omega_F, \Omega) = \frac{1}{\sqrt{N_vN_h}} \sum_{m=1}^{N_v} \sum_{n=1}^{N_h} e^{j[2\pi\lambda^{-1}((n-1)d\sin\theta\cos\phi+(m-1)d\sin\phi)-\beta_{m,n}(\Omega_F)]},
\] (3.12)

where \( \beta_{m,n}(\Omega_F) \) is the phase of the phase shifter connected to antenna \((m, n)\). Again, we consider fine beams designed for the carrier frequency with

\[
\beta_{m,n}(\Omega_F) = 2\pi\lambda^{-1}[(n-1)d\sin\theta_F\cos\phi_F + (m-1)d\sin\phi_F].
\] (3.13)

Similar to a ULA, the UPA array gain for subcarrier \( \xi \) can be expressed as

\[
G(\Omega_F, \Omega, \xi) = \frac{1}{\sqrt{N_vN_h}} \sum_{m=1}^{N_v} \sum_{n=1}^{N_h} e^{j[2\pi\lambda^{-1}((n-1)d\xi\sin\theta\cos\phi+(m-1)d\xi\sin\phi)-\beta_{m,n}(\Omega_F)]},
\] (3.14)
The UPA is designed with \( d = \lambda_c / 2 \). From (3.13) and (3.14),

\[
G(\Omega_F, \Omega, \xi) = \frac{1}{\sqrt{N_v N_h}} \sum_{m=1}^{N_v} \sum_{n=1}^{N_h} e^{j[\pi(n-1)(\xi \sin \theta \cos \phi - \sin \theta_F \cos \phi_F) + \pi(m-1)(\xi \sin \phi - \sin \phi_F)]}
\]

\[
= \frac{\sin \left( \frac{N_v \pi}{2} (\xi \sin \theta \cos \phi - \sin \theta_F) \right)}{\sqrt{N_h \sin \left( \frac{\pi}{2} (\xi \sin \theta \cos \phi - \sin \theta_F) \right)}} e^{j\left(\frac{(N_h-1)\pi}{2} (\xi \sin \theta \cos \phi - \sin \theta_F) \right)} \cdot \frac{\sin \left( \frac{N_h \pi}{2} (\xi \sin \phi - \sin \phi_F) \right)}{\sqrt{N_v \sin \left( \frac{\pi}{2} (\xi \sin \phi - \sin \phi_F) \right)}} e^{j\left(\frac{(N_v-1)\pi}{2} (\xi \sin \phi - \sin \phi_F) \right)}.
\]  

(3.15)

We denote virtual AoA/AoD as

\[
\Psi = [\psi, \varphi]^T,
\]  

(3.16)

where the virtual azimuth angle and the virtual elevation angle are

\[
\psi = \sin \theta \cos \phi,
\]  

(3.17)

\[
\varphi = \sin \phi,
\]  

(3.18)

respectively. Since \( \theta, \phi \in [-\pi/2, \pi/2] \),

\[
\psi^2 + \varphi^2 = \sin^2 \theta \cos^2 \phi + \sin^2 \phi \leq \cos^2 \phi + \sin^2 \phi \leq 1.
\]  

(3.19)

The feasible region of \((\psi, \varphi)\) in two-dimensional space is a circle with radius 1. The phase of the phase shifter \((m, n)\) for a fine beam with beam focus angle \(\Psi_F\) can then be expressed as

\[
\beta_{m,n}(\Psi_F) = \pi \left[ (n - 1) \psi_F + (m - 1) \varphi_F \right].
\]  

(3.20)
Furthermore, the array gain for beam focus angle $\Psi_F$, AoD/AoD $\Psi$ and $\xi$ is

$$G(\Psi_F, \Psi, \xi) = \frac{\sin \left( \frac{N_h \pi}{2} (\xi \psi - \psi_F) \right)}{\sqrt{N_h \sin \left( \frac{\pi}{2} (\xi \psi - \psi_F) \right)}} e^{j(\frac{Nh-1}{2})\pi (\xi \psi - \psi_F)} \cdot \frac{\sin \left( \frac{N_v \pi}{2} (\xi \varphi - \varphi_F) \right)}{\sqrt{N_v \sin \left( \frac{\pi}{2} (\xi \varphi - \varphi_F) \right)}} e^{j(\frac{Nv-1}{2})\pi (\xi \varphi - \varphi_F)} \cdot g^{(Nh)}(\xi \psi - \psi_F) g^{(Nv)}(\xi \varphi - \varphi_F), \quad (3.21)$$

where $g^{(N)}(x)$ is the array gain of a ULA with $N$ antennas in (3.7). Thus, the UPA’s array gain is the product of two ULA’s array gains with $N_h$ and $N_v$ antennas. The maximum array gain is therefore

$$G_m = \sqrt{N_h \times N_v}. \quad (3.22)$$

### 3.3 Discrete-Time Equivalent Channel Model with Beam Squint

The discrete-time equivalent channel model with general beamforming vector has been derived in Section 2.3.4. In this section, we illustrate the discrete-time equivalent channel model based on the beam squint model of fine beams in Section 3.1. Although we only consider a ULA here, similar conclusions can be made for a UPA.

As in Section 2.3.4, we consider an OFDM system with bandwidth $B$ and $N_f$ subcarriers. Index all the subcarriers as $0, 1, \ldots, N_f - 1$ with increasing frequency. Subcarrier $N_f/2$ is converted to DC tone. Suppose the discrete delay spread is $L$. We can assume that there are $L$ paths, and that some of them may have zero channel gain, and we label the paths as $l = 0, 1, \ldots, L - 1$.

Suppose $\xi_{n_f}$ is the ratio of subcarrier frequency to the carrier frequency for subcarrier $n_f$. Then

$$\xi_{n_f} = 1 + \frac{(2n - N_f + 1) b}{2N_f}, \quad n = 0, 1, \ldots, N_f - 1, \quad (3.23)$$
so that

\[
\frac{1}{N_f} \sum_{n=0}^{N_f} \xi_{n_f} = 1. 
\]

(3.24)

The transmitter and receiver array gains for subcarrier \( n_f \) of path \( l \) can be denoted as \( g_T (\xi_{n_f} \psi_{T,l} - \psi_{T,F}) \) and \( g_R (\xi_{n_f} \psi_{R,l} - \psi_{R,F}) \), respectively, where \( g(x) \) is defined in (3.9) and subscripts \((\cdot)_T\) and \((\cdot)_R\) denote transmitter and receiver, respectively. \( \psi_{T,F} \) and \( \psi_{R,F} \) are the beam focus angles for the transmitter and receiver, respectively. \( \psi_{T,l} \) is the AoD for path \( l \), and \( \psi_{R,l} \) is the AoA for path \( l \).

Let \( h_{l,n_f} \) be the channel gain for path \( l \) and subcarrier \( n_f \). Then

\[
h_{l,n_f} = g_{l,R} (\xi_{n_f} \psi_{R,l} - \psi_{R,F}) g_{T} (\xi_{n_f} \psi_{T,l} - \psi_{T,F}),
\]

(3.25)

where \( l = 0, 2, \ldots, L - 1 \), \( n_f = 0, 1, \ldots, N_f - 1 \), and \( g_l \) is the path gain between the transmitter and receiver arrays. Similar to (2.38), the channel gain between the transmitter and receiver RF chains for subcarrier \( n_f \) is

\[
\mathcal{H} (n_f) = \sum_{l=0}^{L-1} g_{l,R} (\xi_{n_f} \psi_{R,l} - \psi_{R,F}) g_{T} (\xi_{n_f} \psi_{T,l} - \psi_{T,F}) e^{-j2\pi l n_f + N_f/2} N_f
\]

(3.26)

We denote the channel impulse response matrix for signal between transmitter and receiver RF chains as

\[
\Theta = \begin{bmatrix}
        h_{0,0} & h_{0,1} & \cdots & h_{0,N_f-1} \\
        h_{1,0} & \cdots & \cdots & h_{1,N_f-1} \\
        \vdots & \cdots & \cdots & \vdots \\
        h_{L-1,0} & \cdots & \cdots & h_{L-1,N_f-1}
\end{bmatrix},
\]

(3.27)
where $\Theta$ is a $L \times N_f$ matrix. Then

$$\mathcal{H} = [\mathcal{H}(0), \mathcal{H}(1), \ldots, \mathcal{H}(N_f - 1)]^T = \text{Diag}(Q\Theta),$$  \hspace{1cm} (3.28)

where the operation $\text{Diag}(\cdot)$ generates a vector with the diagonal elements of a matrix, and $Q$ is an $N_f \times L$ matrix with entries

$$[Q]_{n,l} = e^{-j2\pi l \frac{n + N_f/2}{N_f}}, \hspace{0.5cm} 0 \leq n \leq N_f - 1, \hspace{0.5cm} 0 \leq l \leq L - 1.$$

$$\text{(3.29)}$$

Figure 3.3 compares two examples of channel magnitude responses, both with and without beam squint. In the examples, the magnitude responses with beam squint have larger variation than those without beam squint.

3.4 Summary

Array gain with beam squint was modeled for both a ULA and a UPA. The array gain variation over frequency for the same AoA/AoD is converted to the array gain variation over AoA/AoD for the carrier frequency. We also modeled the baseband-equivalent discrete-time channel with beam squint and demonstrated that its magnitude response varies more substantially over frequency with beam squint than without.
Figure 3.3. Examples of frequency-domain channel gains with and without beam squint for a ULA. There is no beam squint at the transmitter, and noise is not considered. The number of antennas $N = 16$, the beam focus angle $\psi_{R,F} = 0.9$, AoA $\psi_{R,I} = 0.94$, fractional bandwidth $b = 0.0342$ ($B = 2.5$ GHz, $f_c = 73$ GHz), and the number of subcarriers in an OFDM symbol $N_f = 2048$. (a) One path. (b) Two paths, path 0 and 15, with the same AoAs and with channel gain of path 15 being 10% that of path 0.
CHAPTER 4

EFFECTS OF BEAM SQUINT

Base on the beam squint model developed in Chapter 3, some effects of beam squint on wireless communication system are illustrated in this chapter, including effects on channel capacity, effects on path selection, and effects on channel estimation.

4.1 Effect on Channel Capacity

This section illustrates the effects of beam squint on channel capacity for a ULA, with the understanding that there are similar effects in the case of a UPA. Although in general the effect of beam squint on capacity depends upon the AoD, the transmitter beam focus angle, the AoA, and the receiver beam focus angle, we consider for simplicity the case of no beam squint at the transmitter and study the beam squint of a single array at the receiver. Therefore, only the AoA and the receiver beam focus angle are examined here.

Continuing from Section 3.3, we consider a complex channel and an OFDM system with bandwidth $B$ and $N_f$ subcarriers, labeled as $0, 1, \ldots, N_f - 1$. Equal power is allocated to all subcarriers at the transmitter, which is common if channel state information is not available at the transmitter. The total transmit power is denoted $P$, and the two-sided power spectral density of white Gaussian noise modeling receiver thermal noise is $\sigma^2/2$.

Due to poor scattering and significant attenuation of mmWave signals, the mmWave channel is sparse, that is, there is smaller number of significant signal paths compared
to that of microwave channels [2, 25, 43]. If fine beams are used at both the transmitter and receiver, most of the paths are filtered out by the two spatial filters. For simplicity here, we assume there is only one signal path, that is, \( L = 1 \) in the model of (3.26) with path gain \( g_1 \). The received power within bandwidth \( B \) at each receiver antenna before beamforming is denoted by

\[
P_R = P |g_T g_1|^2, \quad (4.1)
\]

where \( g_T \) is transmitter array gain.

Viewing the OFDM system as a set of parallel Gaussian noise channels [15], the channel capacity is the sum of the capacities of each subchannel having bandwidth \( B/N_f \). Incorporating beamforming with fine beams and accounting for beam squint in the receive array, the received power on subcarrier \( n \) with AoA \( \psi \), beam focus angle \( \psi_F \), and fractional bandwidth \( b \) is \( P_R|g(\xi_n \psi - \psi_F)|^2 \), where \( g(\cdot) \) is defined in (3.7). Similar to \( \xi_{n_f} \) in (3.23), define

\[
\xi_n = 1 + \frac{(2n - N_f + 1)b}{2N_f}, \quad n = 0, 1, \ldots, N_f - 1, \quad (4.2)
\]

so that

\[
\frac{1}{N_f} \sum_{n=0}^{N_f} \xi_n = 1. \quad (4.3)
\]

Following Section 5.2.1 in [52], the channel capacity with beam squint at AoA \( \psi \), beam focus angle \( \psi_F \), and fractional bandwidth \( b \) is

\[
C_{BS}(\psi_F, \psi, b) = \frac{B}{N_f} \sum_{n=0}^{N_f-1} \log \left( 1 + \frac{P_R|g(\xi_n \psi - \psi_F)|^2}{B\sigma^2} \right). \quad (4.4)
\]
For comparison, the channel capacity without beam squint at AoA $\psi$ and beam focus angle $\psi_F$ is

$$C_{NBS}(\psi_F, \psi) = B \log \left( 1 + \frac{P_R |g(\psi - \psi_F)|^2}{B\sigma^2} \right),$$  

(4.5)

and we note that $C_{NBS}(\psi_F, \psi)$ is maximized if $\psi = \psi_F$. The proof is straightforward. All subcarriers in the case without beam squint achieve their largest array gain when $\psi = \psi_F$.

For further discussion, we want to highlight some of the properties of the array gain function $g(x)$ for $x \in [-2/N, +2/N]$, where $g(\cdot)$ is defined in (3.7).

First, an important property to establish is the 3 dB point, i.e., the positive solution $x_{3\text{dB}}$ to the equation

$$|g(x)|^2 = \frac{|g_m|^2}{2}, \quad x \in (0, +2/N],$$

(4.6)

where $g_m$ is the maximum array gain of a ULA defined in (3.10). Numeric solution of (4.6) suggests that $x_{3\text{dB}} \approx 2.78/(N\pi)$ [40]. For fixed fractional bandwidth $b$ and beam focus angle $\psi_F$, we define the 3 dB beamwidth for all frequencies as the set of angles

$$\mathcal{R}_{3\text{dB}}(\psi_F, b) = \left\{ \psi : |g(\xi \psi_F - \psi)|^2 \geq \frac{|g_m|^2}{2}, \quad \forall \xi \in [1 - b/2, 1 + b/2] \right\},$$

(4.7)

and with some manipulations, we obtain

$$\mathcal{R}_{3\text{dB}}(\psi_F, b) = \left\{ \psi : \frac{\psi_F - x_{3\text{dB}}}{1 - b/2} \leq \psi \leq \frac{\psi_F + x_{3\text{dB}}}{1 + b/2} \right\},$$

(4.8)
and

\[ |\mathcal{R}_{3\text{dB}}(\psi_F, b)| = \frac{2x_{3\text{dB}} - b\psi_F}{(1 - b^2/4)}. \quad (4.9) \]

Note that we focus on the 3dB beamwidth in the main lobe due to its large array gain. As we would expect, for the narrowband approximation \( b \approx 0 \), \( |\mathcal{R}_{3\text{dB}}(\psi_F, 0)| = 2x_{3\text{dB}} \).

Second, we characterize the concavity of \(|g(x)|^2\). Figure 4.1 illustrates the squared magnitude \(|g(x)|^2\), its derivative, and its second-order derivative as functions of \( x \) in the main lobe for \( N = 16 \). For the red portion of the curve in Figure 4.1(c),

\[ \frac{d^2}{dx^2} |g(x)|^2 \leq 0, \]

illustrating that \(|g(x)|^2\) is concave in that interval. The interval of concavity \([-x_\cap, +x_\cap]\) can be obtained by solving for the positive solution \( x_\cap \) to the equation

\[ \frac{d^2}{dx^2} |g(x)|^2 = 0, \quad x \in (0, +2/N]. \quad (4.10) \]

Numeric solution of (4.10) suggests that \( x_\cap \approx 2.61/(N\pi) \). For fixed fractional bandwidth \( b \) and beam focus angle \( \psi_F \), we define the concavity beamwidth for all frequencies as the set of angles

\[ \mathcal{R}_\cap(\psi_F, b) = \left\{ \psi : \frac{d^2}{dx^2} |g(x)|^2 \bigg|_{x=(\xi_\psi-\psi_F/2)} \leq 0, \quad \forall \xi \in [1 - b/2, 1 + b/2] \right\}, \quad (4.11) \]

and with some manipulations, we obtain

\[ \mathcal{R}_\cap(\psi_F, b) = \left\{ \psi : \frac{\psi_F - x_\cap}{(1 - b/2)} \leq \psi \leq \frac{\psi_F + x_\cap}{(1 + b/2)} \right\}, \quad (4.12) \]
Figure 4.1. Characteristics of $|g(x)|^2$ from (3.7), its derivative, and its second-order derivative within the main lobe. The number of antennas in the ULA $N = 16$. 


\[ |R \cap (\psi F, b)| = \frac{2x_{\cap} - b\psi_F}{(1 - b^2/4)}. \quad (4.13) \]

To study the relationship between \( C_{BS}(\psi F, \psi, b) \) and \( C_{NBS}(\psi F, \psi) \), we focus on angles \( \psi \) in the concavity width \( R \cap (\psi F, b) \). Although this set of angles is smaller than the 3 dB beamwidth \( R_{3\text{dB}}(\psi F, b) \), which may seem more natural from a beam-forming design perspective, the two are comparable. For example, in the narrowband approximation \( b \approx 0 \),

\[ \frac{|R \cap (\psi F, 0)|}{|R_{3\text{dB}}(\psi F, 0)|} \approx 93.9\% \]

**Fact 1.** For fixed fractional bandwidth \( b \) and beam focus angle \( \psi_F \), if the AoA \( \psi \in R \cap (\psi F, b) \) then

\[ C_{BS}(\psi F, \psi, b) \leq C_{NBS}(\psi F, \psi), \quad (4.14) \]

with equality if and only if \( \psi = 0 \).

**Proof.** For fixed fractional bandwidth \( b \) and beam focus angle \( \psi_F \), if the AoA \( \psi \in R \cap (\psi F, b) \), we have

\[
C_{BS}(\psi F, \psi, b) = \frac{B}{N_f} \sum_{n=0}^{N_f-1} \log \left( 1 + \frac{P_R |g(\xi_n \psi - \psi_F)|^2}{B \sigma^2} \right) \]

\[
\leq B \log \left( 1 + \frac{P_R \sum_{n=0}^{N_f-1} |g(\xi_n \psi - \psi_F)|^2}{N_f B \sigma^2} \right) \]
\[
\begin{align*}
&\leq B \log \left( 1 + \frac{P_R \left| g \left( \frac{1}{N_f} \sum_{n=0}^{N_f-1} (\xi_n \psi - \psi_F) \right) \right|^2}{B \sigma^2} \right) \\
&= B \log \left( 1 + \frac{P_R |g(\psi - \psi_F)|^2}{B \sigma^2} \right) \\
&= C_{\text{NBS}}(\psi_F, \psi). \quad (4.15)
\end{align*}
\]

The inequalities (a) and (b) follow from Jensen’s Inequality [39], where (a) is due to the concavity of the log function, and (b) is due to the concavity of \(|g(x)|^2\) for \(x \in [-x_r, x_r]\). The equality holds if and only if \(\psi = 0\) for both (a) and (b), i.e., when there is no beam squint.

We note that, outside \(R_\cap(\psi_F, b)\), beam squint could actually increase the channel capacity. For example, if \(\psi = \pm 2/N\), \(C_{\text{NBS}}(\psi_F, \psi) = 0\) and \(C_{\text{BS}}(\psi_F, \psi, b) > 0\). For \(\psi \in R_{3ib}(\psi_F, b)\) but \(\psi \notin R_\cap(\psi_F, b)\), the relation between \(C_{\text{NBS}}(\psi_F, \psi)\) and \(C_{\text{BS}}(\psi_F, \psi, b)\) must be examined on a case by case basis due to the concavity of log function and convexity of \(|g(x)|^2\). We do not emphasize scenarios in which \(\psi \notin R_{3ib}(\psi_F, b)\) in this dissertation, because the array gain is low.

For a given receive power per transmit antenna \(P_R\), \(C_{\text{NBS}}(\psi_F, \psi)\) increases with bandwidth \(B\) [15]. As \(B \to \infty\), \(C_{\text{NBS}}(\psi_F, \psi) \to \frac{P_R |g(\psi - \psi_F)|^2}{\sigma^2} \log_2 e [15]\). However, this increasing trend does not apply to \(C_{\text{BS}}(\psi_F, \psi, b)\). As \(B\) increases, \(C_{\text{BS}}(\psi_F, \psi, b)\) has trends of first increasing and then decreasing. Figure 4.2 illustrates three examples. For small \(B\), the effect of beam squint is limited, and \(C_{\text{BS}}(\psi_F, \psi, b)\) increases due to the dominant effect of the growing bandwidth, similar to \(C_{\text{NBS}}(\psi_F, \psi)\). As \(B\) further increases, the effect of beam squint dominates, and a larger portion of the subcarriers have smaller array gains. Some subcarriers may even fall outside the main lobe. Therefore, \(C_{\text{BS}}(\psi_F, \psi, b)\) begins to decrease.

Fact 4 suggests that beam squint decreases channel capacity for most practical
Figure 4.2. Capacity with beam squint $C_{BS}(\psi_F, \psi, b)$ and without beam squint $C_{NBS}(\psi_F, \psi, b)$ as a function of bandwidth for $\psi = \psi_F = 0.9$ with fixed $P_R/\sigma^2 = 2 \times 10^9$ Hz, and the number of subcarriers in an OFDM symbol $N_f = 2048$.

scenarios. To further illustrate the degree to which capacity decreases due to beam squint, we define the relative capacity loss as

$$D(\psi_F, \psi, b) = \frac{C_{NBS}(\psi_F, \psi) - C_{BS}(\psi_F, \psi, b)}{C_{NBS}(\psi_F, \psi)}.$$  \hspace{1cm} (4.16)

To study the effect of fractional bandwidth on channel capacity, 3 dB beamwidth without beam squint for $\psi_F$ is defined as

$$\mathcal{R}_{3\text{db}}(\psi_F) = \mathcal{R}_{3\text{db}}(\psi_F, b)|_{b=0}.$$  \hspace{1cm} (4.17)

For fixed $\psi_F$ and $b$, we defined the maximum relative capacity loss across the 3 dB beamwidth $\mathcal{R}_{3\text{db}}(\psi_F)$ as

$$D_{\text{max}, \psi}(\psi_F, b) = \max_{\psi \in \mathcal{R}_{3\text{db}}(\psi_F)} D(\psi_F, \psi, b),$$ \hspace{1cm} (4.18)
Similarly, we define the minimum relative capacity loss across the 3 dB beamwidth $\mathcal{R}_{3\text{dB}} (\psi_F)$ as

$$D_{\min, \psi} (\psi_F, b) = \min_{\psi \in \mathcal{R}_{3\text{dB}} (\psi_F)} D (\psi_F, \psi, b) , \quad (4.19)$$

For fixed $b$, the maximum relative capacity loss for all $\psi \in \mathcal{R}_{3\text{dB}} (\psi_F)$ and $\psi_F \in [-1, 1]$ is defined as

$$D_{\max, \psi, \psi_F} (b) = \max_{\psi_F \in [-1,1]} D_{\max, \psi} (\psi_F, b) . \quad (4.20)$$

Figure 4.3 shows examples of the maximum and minimum relative capacity loss across $\psi \in \mathcal{R}_{3\text{dB}} (\psi_F)$. Both $D_{\max, \psi} (\psi_F, b)$ and $D_{\min, \psi} (\psi_F, b)$ increase as $N$ and $|\psi_F|$ increase. Figure 4.4 illustrates that $D_{\max, \psi, \psi_F} (b)$ increases as the fractional bandwith $b$ increases, i.e., as the bandwidth increases for fixed carrier frequency.
Figure 4.4. Maximum relative capacity loss among all fine beams, $D_{\max,\psi,\psi_F}(b)$, as a function of fractional bandwidth $b$. Here the number of subcarriers in an OFDM symbol $N_f = 2048$, and $\frac{P_R}{B\sigma^2} = 0$ dB.

In summary, the reduction in channel capacity increases with increases in any of the following:

(a) the number of antennas in the array, $N$;
(b) the fractional bandwidth, $b$;
(c) the magnitude of the beam focus angle, $|\psi_F|$.

4.2 Effect on Path Selection

For environments in which there are multiple transmission paths between the transmitter and receiver, we might expect only one path to pass through the spatially selective fine beams on both sides. For the selected path, the transmitter and receiver will focus on the AoD and AoA of that path, respectively. This section demonstrates the need to consider path and beam selection jointly, and using channel capacity as the metric, when beam squint is significant. We consider again a ULA, with the understanding that analogous conclusions apply to UPA.
For concreteness, suppose the environment exhibits $N_l$ transmission paths, labeled as $l = 1, 2, \ldots, N_l$. Path $l$ has AoD $\theta_{T,l}$, AoA $\theta_{R,l}$, and complex path gain $g_l$. For simplicity, Doppler shift $\nu_l$ is ignored.

Traditional path selection methods choose a path based on the largest end-to-end channel gain incorporating the path gain as well as beamforming as in Design Criterion 1 [11].

**Design Criterion 1 (Max Gain Path Selection).** The selected path

$$l^* = \arg \max_{l \in \{1, 2, \ldots, N_l\}} |g_l g_R (\psi_{R,l} - \psi_{R,F,l}) g_T (\psi_{T,l} - \psi_{T,F,l})|,$$  \hspace{1cm} (4.21)

where $\psi_{T,F,l}$ and $\psi_{R,F,l}$ are the transmitter and receiver beam focus angles designed for path $l$, respectively.

If the AoD and AoA are known precisely, we can set $\psi_{T,F,l} = \psi_{T,l}$ and $\psi_{R,F,l} = \psi_{R,l}$, so that the traditional path selection rule (4.21) becomes

$$l^* = \arg \max_{l \in \{1, 2, \ldots, N_l\}} |g_l|,$$ \hspace{1cm} (4.22)

corresponding to selecting the path with the largest gain.

Neither of the path selection rules in (4.21) and (4.22) consider beam squint, but we have shown in the previous section that beam squint reduces channel capacity. Taking into account beam squint at both the transmitter and receiver, we conjecture that the path with the highest gain may not lead to the highest channel capacity. To motivate the conjecture, Figure 4.5 illustrates an example scenario with two paths, path 1 and path 2. The AoD and AoA of path 1 are both 0, and path 1 is shorter than path 2 but passes through a wall. Suppose the path gain of path 2 is slighter larger than that of path 1, i.e., $|g_2| > |g_1|$, due to the attenuation from the wall in path 1. If path selection is based on path gain only, path 2 would be selected. However,
based upon our results, we expect there to be beam squint and a reduction of channel capacity $C_2$ on path 2, while there is no beam squint or reduction of channel capacity $C_1$ on path 1. We expect that there will be scenarios in which $C_2 < C_1$.

With beam squint considered, the path selection criterion is stated in Design Criterion 2:

**Design Criterion 2 (Max Capacity Path Selection).** *For a scenario with $N_l$ paths, consider beams focusing on path $l$, that is, $\psi_{T,F} = \psi_{T,l}$ and $\psi_{R,F} = \psi_{R,l}$, resulting in channel capacity $C_l$, $l = 1, 2, \ldots, N_l$. The selected path $l^*$ is*

$$l^* = \arg \max_{l \in \{1, 2, \ldots, N_l\}} C_l. \quad (4.23)$$

We note that water filling [15] should be used to calculate channel capacity $C_l$, because AoA and AoD are known.
Figure 4.6. Channel capacities from path 1 and path 2 as function of AoA and AoD. ULAs are used in both the transmitter and receiver. 
\[
\psi_{R,l} = \psi_{T,l}, \text{ and } |g_2| = 1.1 |g_1|.
\]
The number of transmitter antennas \(N_T = 64\), and the number of receiver antennas \(N_R = 16\), the number of subcarriers in an OFDM symbol \(N_f = 2048\), fractional bandwidth \(b = 0.0714\) (\(B = 2\) GHz and \(f_c = 28\) GHz, or \(B = 5.2\) GHz and \(f_c = 73\) GHz), and \(\frac{P|g_1|^2}{B\sigma^2} = 0\) dB.

Figure 4.6 continues exploring the example in Figure 4.5. It shows an example of the variations of \(C_1\) and \(C_2\) as AoA \(\psi_{R,l}\) and AoD \(\psi_{T,l}\) increase with \(\psi_{R,l} = \psi_{T,l}\). The channel capacity is labeled on the left. \(C_2\) decreases significantly as AoA and AoD increase. If AoA and AoD are close to the broadside, i.e., \(\psi_T = \psi_R = 0\), it is better to use path 2 for higher channel capacity due to \(g_2 > g_1\), which arrives at the same conclusion as using channel gain as the metric. However, as AoA and AoD increase, the channel capacity of path 2, \(C_2\), decreases. To a certain value of AoA and AoD, it is better to use path 1 even if \(|g_2| = 1.1 |g_1|\) remains. \(C_2/C_1\) is also calculated with label on the right of the plot. \(C_2\) decreases significantly as AoA and AoD increase, further demonstrating the significance of using channel capacity for path selection.
4.3 Effect on Channel Estimation

In the previous sections, channel gain is directly used for channel capacity calculation. However, in a practical system, channel gain is usually an unknown parameter. Often the receiver and sometimes the transmitter operate their modulation and coding scheme as well as beam patterns based upon estimates of the channel obtained, for example, from pilot or reference signals.

Among the pilot-aided channel estimation algorithms for OFDM, MMSE estimator and MLE are widely used [17, 47]. In this dissertation, slow fading is assumed so that $h$ is modeled as a fixed vector during the period of channel estimation and related transmissions. We use MLE to study the effect of beam squint on channel estimation. We will show beam squint significantly increases channel estimation errors of MLE, even without any noise.

Again, we consider a ULA and an OFDM system with bandwidth $B$ and $N_f$ subcarriers, labeled $0, 1, \ldots, N_f - 1$, such that subcarrier $N_f/2$ is the DC tone.

We expect the channel capacity would decrease as channel estimation error increases. Usually, the performance metric for channel estimation error is mean-square error (MSE) of the estimated channel between the Tx and Rx RF chains [37],

$$\Gamma_0 = \mathbb{E} \left[ \frac{1}{N_f} \sum_{n=0}^{N_f-1} \left| \hat{\mathcal{H}}(n) - \mathcal{H}(n) \right|^2 \right], \quad (4.24)$$

where $\mathcal{H}(n)$ is the true channel, $\hat{\mathcal{H}}(n)$ is the estimated channel gain of subcarrier $n$, and the expectation is over joint distribution. However, $\Gamma_0$ varies with $\mathcal{H}(n)$. Therefore, the comparison with $\Gamma_0$ is not fair for channel estimation errors of different array gains. To make fair comparison among different array sizes and AoA, we use
MSE of path gain between transmitter and receiver arrays, i.e,

\[
\Gamma (\psi_F, \psi, b) = \mathbb{E} \left[ \sum_{l=0}^{N_f-1} |\hat{g}_l (\psi_F, \psi, b) - g_l|^2 \right],
\]

(4.25)

where \( g_l \) is the path gain between arrays, \( \hat{g}_l \) is the estimated path gain between arrays.

### 4.3.1 Channel Estimation with MLE

In OFDM systems, pilot or reference signals are usually used to assist in channel estimation. Suppose we have \( N_p \) pilots uniformly located in the bandwidth \( B \). Label the pilots as \( 0, 1, \ldots, N_p - 1 \), and let \( i_p(n) \in \{0, 1, \ldots, N_f - 1\}, n = 0, 1, \ldots, N_p - 1 \), denote the index of pilot \( n \) in the OFDM symbol. Denote the transmitted known pilot vector as \( s_p \), the received pilot vector as \( r_p \), and frequency-domain channel vector between the transmitter and receiver RF chains for pilots as \( \mathcal{H}_p \) with beam squint embedded. \( \mathcal{H}_p \in \mathbb{C}^{N_p \times 1} \) and

\[
\mathcal{H}_p = [\mathcal{H}(i_p(0)) \ \mathcal{H}(i_p(1)) \ \ldots \ \mathcal{H}(i_p(N_p-1))]^T,
\]

(4.26)

where \( i_p(n) \) is the index of pilot \( n \) in the OFDM symbol. We also have

\[
s_p = [s_{i_p(0)} \ s_{i_p(1)} \ \ldots \ s_{i_p(N_p-1)}]^T,
\]

(4.27)

\[
r_p = [r_{i_p(0)} \ r_{i_p(1)} \ \ldots \ r_{i_p(N_p-1)}]^T.
\]

(4.28)

The received pilot vector \( r_p \) can be expressed as

\[
r_p = \mathcal{H}_p \odot s_p + z_p,
\]

(4.29)

where \( \odot \) denotes Hadamard or element-wise product, and \( z_p \) is independent and identically distributed Gaussian noise vector for the pilots.
From [37], the MLE of the channel impulse response in model (4.29) is

$$\hat{h}_{\text{MLE}} = (Q_p^H Q_p)^{-1} Q_p^H (s_p^* \odot r_p), \quad (4.30)$$

where \((\cdot)^H\) denotes conjugate transpose, \((\cdot)^*\) denotes complex conjugate operation, and \(Q_p\) is an \(N_p \times L\) matrix with entries

$$[Q_p]_{n,l} = e^{-j2\pi \frac{n l + N_f / 2}{N_f}}, \quad 0 \leq n \leq N_p - 1, \quad 0 \leq l \leq L - 1, \quad (4.31)$$

where \(L\) is discrete delay spread. The corresponding MLE of \(H\) is

$$\hat{H}_{\text{MLE}} = A \hat{h}_{\text{MLE}} = A (Q_p^H Q_p)^{-1} Q_p^H (s_p^* \odot r_p), \quad (4.32)$$

where \(A\) is a transform matrix with entries

$$[A]_{n,l} = e^{-j2\pi \frac{n l + N_f / 2}{N_f}}, \quad 0 \leq n \leq N_f - 1, \quad 0 \leq l \leq L - 1. \quad (4.33)$$

The MLE requires the invertibility of \(Q_p^H Q_p\), and the condition is met if and only if \(Q_p\) is full rank and \(N_p > L\).

In summary, the inputs of the MLE are \(s_p, r_p\), and the indexes of the pilots. The output is the estimated channel vector \(\hat{H}_{\text{MLE}}\). Figure 4.7 illustrates an example of MLE channel estimation. There is significant difference between the estimated channel and the actual channel, even without considering noise.

4.3.2 Channel Estimation of ULA

In this section, we further show more numerical examples of the effect of beam squint on channel estimation when beam squint is not considered in MLE. A flat channel is considered. For simplicity, we assume there is no beam squint at the
Figure 4.7. Example of MLE channel estimation with beam squint in the receiver. There is only one path and there is no beam squint in the transmitter. There is no additive noise. The number of antennas $N = 16$, beam focus angle $\psi_F = 0.9$, AoA $\psi = 0.94$, bandwidth $B = 2.5$ GHz, carrier frequency $f_c = 73$ GHz, fractional bandwidth $b = 0.0342$, discrete delay spread $L = 128$, the number of subcarriers in an OFDM symbol $N_f = 2048$, and the number of pilots $N_p = 256$.

Transmitter and the power in the transmitter is equally assigned to each subcarrier. Therefore, the signal arrived at the receiver array is flat without beams squint while the receiver array has beam squint.

From (4.25),

$$
\Gamma (\psi_F, \psi, b) = \mathbb{E} \left[ |\hat{g}_0 (\psi_F, \psi, b) - g_0|^2 + \sum_{l=1}^{N_f-1} |\hat{g}_l (\psi_F, \psi, b)|^2 \right]
$$

$$
= \mathbb{E} \left[ \frac{1}{N_f} \sum_{n=0}^{N_f-1} \left| \frac{\hat{H} (n)}{g (\xi_n \psi - \psi_F)} - \frac{H (n)}{g (\xi_n \psi - \psi_F)} \right|^2 \right], \quad (4.34)
$$

where $g_l$ is the path gain between arrays, $\hat{g}_l$ is the estimated path gain between arrays,
\( \xi_n \) is defined in [4.2], \( \hat{H}(n) \) is estimated frequency-domain channel gain, and equality in (4.34) can be derived from Parseval’s theorem. The MSE is a function of \( \psi_F, \psi \) and \( b \).

For fixed beam focus angle \( \psi_F \), the angle set within 3 dB beamwidth is \( \mathcal{R}_{3\text{dB}}(\psi_F) \) as in (4.17). Figure 4.8 illustrates three examples of \( \Gamma(\psi_F, \psi, b) \) as a function of \( \psi \) for fixed beam focus angles \( \psi_F \). There is no beam squint if \( \psi = 0 \). Within the 3 dB beamwidth, the MSEs on two sides tend to be higher than in the center. If \( \psi_F \neq 0 \), the smallest MSE in a beam is not located at \( \psi_F \), but at a \( \psi \) with \( |\psi| < |\psi_F| \).

For certain \( \psi_F \) and \( b \), the maximum MSE across \( \psi \in \mathcal{R}_{3\text{dB}} \) is defined as

\[
\Gamma_{\text{max},\psi}(\psi_F, b) = \max_{\psi \in \mathcal{R}_{3\text{dB}}(\psi_F)} \Gamma(\psi_F, \psi, b),
\]

and the minimum MSE across \( \psi \in \mathcal{R}_{3\text{dB}} \) is defined as

\[
\Gamma_{\text{min},\psi}(\psi_F, b) = \min_{\psi \in \mathcal{R}_{3\text{dB}}(\psi_F)} \Gamma(\psi_F, \psi, b).
\]

Numerical examples of \( \Gamma_{\text{max},\psi}(\psi_F, b) \) and \( \Gamma_{\text{min},\psi}(\psi_F, b) \) are shown in Figure 4.9 and 4.10, respectively. \( \Gamma_{\text{max},\psi}(\psi_F, b) \) and \( \Gamma_{\text{min},\psi}(\psi_F, b) \) increase with \( \psi_F \) and \( N \). If \( \psi_F > 0.9 \), the curve for \( N = 64 \) in Figure 4.9 is not smooth because the equivalent AoAs of some subcarriers are out of the main lobe.

For certain \( b \), the maximum MSE across \( \psi \in \mathcal{R}_{3\text{dB}}(\psi_F) \) and \( \psi_F \in [-1, 1] \) is denoted as

\[
\Gamma_{\text{max},\psi,\psi_F}(b) = \max_{\psi_F \in [-1, 1]} \Gamma_{\text{max},\psi}(\psi_F, b).
\]

Figure 4.11 shows \( \Gamma_{\text{max},\psi,\psi_F}(b) \) as a function of \( b \). Within the main lobe, \( \Gamma_{\text{max},\psi,\psi_F}(b) \) increases as \( b \) increases, that is, \( \Gamma_{\text{max},\psi,\psi_F}(b) \) increases with the bandwidth for fixed carrier frequency. When \( b > 0.9 \), the curve for \( N = 64 \) is not smooth because the
Figure 4.8. MSE $\Gamma (\psi_F, \psi, b)$ in a fine beam as a function of AoA $\psi$ for fixed beam focus angle $\psi_F$. The dotted lines indicate $\Gamma_{\text{max,}\psi_F} (\psi_F, b)$ and $\Gamma_{\text{min,}\psi_F} (\psi_F, b)$, respectively. The number of antennas $N = 16$, fractional bandwidth $b=0.0342$ ($B = 2.5$ GHz, $f_c = 73$ GHz), discrete delay spread $L = 128$, the number of subcarriers in an OFDM symbol $N_f = 2048$, and the number of pilots $N_p = 256$. (a) Beam focus angle $\psi_F = 0$. (b) Beam focus angle $\psi_F = 0.9$. (c) Beam focus angle $\psi_F = -0.9$. 

58
Figure 4.9. The maximum MSE across 3 dB beamwidth in a fine beam, $\Gamma_{\text{max,}\psi}(\psi_F, b)$, as a function of beam focus angle $\psi_F$. Fractional bandwidth $b=0.0342$ ($B = 2.5$ GHz, $f_c = 73$ GHz), discrete delay spread $L = 128$, the number of subcarriers in an OFDM symbol $N_f = 2048$, and the number of pilots $N_p = 256$.

equivalent AoAs of some subcarriers are out of the main lobe.

To summarize these empirical results, the channel estimation error using MLE increases with increases in

(a) the number of antennas in the array, $N$;

(b) the fractional bandwidth, $b$;

(c) the magnitude of beam focus angle, $|\psi_F|$.

4.4 Summary

We illustrated that beam squint reduces channel capacity, and the reduction of channel capacity increases with the number of antennas in the array, fractional bandwidth, and the magnitude of beam focus angle. In system with beam squint, channel capacity is also shown to be a better path selection criterion than channel gain. With
Figure 4.10. The minimum MSE across 3 dB beamwidth in a fine beam, $\Gamma_{\min,\psi}(\psi_F, b)$, as a function of beam focus angle $\psi_F$. Fractional bandwidth $b=0.0342$ ($B = 2.5$ GHz, $f_c = 73$ GHz), discrete delay spread $L = 128$, the number of subcarriers in an OFDM symbol $N_f = 2048$, and the number of pilots $N_p = 256$.

beam squint, the channel estimation error using MLE increases with the number of antennas in the array, fractional bandwidth, and the magnitude of beam focus angle.
Figure 4.11. The maximum MSE across 3 dB beamwidth and beam focus angle range $[-1, 1]$, $\Gamma_{\max,\psi,\psi_F}(b)$, as a function of fractional bandwidth $b$. Discrete delay spread $L = 128$, the number of subcarriers in an OFDM symbol $N_f = 2048$, and the number of pilots $N_p = 256$. 
CHAPTER 5
CARRIER AGGREGATION WITH BEAM SQUINT

To the best of our knowledge, most existing works on beamforming only study a single band. In practice, the transmitter and receiver may access multiple non-contiguous bands to increase the throughput, i.e., carrier aggregation. Even if the bandwidth of each individual band is small, the aggregated channels from widely separated carriers can also introduce significant beam squint.

In this chapter, we study carrier aggregation with beam squint for a ULA transmitter and a ULA receiver. Note that we consider only one RF chain for multiple bands considering the size and cost constraints of the user equipment (UE).

Two beam squint scenarios are considered, carrier aggregation with one-sided beam squint and carrier aggregation with two-sided beam squint. Carrier aggregation with one-sided beam squint only considers beam squint in the receiver; carrier aggregation with two-sided beam squint assumes beam squint in both transmitter and receiver. We first study carrier aggregation with one-sided beam squint. Then the conclusion is applied to the study of carrier aggregation with two-sided beam squint.

5.1 System Setup

The system setup in this section can be used for carrier aggregation with both one-sided and two-sided beam squint.

We consider an OFDM system as in Chapter 3.3 to aggregate multiple non-contiguous bands. Suppose there are $N_B$ bands, labeled $1, 2, \ldots, N_B$ from the lowest
frequency to the highest frequency. The transmitter uses the $N_B$ bands to transmit data to the receiver. The bandwidths of the $N_B$ bands are $B_1, B_2, \ldots, B_{N_B}$, respectively. The center frequencies of the $N_B$ bands are $f_1, f_2, \ldots, f_{N_B}$, respectively, and

$$f_1 < f_2 < \cdots < f_{N_B}. \quad (5.1)$$

To avoid overlap of any two bands, we must have

$$f_i + \frac{B_i}{2} \leq f_{i+1} - \frac{B_{i+1}}{2}, \quad i = 1, 2, \ldots, N_B - 1. \quad (5.2)$$

We define the aggregate center frequency $f_a$ to be the center of all the bands, that is,

$$f_a = \frac{(f_1 - \frac{B_1}{2}) + (f_{N_B} + \frac{B_{N_B}}{2})}{2}. \quad (5.3)$$

Assume the ULA is designed based on the aggregate center frequency, that is, the distance between two adjacent antennas $d = \lambda_a/2$, where

$$\lambda_a = \frac{c}{f_a}. \quad (5.4)$$

c is speed of light.

Suppose there are $N_{f,i}$, $i = 1, 2, \ldots, N_B$ subcarriers in band $i$. The subcarriers of band $i$ are indexed by $n \in 0, 1, \ldots, N_{f,i} - 1$ with frequency increasing. The ratios of the subcarrier frequencies to the aggregate center frequency are defined as

$$\xi_{i,n} = \frac{f_i}{f_a} + \frac{(2n - N_{f,i} + 1) B_i}{2N_{f,i} f_a}, \quad (5.5)$$

where $i = 1, 2, \ldots, N_B$, and $n = 0, 1, \ldots, N_{f,i} - 1$.  

63
Assume the maximum transmission power across all bands is $P$. For simplicity, we assume that there is one dominant signal path with the same channel gain for all bands and that the remaining paths are filtered out by the transmitter and receiver array. Suppose the channel gain between the transmitter and receiver arrays is $g_p$. The two-sided power spectral density of the stationary white Gaussian noise modeling thermal noise is $\sigma^2/2$.

5.2 Carrier Aggregation with One-Sided Beam Squint

For one-sided beam squint, beam squint can be on either the transmitter or the receiver side. In this dissertation, we consider carrier aggregation with beam squint on the receiver side, with the understanding that similar conclusions can be drawn with beam squint only at the transmitter. This scenario of no beam squint at the transmitter can be achieved if the AoD is $0^\circ$ or true-time-delay devices are used at the transmitter.

Suppose the transmitter array gain is $g_T$. To maximize the channel capacity, water filling power allocation is used across bands. Based upon [52], the total channel capacity of all bands for the complex channel with beam squint at AoA $\psi$ and beam focus angle $\psi_F$ is

$$C_{WF}^{(NB)}(\psi_F, \psi) = \sum_{i=1}^{NB} \frac{B_i}{N_{f,i}} \sum_{n=0}^{-1} \log \left( 1 + \frac{P_{i,n}^* |g_T g_p g(\xi_{i,n} \psi - \psi_F)|^2}{B_i \sigma^2} \right), \quad (5.6)$$

where subscript WF denotes water filling, and $P_{i,n}^*$ is the power allocated to subcarrier $n$ of band $i$ such that

$$P_{i,n}^* = \left( \nu - \frac{B_i \xi_{i,n}^2}{N_{f,i}} \right)^+, \quad n = 0, 1, \ldots, N_{f,i} - 1. \quad (5.7)$$
Here for $i = 1, 2, \ldots, N_B$ and $n = 0, 1, \ldots, N_{f,i} - 1$,

$$\zeta_{i,n}^2 = \frac{\sigma^2}{|g_i g_p g_\psi (\zeta_{i,n} \psi - \psi_F)|^2}; \quad (5.8)$$

and $\nu$ is chosen such that

$$\sum_{i=1}^{N_B} \sum_{n=0}^{N_{f,i}-1} \left( \nu - \frac{B_i \varsigma_{i,n}^2}{N_{f,i}} \right)^+ = P, \quad (5.9)$$

where $(x)^+ := \max(x, 0)$. For a given AoA $\psi$, we would like to find the optimal beam focus angle $\psi^*_F$ to maximize the channel capacity in (5.6), that is,

$$\psi^*_F (\psi) = \arg \max_{\psi_F} C_{WF}^{(N_B)} (\psi_F, \psi). \quad (5.10)$$

Note that the optimal beam focus angle could be beyond the (virtual) angle range of $[-1, 1]$.

**5.2.1 Numerical Optimization**

It turns out that the transmission power allocations vary with the beam focus angle $\psi_F$, increasing the difficulty to find the optimal beam focus angle. The optimal beam focus angle falls within the interval

$$\mathcal{R}_{BF} = \left[ \frac{f_1}{f_a} \psi, \frac{f_N}{f_a} \psi \right], \quad (5.11)$$

where subscript BF denotes beam focus angle. If $\psi_F < \frac{f_1}{f_a} \psi$, the channel capacity is not maximized even with band 1 alone, and the subcarriers’ array gains of all other bands are smaller than those of band 1; therefore, the channel capacity cannot be maximized. Similarly, if $\psi_F > \frac{f_N}{f_a} \psi$, the channel capacity cannot be maximized, either.
Numerical optimization can be used to obtain an effective, though sub-optimal, beam focus angle. In the numerical optimization, the interval (5.11) is sampled into a finite set $\mathcal{R}_{BF}'$, defined as

$$
\mathcal{R}_{BF}' = \left\{ \psi_{F,i} : \psi_{F,i} = \frac{(i - 1)(f_{NB} - f_1)}{(N_{BF} - 1)f_a} \psi + \frac{f_1}{f_a} \psi, \; i = 1, 2, \ldots, N_{BF} \right\}, \quad (5.12)
$$

where there are $N_{BF}$ samples in $\mathcal{R}_{BF}'$. $N_{BF}$ is chosen to be an odd number such that the AoA $\psi$ will be included in $\mathcal{R}_{BF}'$.

The beam focus angle among $\mathcal{R}_{BF}'$ that has the largest channel capacity is defined as

$$
\psi_{F,NO}^* = \arg \max_{\psi_F \in \mathcal{R}_{BF}'} C_{WF}^{(N_B)}(\psi_F, \psi). \quad (5.13)
$$

The subscript NO denotes numerical optimization.

If there is more information on the band allocation, the information can be used to reduce the complexity of the numerical optimization. For example, if all the $N_B$ bands are symmetric with respect to the aggregate center frequency $f_a$, i.e. $B_i = B_{NB-i}$ and $f_i + f_{NB-i} = 2f_a$ for $i = 1, 2, \ldots, N_B$, there should be two $\psi_F^*$ that are symmetric to $\psi$ if $\psi_F^* \neq \psi$, and only one $\psi_F^*$ if $\psi_F^* = \psi$. Therefore, the range of the numerical optimization can be reduced to half the case without symmetry, and the number of samples required for the same resolution is reduced to $\frac{N_{BF}+1}{2}$, where $N_{BF}$ is odd.

5.2.2 Carrier Aggregation of Two Symmetric Bands

In this section, we study carrier aggregation for the special case of two non-contiguous bands, band 1 and band 2. Both band 1 and band 2 have bandwidth $B$ and their center frequencies are $f_1$ and $f_2$, respectively. The system setup is similar to that in Section 5.1 $f_1 < f_2$, and the two bands are located symmetrically with
respect to the aggregate center frequency \( f_a \), i.e.,
\[
f_a = \frac{f_1 + f_2}{2}.
\] (5.14)

The band separation or carrier separation, \( B_s \), is the difference between \( f_2 \) and \( f_1 \), i.e., \( B_s = f_2 - f_1 \). To ensure there is no overlap between the two bands, we must have \( B_s \geq B \). Note that if \( B_s = B \), the two bands can be treated as a single band.

The ratio of subcarrier frequencies to aggregate center frequency for subcarrier \( n \) in band \( i \), \( i = 1, 2 \) is
\[
\xi_{i,n} = 1 + (-1)^i \frac{B_s}{2f_a} + \frac{(2n - N_f + 1)B}{2N_ff_a}.
\] (5.15)

Water filling power allocation is also employed to maximize the channel capacity. The total channel capacity of band 1 and band 2 with beam squint at AoA \( \psi \) and beam focus angle \( \psi_F \) is
\[
C_{\mathrm{WF}}^{(2)}(\psi_F, \psi, B_s) = \frac{B}{N_f} \sum_{i=1}^{2} \sum_{n=0}^{N_f-1} \log \left( 1 + \frac{P_{n,i} |g_T g_p g (\xi_{i,n} \psi - \psi_F)|^2}{B/N_f \sigma^2} \right),
\] (5.16)

where \( i \) is the band index. The water filling power allocation is the same as (5.7).

\( C_{\mathrm{WF}}^{(2)}(\psi_F, \psi, B_s) \) is a function of the band separation. The optimal beam focus angle becomes
\[
\psi_F^*(\psi, B_s) = \arg \max_{\psi_F \in R_{\mathrm{BF}}} C_{\mathrm{WF}}^{(2)}(\psi_F, \psi, B_s).
\] (5.17)

Again, the optimal beam focus angle could be beyond the (virtual) angle range of \([-1, 1]\).
5.2.3 To Aggregate or Not to Aggregate

Beyond certain values of the system parameters, such as band separation, AoA, and SNR, it is preferable not to aggregate. We develop these results in this section.

5.2.3.1 Increasing Band Separation

Figure 5.1 shows examples of $C_{\text{WF}}^{(2)}(\psi, \psi, B_s)$ and $C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s)$ as functions of the band separation $B_s$. For $C_{\text{WF}}^{(2)}(\psi, \psi, B_s)$, the beam focuses on the center of the two bands; for $C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s)$ the beam focuses on the center of band 1. The horizontal dotted line is $C_{\text{WF}}^{(1)}(\psi, \psi)$, in which only one channel with bandwidth $B$ is used and the beam focuses on the center of that band. In general, $C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s) \geq C_{\text{WF}}^{(1)}(\psi, \psi)$. Note that the fluctuation of channel capacity in Figure 5.1 occurs because one or two bands fall in a side lobe of the array response, which is beyond the scope of our study. As SNR decreases, the gap between $C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s)$ and $C_{\text{WF}}^{(1)}(\psi, \psi)$ also decreases.

If the band separation $B_s$ is small, the two bands perform similarly to a single band. The maximum channel capacity is achieved if the beam focuses on the center of the two bands, that is, $\psi_F^* = \psi$, because of symmetry and the concavity of $|g(x)|^2$ for $x \in [-x_\gamma, x_\gamma]$. This system is said to be in the Aggregation Regime. The corresponding channel capacity is $C_{\text{WF}}^{(2)}(\psi, \psi, B_s)$.

Fix $\psi_F = \psi$. As $B_s$ increases, the two bands move away from the center of the main lobe of the array, the array gain decreases for both bands, and the channel capacity decreases. Eventually, as $B_s$ continues to increase, the beam can be focused on one of the bands, i.e, $\psi_F = \frac{f_1}{f_a} \psi$ or $\psi_F = \frac{f_2}{f_a} \psi$, resulting in higher channel capacity as illustrated in Figure 5.1. The system is then said to be in the Disaggregation Regime. Although there could still be power allocated to the other band having low array gain, the marginal increase in channel capacity obtained from that band is small. The corresponding channel capacities for the two cases are $C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s) =$
Figure 5.1. Example of aggregated channel capacities $C_{WF}^{(2)}(\psi, \psi, B_s)$ and $C_{WF}^{(2)}(\psi f_1/f_a, \psi, B_s)$ as functions of the band separation $B_s$. Critical band separation $B_{s,c} = 7.3$ GHz. The horizontal dotted line is $C_{WF}^{(1)}(\psi, \psi)$, in which only one channel with bandwidth $B$ is used and the beam focuses on the center of the band. If $B_s < B_{s,c}$, $C_{WF}^{(2)}(\psi, \psi, B_s) > C_{WF}^{(2)}(\psi f_1/f_a, \psi, B_s)$; if $B_s > B_{s,c}$, $C_{WF}^{(2)}(\psi, \psi, B_s) < C_{WF}^{(2)}(\psi f_1/f_a, \psi, B_s)$. The number of antennas in the array $N = 32$, AoA $\psi = 0.8$, channel bandwidth $B = 2.5$ GHz, and aggregate center frequency $f_a = 73$ GHz. $P_0g_Tg_R/\sigma_B^2 = 5$ dB, which is defined as $\text{SNR}_r$ in (5.24).
\[ C_{\text{WF}}^{(2)}(\psi f_2/f_a, \psi, B_s), \] where the equality results from the assumed symmetry about the aggregate center frequency \( f_a \).

Without loss of generality, we focus on \( C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s) \). Note that, in practice, band 1 may exhibit less signal attenuation than band 2 due to its lower frequency; however, the difference is not considered in this dissertation. If \( B_s \) increases to a certain point, \( C_{\text{WF}}^{(2)}(\psi, \psi, B_s) = C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s) \). The corresponding band separation is called the critical band separation denoted by \( B_{s,c} \). If \( B_s > B_{s,c} \), \( C_{\text{WF}}^{(2)}(\psi, \psi, B_s) < C_{\text{WF}}^{(2)}(\psi f_1/f_a, \psi, B_s) \). The natural beam focus angles in these two regimes are

\[
\psi_{F,sub}^* (\psi, B_s) = \begin{cases} 
\psi, & B_s \leq B_{s,c}, \\
\frac{f_2}{f_a} \psi \text{ or } \frac{f_2}{f_a} \psi, & B_s > B_{s,c},
\end{cases} \tag{5.18}
\]

where the subscript \text{sub} denotes sub-optimal.

The problem of determining an effective, though sub-optimal in general, beam focus angle then becomes identifying the critical band separation \( B_{s,c} \). The critical band separation \( B_{s,c} \) can be obtained by numerically solving

\[
C_{\text{WF}}^{(2)} \left( \psi \frac{f_1}{f_a}, \psi, B_{s,c} \right) = C_{\text{WF}}^{(2)} \left( \psi, \psi, B_{s,c} \right), \tag{5.19}
\]

for \( B_{s,c} \), where

\[
f_1 = f_a - \frac{B_{s,c}}{2}. \tag{5.20}
\]

### 5.2.3.2 Approximating the Critical Band Separation

The critical band separation \( B_{s,c} \) can be obtained through numerical optimization if the number of samples is large. In this section, an approximation of the critical band separation is used.
Within each band, the beam squint may not be a dominant factor in channel capacity, especially if the bandwidth $B$ is small. The array gain variation within band is smaller than that among bands. Therefore, we use the following approximation.

**Approximation 1.** Beam squint is ignored within a band for both the Aggregation Regime and Disaggregation Regime, that is, different subcarriers within each band have the same array gain, and the channel capacity of each band can be approximated as the channel capacity without beam squint in that band.

We note that beam squint is still considered across different bands. Water filling is used to allocate power to the bands to maximize the total channel capacity.

$$C^{(2)}_{\text{WF}}(\psi_F, \psi, B_s) \approx B \log \left( 1 + \frac{P^*_1 |gr g_p g (\psi f_1 / f_a - \psi_F)|^2}{B \sigma^2} \right) + B \log \left( 1 + \frac{P^*_2 |gr g_p g (\psi f_2 / f_a - \psi_F)|^2}{B \sigma^2} \right),$$

(5.21)

where $P^*_1$ and $P^*_2$ are power allocated to band 1 and band 2 according to water filling, respectively, similar to (5.7).

If the beam focuses on the center of the two bands, i.e., $\psi_F = \psi$, equal power should be allocated to both bands due to symmetry. Therefore, the channel capacity can be approximated as

$$C^{(2)}_{\text{WF}}(\psi, \psi, B_s) \approx 2B \log \left( 1 + \frac{P |gr g_p g (\psi f_1 / f_a - \psi)|^2}{2B \sigma^2} \right)$$

$$= 2B \log \left( 1 + \frac{P |gr g_p g (\psi f_a / 2f_a)|^2}{2B \sigma^2} \right).$$

(5.22)
Combining (5.19), (5.21), and (5.22),

\[
2B \log \left( 1 + \frac{P_1 |g_r g_p \sqrt{N}|^2}{B \sigma^2} \right) = B \log \left( 1 + \frac{P_2 |g_r g_p (\psi_{B_{s,c}}/f_a)|^2}{B \sigma^2} \right).
\]

(5.23)

**Lemma 1.** With Approximation 1, the critical band separation \( B_{s,c} \) is proportional to the aggregate center frequency \( f_a \) and inversely proportional to the absolute value of AoA \( \psi \).

The proof is also straightforward. \( B_{s,c} \) only appears in (5.23) together with \( f_a \) as \( \frac{B_{s,c}}{f_a} \) and together with \( \psi \) as \( B_{s,c} \psi \). Note that Lemma 1 needs not to apply to (5.19), but does so approximately if the bandwidth \( B \) is small.

Let

\[
\text{SNR}_r = \frac{P |g_r g_p|^2}{B \sigma^2}.
\]

(5.24)

\( \text{SNR}_r \) measures the signal-to-noise ratio before the receiver array.

**Theorem 1.** With Approximation 1, the critical band separation \( B_{s,c} \) monotonically increases as \( \text{SNR}_r \) increases.

The proof of Theorem 1 is in Appendix A.1. Theorem 1 shows that with lower SNR, the system is easier to enter Disaggregation Regime.

5.2.3.3 Approximating the Critical AoA

If the band separation \( B_s \), aggregate center frequency \( f_a \), and \( \text{SNR}_r \) are fixed, the optimal beam focus angle depends on the AoA \( \psi \). In this subsection, we also use Approximation 1. Similar to the critical band separation, there is also a positive
critical AoA defined as $\psi_c$. The channel capacities with various AoAs follow

\[
\begin{align*}
C_{WF}^{(2)}(\psi_c, \psi_f, B_s) &> C_{WF}^{(2)}(\psi_c f_1/f_a, \psi_c, B_s), \ |\psi| < \psi_c, \\
C_{WF}^{(2)}(\psi_c, \psi_c, B_s) &> C_{WF}^{(2)}(\psi_c f_1/f_a, \psi_c, B_s), \ |\psi| = \psi_c, \\
C_{WF}^{(2)}(\psi_c, \psi_c, B_s) &< C_{WF}^{(2)}(\psi_c f_1/f_a, \psi_c, B_s), \ |\psi| > \psi_c.
\end{align*}
\]

If the absolute value of the AoA is smaller than $\psi_c$, it is better to focus on the center of the two bands; if the absolute value of the AoA is larger than $\psi_c$, it is better to focus on the center of one of the two bands. The sub-optimal beam focus angle should follow

\[
\psi_{F,\text{sub}}^* = \begin{cases} 
\psi, & |\psi| \leq \psi_c, \\
\frac{L}{f_a} \psi, & |\psi| > \psi_c,
\end{cases}
\]

where $i = 1, 2$. Note that the critical AoA $\psi_c$ could be larger than 1. In this case, $|\psi|$ is always smaller than $\psi_c$, and the beam focus angle should focus on the center of the two bands, i.e., $\psi_F = \psi$, to maximize the total channel capacity.

Similarly, with Approximation 1, the critical AoA $\psi_c$ is inversely proportional to the band separation $B_s$, and proportional to the aggregate center frequency $f_a$. With Approximation 1, the critical AoA $\psi$ monotonically increases as SNR increases.

5.2.3.4 Criterion for Whether or Not to Aggregate

If the beam focuses on the center of one of the two bands, for example, band 1, the power allocated to band 2 is very small, and band 2 can be ignored in practice. Therefore, we use the following approximation.

**Approximation 2.** In addition to Approximation 1, in the Disaggregation Regime, we further ignore the band with smaller array gain.

In our case, band 2 is ignored and the total channel capacity can be approximated...
as

\[ C_{WF}^{(2)}(\psi_1 f_{a}, \psi, B_s) \approx B \log \left( 1 + \frac{P|g_T g_P \sqrt{N}|^2}{B \sigma^2} \right). \]  

(5.27)

To solve for \( B_{s,c} \), combine (5.19), (5.24), and (5.27) to obtain

\[ \left( 1 + \frac{\text{SNR}_r |g(\psi_2 f_{a,c})|^2}{2} \right)^2 = 1 + \text{SNR}_r N. \]  

(5.28)

Therefore,

\[ \frac{B_{s,c} |\psi|}{f_{a}} = 2g^{-1} \left( \sqrt{\frac{2\sqrt{\text{SNR}_r N + 1 - 2}}{\text{SNR}_r}} \right) \]  

(5.29)

where \( g^{-1}(y) \) is defined as the inverse function of \( |g(x)| \) for \( x \in [0, \frac{2}{N}] \), and \( y \in \mathbb{R}^+ \) is the magnitude of the array gain, \( g^{-1}(y) \) outputs the corresponding angle in the positive part of the main lobe. \(-g^{-1}(y)\) can be used for the angle in the negative part of the main lobe.

Fix band separation \( B_s \), the critical AoA can also be obtained from (5.29). Based on (5.29), we have the following design criterion to determine whether to aggregate two bands or to use only one of them.

**Design Criterion 3.** Given band separation \( B_s \), AoA \( \psi \), and \( \text{SNR}_r \), carrier aggregation should be applied if

\[ \frac{B_s |\psi|}{f_{a}} \leq 2g^{-1} \left( \sqrt{\frac{2\sqrt{\text{SNR}_r N + 1 - 2}}{\text{SNR}_r}} \right). \]  

(5.30)

Otherwise, it is sensible to use only one band with the beam focused on the center of that band, i.e., \( \frac{B}{f_{a}} \psi \) or \( \frac{B}{f_{a}} \psi \).
From (5.29), the critical band separation $B_{s,c}$ is a function of $\text{SNR}_r$. Define the function

$$f (\text{SNR}_r) = \sqrt{2 \sqrt{\text{SNR}_r N + 1} - 2 \text{SNR}_r},$$

(5.31)

$f (\text{SNR}_r)$ is a monotonically decreasing function of $\text{SNR}_r$ for $\text{SNR}_r > 0$, and

$$\lim_{\text{SNR}_r \to 0} f (\text{SNR}_r) = \sqrt{N},$$

(5.32)

$$\lim_{\text{SNR}_r \to \infty} f (\text{SNR}_r) = 0.$$  

(5.33)

Since $g^{-1} (y)$ is also a monotonically decreasing function of $y$ for $y \in [0, \sqrt{N}]$, $B_{s,c} \psi_{fa}$ monotonically increases as $\text{SNR}_r$ increases. Furthermore,

$$\lim_{\text{SNR}_r \to 0} \frac{B_{s,c} |\psi|}{fa} = 2g^{-1} (\sqrt{N}) = 0,$$

(5.34)

$$\lim_{\text{SNR}_r \to \infty} \frac{B_{s,c} |\psi|}{fa} = 2g^{-1} (0) = \frac{4}{N}.$$  

(5.35)

The critical AoA exhibits similar conclusions as (5.34) and (5.35) for a fixed $B_s$.

5.2.3.5 Numerical Results

Figure 5.2 illustrates the estimated optimal beam focus angles with numerical optimization and Approximation 1. The plot indicates small differences between the two methods. These results suggest that the beam focus angle derived with Approximation 1 is reasonable in practice. Figure 5.3 shows the maximum channel capacities obtained from numerical optimization and Approximation 1. The channel capacities from the two methods exhibit small differences, showing the effectiveness of the approximations.

Figure 5.4 compares the estimated critical band separations from numerical opti-
Figure 5.2. Sub-optimal beam focus angles for two-band carrier aggregation obtained from numerical optimization and Approximation 1 as functions of band separation $B_s$. $N_{BF} = 801$, AoA $\psi = 0.8$, $B = 2.5$ GHz, and $f_a = 73$ GHz.
Figure 5.3. Maximum channel capacity for two-band carrier aggregation obtained from numerical optimization and Approximation 1 as functions of band separation $B_s$. $N_{BF} = 801$, AoA $\psi = 0.8$, $B = 2.5$ GHz, and $f_a = 73$ GHz.
mization and Approximation 2. With Approximation 2, the critical band separation is calculated based on (5.29). $C_{WF}^{(2)}(\psi f_1/f_a, \psi, B_s)$ is obtained when the beam focuses on band 1, and no power is allocated to band 2 in the Disaggregation Regime. In Figure 5.4 (a), the Disaggregation Regime allows power to be allocated to band 2 for numerical optimization. The estimated critical band separations are very close if the $\text{SNR}_r$ is smaller than 5 dB, because power is not allocated to band 2 in numerical optimization due to water filling. As $\text{SNR}_r$ increases from 5 dB, there is an increasing gap between the two estimated critical band separations. The major reason is that power is allocated to band 2 in the Disaggregation Regime for numerical optimization and not by Approximation 2. To further verify this claim, we run another simulation to ensure that no power is allocated to band 2 in the Disaggregation Regime for numerical optimization. The results are illustrated in Figure 5.4 (b). The estimated critical band separations are very close, further demonstrating our claim. From Figure 5.4 (a), we can also observe that the critical band separation from numerical optimization is close to a constant if $\text{SNR}_r > 5$ dB, which could be used to estimate the critical band separation for high $\text{SNR}_r$.

5.3 Carrier Aggregation with Two-Sided Beam Squint

In this section, we consider beam squint at both the transmitter and receiver ULAs. To maximize the channel capacity, water filling power allocation is also used across bands. Based on [52], the channel capacity of all bands with beam squint at AoD $\psi_T$, transmitter beam focus angle $\psi_{T,F}$, AoA $\psi_R$, and receiver beam focus angle $\psi_{R,F}$ is

78
Figure 5.4. Critical band separations for two-band carrier aggregation of numerical optimization and Approximation 2 as functions of $\text{SNR}_r$.

$N_{BF} = 801$, AoA $\psi = 0.8$, $B = 2.5$ GHz, and $f_a = 73$ GHz. In the Disaggregation Regime, the beam focuses on band 1. Power is (a) allowed and (b) not allowed on band 2 in numerical optimization.
\[ C_{WF}^{(N_B)} (\psi_{T,F}, \psi_T, \psi_{R,F}, \psi_R) = \sum_{i=1}^{N_B} \frac{B_i}{N_{f,i}} \sum_{n=0}^{N_{f,i}-1} \log \left( 1 + \frac{P_{i,n}^* \left| g_T (\xi_{i,n} \psi_T - \psi_{T,F}) g_p g_R (\xi_{i,n} \psi_R - \psi_{R,F}) \right|^2}{B_i \sigma^2} \right), \quad (5.36) \]

where \( P_{i,n}^* \) is the power allocated to subcarrier \( n \) of band \( i \) such that

\[ P_{i,n}^* = \left( \nu - \frac{B_i \xi_{i,n}^2}{N_{f,i}} \right)^+, \quad n = 0, 1, \ldots, N_{f,i} - 1. \quad (5.37) \]

Here for \( i = 1, 2, \ldots, N_B \) and \( n = 0, 1, \ldots, N_{f,i} - 1, \)

\[ \xi_{i,n}^2 = \frac{\sigma^2}{\left| g_T (\xi_{i,n} \psi_T - \psi_{T,F}) g_p g_R (\xi_{i,n} \psi_R - \psi_{R,F}) \right|^2}, \quad (5.38) \]

and \( \nu \) is chosen such that

\[ \sum_{i=1}^{N_B} \sum_{n=0}^{N_{f,i}-1} \left( \nu - \frac{B_i \xi_{i,n}^2}{N_{f,i}} \right)^+ = P. \quad (5.39) \]

For fixed AoD \( \psi_T \) and AoA \( \psi_R \), we would like to find the optimal beam focus angle vector \( [\psi_{T,F}^*, \psi_{R,F}^*]^T \) to maximize the channel capacity in (5.36), that is,

\[ [\psi_{T,F}^*, \psi_{R,F}^*]^T = \arg \max_{[\psi_{T,F}^*, \psi_{R,F}^*]^T} C_{WF}^{(N_B)} (\psi_{T,F}, \psi_T, \psi_{R,F}, \psi_R), \quad (5.40) \]

where \( \psi_{T,F}^* \) and \( \psi_{R,F}^* \) are functions of \( \psi_T \) and \( \psi_R \). Note that \( \psi_{T,F}^* \) and \( \psi_{R,F}^* \) could be beyond the (virtual) angle range of \([-1, 1]\).
5.3.1 Numerical Optimization

The transmission power allocations are different for different beam focus angle vector \([\psi_{R,F}^*, \psi_{R,F}^*]^T\), increasing the difficulty to find the optimal beam focus angles. The optimal beam focus angles fall within the intervals

\[
R_{T,BF} = \left[ \frac{f_1}{f_a} \psi_T, \frac{f_{N_B}}{f_a} \psi_T \right],
\]

\[
R_{R,BF} = \left[ \frac{f_1}{f_a} \psi_R, \frac{f_{N_B}}{f_a} \psi_R \right],
\]

at the transmitter and receiver, respectively. If \(\psi_{T,F} < \frac{f_1}{f_a} \psi_T, \psi_{T,F} > \frac{f_{N_B}}{f_a} \psi_T, \psi_{R,F} < \frac{f_1}{f_a} \psi_R, \) or \(\psi_{R,F} > \frac{f_{N_B}}{f_a} \psi_R,\) the channel capacity cannot be maximized.

Similar to the case of one-sided beam squint, numerical optimization can also be used to obtain an effective, though sub-optimal, transmitter and receiver beam focus angle vector. In the numerical optimization, both \(R_{T,BF}\) and \(R_{R,BF}\) are sampled into finite sets \(R'_{T,BF}\) and \(R'_{R,BF}\), respectively,

\[
R'_{T,BF} = \left\{ \psi_{T,F,i} : \psi_{T,F,i} = \frac{(i - 1) (f_{N_B} - f_1)}{(N_{T,BF} - 1) f_a} \psi_T + \frac{f_1}{f_a} \psi_T, \ i = 1, 2, \ldots, N_{T,BF} \right\},
\]

\[
R'_{R,BF} = \left\{ \psi_{R,F,i} : \psi_{R,F,i} = \frac{(i - 1) (f_{N_B} - f_1)}{(N_{R,BF} - 1) f_a} \psi_R + \frac{f_1}{f_a} \psi_R, \ i = 1, 2, \ldots, N_{R,BF} \right\},
\]

where there are \(N_{T,BF}\) and \(N_{R,BF}\) samples in \(R'_{T,BF}\) and \(R'_{R,BF}\), respectively. \(N_{T,BF}\) and \(N_{R,BF}\) are chosen to be an odd number such that \(\psi_T\) and \(\psi_R\) will be included in \(R'_{T,BF}\) and \(R'_{T,BF}\), respectively.

With numerical optimization, the beam focus angle vector that achieves the
largest channel capacity is

$$\left[ \psi_{T,F,NO}^*, \psi_{R,F,NO}^* \right]^T = \underset{\{[\psi_{T,F,NO}, \psi_{R,F,NO}]^T: \psi_{T,F,NO} \in \mathbb{R}^T_{T,BF}, \psi_{R,F,NO} \in \mathbb{R}^R_{R,BF} \}}{\arg \max} C_{WF}^{(NB)} (\psi_{T,F,NO}, \psi_{T,F,NO}, \psi_{R,F,NO}, \psi_{R,F,NO}) ;$$

(5.45)

where $\psi_{T,F,NO}^*$ is the transmitter beam focus angle and $\psi_{R,F,NO}^*$ is the receiver beam focus angle. Again, the subscript NO denotes numerical optimization. This optimization problem has not been proven to be convex or concave. Therefore, the number of trials or the time complexity with the numerical optimization is $N_{T,BF}N_{R,BF}$.

Similar to the case of one-sided beam squint, if there is more structure in the band allocation, the structure can be used to reduce the complexity of the numerical optimization. For example, if all the $N_B$ bands are symmetric with respect to the aggregate center frequency $f_a$, i.e. $B_i = B_{N_B-i}$ and $f_i + f_{N_B-i} = 2f_a$ for $i = 1, 2, \ldots, N_B$, there should be two $\psi_{T,F}^*$ that are symmetric to $\psi_T$ for $\psi_{T,F}^* \neq \psi_T$ and two $\psi_{R,F}^*$ that are symmetric to $\psi_R$ for $\psi_{R,F}^* \neq \psi_R$. The two $\psi_{T,F}^*$ could overlap at $\psi_{T,F}^* = \psi_T$, and the two $\psi_{R,F}^*$ could overlap at $\psi_{R,F}^* = \psi_R$. Therefore, the range of the numerical optimization for both the transmitter and receiver can be reduced to half the case without symmetry. Correspondingly, for the same resolution, the number of samples can be reduced by a factor of two, and exhaustive search for the optimal solution can be reduced by a factor of four.

5.3.2 Carrier Aggregation of Two Symmetric Bands

In this section, we study carrier aggregation for the special case of two non-contiguous bands, band 1 and band 2. The carrier setup is the same as that in Section 5.2.2. Each has bandwidth $B$, and there are $N_f$ subcarriers in each band. The carriers satisfy $f_1 < f_2$, the center frequency $f_a$ is still in the center of the two bands as shown in (5.14), and the ratio of subcarrier frequencies to aggregate center
frequency for subcarrier \( n \) in band \( i, \ i = 1, 2 \) is the same as in (5.15). To ensure there is no overlap between the two bands, we must have \( B_s \geq B \). Note that if \( B_s = B \), the two bands can be treated as a single band.

The total channel capacity (5.36) becomes

\[
C^{(2)}_{WF} (\psi_{T,F}, \psi_T, \psi_{R,F}, \psi_R, B_s) = \frac{B}{N_f} \sum_{i=1}^{2} \sum_{n=0}^{N_f-1} \log \left( 1 + \frac{P_{i,n}}{\frac{B}{N_f} \sigma^2} g_T \left( \xi_{i,n} \psi_T - \psi_{T,F} \right) g_p g_R \left( \xi_{i,n} \psi_R - \psi_{R,F} \right) \right),
\]

(5.46)

where we have emphasized that the capacity is a function of the band separation \( B_s \) as in (5.16). The water filling power allocation is the same as (5.37). The optimal beam focus angle vector is then

\[
\begin{bmatrix} \psi_{T,F}^* \\ \psi_{R,F}^* \end{bmatrix}^T = \arg \max \left\{ \begin{bmatrix} \psi_{T,F} \\ \psi_{R,F} \end{bmatrix}^T : \psi_{T,F} \in \mathcal{R}_{T,BF}, \psi_{R,F} \in \mathcal{R}_{R,BF} \right\} C^{(2)}_{WF} (\psi_{T,F}, \psi_T, \psi_{R,F}, \psi_R, B_s).
\]

(5.47)

Again, the optimal beam focus angles could be beyond the (virtual) angle range of \([-1, 1]\).

5.3.3 To Aggregate or Not to Aggregate

Similar to the case of one-sided beam squint, it is preferable not to aggregate beyond certain values of the system parameters, such as band separation, AoD, AoA, and SNR.

5.3.3.1 Increasing Band Separation

Again, if the band separation \( B_s \) is small, the two bands perform similarly to a single band. The maximum channel capacity is achieved if both the transmitter beam focus angle and receiver beam focus angle are on the center of the two bands, that is,
\[ \psi_{T,F} = \psi_T \text{ and } \psi_{R,F} = \psi_R. \] This system is also said to be in the Aggregation Regime. The corresponding maximum channel capacity achieved is \( C_{WF}^{(2)}(\psi_T, \psi_T, \psi_R, \psi_R, B_s). \)

Using Approximation 1, the beam squint within each band is ignored, and different subcarriers within each band have the same array gain. If the band separation \( B_s \) is small, the array gains of the two bands are similar for both the transmitter and receiver arrays. Both transmitter and receiver arrays are in the Aggregation Regime. Therefore, our study of carrier aggregation with two-sided beam squint can be separated into two problems of carrier aggregation with one-sided beam squint. Applying the design criterion for one-sided beam squint to the transmitter and receiver separately, we denote the critical band separations for the transmitter and receiver as \( B_{T,s,c} \) and \( B_{R,s,c} \), respectively. Separation breaks down if either the transmitter or the receiver enters the Disaggregation Regime.

Fix \( \psi_{T,F} = \psi_T \text{ and } \psi_{R,F} = \psi_R. \) As \( B_s \) increases from \( B_s = B, \) the two bands move away from the center of the array's main lobe, the array gain decreases for both arrays and both bands, and the channel capacity decreases. Eventually, as \( B_s \) continues increasing, it is better for either the transmitter array or the receiver array to enter its Disaggregation Regime. In this case, the beam for this array can be focused on one of the bands to maximize the channel capacity. For example, if the receiver array first enters its Disaggregation Regime, we need to let \( \psi_{R,F} = \frac{L}{fa} \psi_R \) or \( \psi_{R,F} = \frac{L}{fa} \psi_R. \) We focus on band 1 for the Disaggregation Regime without loss of generality.

If the receiver focuses on band 1, its gain for band 2 is much smaller. The principle of water filling forces the transmitter to focus its power and beam on band 1, effectively driving it into its Disaggregation Regime as well. Therefore, if one array enters the Disaggregation Regime, we can expect that both will, and the critical band
separation for the system becomes

\[ B_{s,c} \approx \min \{ B_{T,s,c}, B_{R,s,c} \} . \] (5.48)

For simplicity of discussion,

- \( C_{WF}^{(2)}(\psi_T, \psi_T, \psi_R, \psi_R, B_s) \) is denoted as \( C_{WF,bb}^{(2)} \);
- \( C_{WF}^{(2)}(\psi_T f_1/f_{a}, \psi_T, \psi_R, \psi_R, B_s) \) is denoted as \( C_{WF,1b}^{(2)} \);
- \( C_{WF}^{(2)}(\psi_T, \psi_T, \psi_R f_1/f_{a}, \psi_R, B_s) \) is denoted as \( C_{WF,b1}^{(2)} \);
- \( C_{WF}^{(2)}(\psi_T f_1/f_{a}, \psi_T, \psi_R f_1/f_{a}, \psi_R, B_s) \) is denoted as \( C_{WF,11}^{(2)} \).

Subscript 1 denotes focusing on the center of band 1, and subscript \( b \) denotes focusing on the center of the two bands.

Figure 5.5 and Figure 5.6 show examples of \( C_{WF,bb}^{(2)} \), \( C_{WF,1b}^{(2)} \), \( C_{WF,b1}^{(2)} \), and \( C_{WF,11}^{(2)} \) as functions of the band separation \( B_s \). The example in Figure 5.5 has AoD \( \psi_T = 0.8 \), and the example in Figure 5.6 has AoD \( \psi_T = 0.4 \), and the remaining parameters are the same for the two examples in the two figures. In Figure 5.5, \( \psi_T = \psi_R \). Therefore, \( C_{WF,1b}^{(2)} = C_{WF,b1}^{(2)} \). The transmitter and receiver arrays have the same critical band separation, i.e., \( B_{T,s,c} = B_{R,s,c} \). The system critical band separation \( B_{s,c} \) is close to but not exactly the same as \( \min \{ B_{T,s,c}, B_{R,s,c} \} \), because of the interference from the side lobes. If \( B_s < B_{s,c} \), \( C_{WF,bb}^{(2)} \) has the largest channel capacity; otherwise, \( C_{WF,11}^{(2)} \) has the largest channel capacity. In Figure 5.6, \( \psi_T \neq \psi_R \), and the transmitter array and receiver array have different critical band separations. Unlike Figure 5.5, \( C_{WF,b1}^{(2)} \) has the largest channel capacity around \( B_{s,c} \) instead of \( C_{WF,bb}^{(2)} \) or \( C_{WF,11}^{(2)} \). We note that this situation is caused by the side lobes, which is beyond the scope of our study. In any case for \( B_s \gg B_{s,c} \), \( C_{WF,11}^{(2)} \) achieves the largest channel capacity. In both figures, the horizontal dotted lines indicate \( C_{WF}^{(1)} \), in which only one channel with bandwidth \( B \) is used, and the beam focuses on the center of the band.
Figure 5.5. Example of channel capacities $C_{WF,bb}^{(2)}$, $C_{WF,1b}^{(2)}$, $C_{WF,bb}^{(2)}$, and $C_{WF,11}^{(2)}$ for two-band carrier aggregation with two-sided beam squint as functions of the band separation $B_s$. The system critical band separation from numerical search is $B_{s,c} = 7$ GHz. The horizontal dotted line is $C_{WF}^{(1)}(\psi_T, \psi_T, \psi_R, \psi_R)$, in which only one channel with bandwidth $B$ is used and the beam focuses on the center of the band. If $B_s < B_{s,c}$, $C_{WF,bb}^{(2)}$ has the largest channel capacity; if $B_s > B_{s,c}$, $C_{WF,11}^{(2)}$ has the largest channel capacity. The number of antennas in the array $N_T = N_R = 32$, AoD and AoA $\psi_T = \psi_R = 0.8$, bandwidth $B = 2.5$ GHz, aggregate center frequency $f_a = 73$ GHz, and signal-to-noise ratio $\frac{|p|/\sigma^2}{B\sigma^2} = 0$ dB.
Figure 5.6. Example of channel capacities $C_{WF,bb}^{(2)}$, $C_{WF,1b}^{(2)}$, $C_{WF,bb}^{(2)}$, and $C_{WF,11}^{(2)}$ for two-band carrier aggregation with two-sided beam squint as functions of the band separation $B_s$. The system critical band separation from numerical search is $B_{s,c} = 7.8$ GHz. The horizontal dotted line is

$C_{WF}^{(1)}(\psi_T, \psi_T, \psi_R, \psi_R)$, in which only one channel with bandwidth $B$ is used and the beam focuses on the center of the band. Generally, if $B_s < B_{s,c}$, $C_{WF,bb}^{(2)}$ has the largest channel capacity; if $B_s > B_{s,c}$, $C_{WF,11}^{(2)}$ has the largest channel capacity. However, $C_{WF,1b}^{(2)}$ is an outlier around $B_{s,c}$, because of side lobe. Therefore, the outlier is not considered. The number of antennas in the array $N_T = N_R = 32$, AoD $\psi_T = 0.4$, AoA $\psi_R = 0.8$, bandwidth $B = 2.5$ GHz, aggregate center frequency $f_a = 73$ GHz, and signal-to-noise ratio $\frac{P|g_0|^2}{B\sigma^2} = 0$ dB.
The natural beam focus angle vector in these two regimes is

\[
\begin{bmatrix}
\psi_{T,F,\text{sub}}^* \\
\psi_{R,F,\text{sub}}^*
\end{bmatrix}^T = \begin{cases}
\begin{bmatrix}
\psi_T \\
\psi_R
\end{bmatrix}^T, & B_s \leq B_{s,c} \\
\begin{bmatrix}
\frac{f_a}{f} \psi_T \\
\frac{f_a}{f} \psi_R
\end{bmatrix}^T \text{ or } \begin{bmatrix}
\frac{f_a}{f} \psi_T \\
\frac{f_a}{f} \psi_R
\end{bmatrix}^T, & B_s > B_{s,c}
\end{cases}
\] (5.49)

The problem of determining an effective, though sub-optimal in general, beam focus angle vector then becomes identifying the critical band separation \(B_{s,c}\). The critical band separation can be obtained by numerically solving \(C_{WF,bb}^{(2)} = C_{WF,11}^{(2)}\), i.e.,

\[
C_{WF}^{(2)}(\psi_T, \psi_T, \psi_R, \psi_R, B_s) = C_{WF}^{(2)}(\psi_T f_1 / f_a, \psi_T, \psi_R f_1 / f_a, \psi_R, B_s),
\] (5.50)

with \(f_1 = f_a - \frac{B_{s,c}}{2}\), or approximating it using (5.48).

### 5.3.3.2 Criterion for Whether or Not to Aggregate

If the beams focus on the center of one of the bands, for example, band 1, the power allocated to the other band is very small and can essentially be ignored. Therefore, we apply Approximation 2, that is, in the Disaggregation Regime, we further ignore the band with smaller array gain in addition to Approximation 1.

Define

\[
\begin{align*}
\text{SNR}_1 &= \frac{P \left| g_p g_R \left( \frac{f_a}{f} \psi_R - \psi_R \right) \right|^2}{B \sigma^2}, \\
\text{SNR}_2 &= \frac{P \left| g_p g_T \left( \frac{f_a}{f} \psi_T - \psi_T \right) \right|^2}{B \sigma^2}.
\end{align*}
\] (5.51, 5.52)

To obtain the critical band separation for the whole system, the critical band separations of each array is calculated separately assuming there is no beam squint in the other one. From Design Criterion 3 for one-sided beam squint, we have Design Criterion 4 for the two-sided beam squint to determine to aggregate or not to
aggregate.

**Design Criterion 4.** Given band separation $B_s$, AoD $\psi_T$, AoA $\psi_R$, and SNR$_1$ and SNR$_2$, we have two inequalities

$$\frac{B_s |\psi_T|}{f_a} \leq 2g^{-1} \left( \sqrt{\frac{2\sqrt{\text{SNR}_1 N_T + 1} - 2}{\text{SNR}_1}} \right),$$  \hspace{1cm} (5.53)

$$\frac{B_s |\psi_R|}{f_a} \leq 2g^{-1} \left( \sqrt{\frac{2\sqrt{\text{SNR}_2 N_R + 1} - 2}{\text{SNR}_2}} \right),$$  \hspace{1cm} (5.54)

where $N_T$ and $N_R$ are number of antennas in the transmitter array and receiver array, respectively. If both (5.53) and (5.54) are satisfied, carrier aggregation should be used with beam focus angles $\psi_T$ and $\psi_R$ at the transmitter and receiver, respectively. If either (5.53) or (5.54) is not satisfied, it is sensible to use only one band with the beam focus angles to be on the center of the used band, i.e., $\frac{f_1}{f_a} \psi_T$ and $\frac{f_1}{f_a} \psi_R$ for the transmitter and receiver, or $\frac{f_2}{f_a} \psi_T$ and $\frac{f_2}{f_a} \psi_R$ for the transmitter and receiver, respectively.

Either band 1 or 2 can be used. Usually, band 1 is preferred due to its lower frequency and smaller signal attenuation.

### 5.4 Summary

We set up a one-sided carrier aggregation model for analog beamforming to maximize the overall channel capacity through the alignment of beam focus angle. Numerical optimization was used to obtain a sub-optimal beam focus angle because of the mathematical difficulty of solving for the optimal one. As a special case, we studied the carrier aggregation problem with two symmetric bands. In this scenario, a practical selection of the beam focus angle is either focusing on the center of two bands (Aggregation Regime) or one of the bands (Disaggregation Regime), depending
on the band separation, AoA, and SNR. Design criteria were developed to determine whether to use carrier aggregation or not and to approximate the critical parameter values beyond which we do not aggregate.

We also studied two-sided carrier aggregation with beam squint. For arbitrary combinations of multiple bands, numerical optimization was also proposed to obtain sub-optimal beam focus angles for the transmitter and receiver. For the case of two symmetric bands, the two-sided carrier aggregation problem was treated as two one-sided carrier aggregation problems. If either the transmitter or the receiver enters the Disaggregation Regime, the whole system enters Disaggregation Regime, according to the proposed design criteria.
CHAPTER 6

BEAMFORMING CODEBOOK WITH CHANNEL CAPACITY CONSTRAINT

This chapter treats the problem of designing a set of beams, called a beamforming codebook, to cover a given range of AoD or AoA. The beams (codewords) in the codebook should meet some minimum performance requirements, such as minimum channel capacity or minimum array gain. Current codebook design algorithms do not take beam squint into consideration, but as we have seen in Section 4.1, beam squint reduces channel capacity. Therefore, in this chapter, we design and analyze beamforming codebooks with consideration of beam squint subject to a constraint on the minimum channel capacity.

6.1 Definition of a Beamforming Codebook

The system setup is the same as that for a ULA in Section 4.1. The (virtual) angle range to be covered by the codebook is denoted as \( \mathcal{R}_T \), which we consider to be symmetric to the broadside,

\[
\mathcal{R}_T = [-\psi_m, \psi_m],
\]  
(6.1)

where \( 0 < \psi_m \leq 1 \).

The coverage of a fine beam \( \mathbf{w} \) with beam focus angle \( \psi_F(\mathbf{w}) \) and target rate \( C_t \) is the set of (virtual) angles

\[
\mathcal{R}_c(\mathbf{w}, C_t, b) := \{ \psi \in \mathcal{R}_T : C_{BS}(\psi_F(\mathbf{w}), \psi, b) \geq C_t \},
\]  
(6.2)
where $C_{\text{BS}}(\psi_F(w), \psi, b)$ is defined in (4.4). Selection of the capacity threshold $C_t$ depends on the system design requirements. Generally, we can select the channel capacity without beam squint as a benchmark for $C_t$. For example, select

$$C_t(r) = B \log \left( 1 + \frac{(rg_m)^2 P_R}{B\sigma^2} \right) = B \log \left( 1 + \frac{r^2NP_R}{B\sigma^2} \right),$$

where $0 < r < 1$, $P_R$ is the received power within bandwidth $B$ at one antenna before beamforming defined in (4.1), and $g_m$ is the maximum array gain defined in (3.10). A reasonable choice is $C_{t,3\text{dB}} = C_t \left( \frac{\sqrt{2}}{2} \right)$. $C_{t,3\text{dB}}$ is the capacity without beam squint at the edge of 3 dB beamwidth, i.e., where the array gain is 3 dB below $g_m$.

The codebook is defined as the set of fine beams

$$C_{\text{cap}} := \left\{ w_1, w_2, \ldots, w_M : \bigcup_{i=1}^{M} \mathcal{R}_c(w_i, C_t, b) \supseteq [-\psi_m, \psi_m] \right\},$$

where subscript cap denotes capacity. The $i$th beam in the codebook has weights $w_i$ and beam focus angle $\psi_{i,F} = \psi_F(w_i)$. For convenience, we order the beam focus angles such that

$$-\psi_m \leq \psi_{1,F} < \psi_{2,F} < \cdots < \psi_{M,F} \leq \psi_m.$$

The codebook size $|C_{\text{cap}}| = M$. Let $\Omega_c$ be the set of all codebooks that meet the requirements of (6.4) for fixed values of $N$, $b$, $C_t$, $P_R/\sigma^2$, and $\psi_m$. There is an infinite number of codebooks in $\Omega_c$. Among them, we focus on the ones that achieve the minimum codebook size. The problem that needs to be solved is to find

$$C_{\text{cap},\text{min}} = \arg \min_{C_{\text{cap}} \in \Omega_c} |C_{\text{cap}}|.$$
6.2 Properties of the Smallest Codebook

Define the left and right edges of the coverage of beam $i$ as

$$
\psi_{i,l} = \min R_c(w_i, C_t, b), \quad (6.7)
$$

$$
\psi_{i,r} = \max R_c(w_i, C_t, b), \quad (6.8)
$$

respectively. The corresponding beamwidth is

$$
\Delta \psi_i = \psi_{i,r} - \psi_{i,l}. \quad (6.9)
$$

To achieve the minimum codebook size, there should be as little overlap as possible in the coverages among all the beams. One solution is

$$
R_c(w_i, C_t, b) \cap R_c(w_j, C_t, b) = \emptyset, \quad \text{for } i \neq j, \quad (6.10)
$$

so that

$$
\begin{align*}
\psi_{1,l} & \leq -\psi_m \\
\psi_{i+1,l} &= \psi_{i,r}, \quad i = 1, 2, \ldots, M - 1 \\
\psi_{M,r} & \geq \psi_m
\end{align*} \quad (6.11)
$$

The codebook can therefore be designed by aligning the beams one by one.

6.3 Codebook Design

There could be multiple codebook designs meeting the requirements in (6.11). Among them, we are more interested in the ones that are symmetric to the broadside. Before beam alignment, it is unknown whether the codebook with odd or even number of beams achieves the minimum codebook size. The codebook with odd number of
beams is called odd-number codebook, denoted as \( C_\text{o} \); the codebook with even number of beams is called even-number codebook, denoted as \( C_\text{e} \). Both the odd and even cases should be considered.

6.3.1 Odd-Number Codebook

In the odd-number codebook \( C_\text{o} \), \( M \) is odd. Beam \((\frac{M+1}{2})\) is in the middle of all beams, and its beam focus angle \( \psi_{\frac{M+1}{2},F} = 0 \). Given \( \psi_{i,F} \), \( \psi_{i,r} \) can be derived by solving

\[
C_{\text{BS}}(\psi_{i,F}, \psi, b) = C_t, \tag{6.12}
\]

where \( C_{\text{BS}}(\psi_{i,F}, \psi, b) \) is defined in \([4.4]\). Because of the symmetry of the main lobe with respect to the broadside, there should be two solutions. The larger one is \( \psi_{i,r} \).

We need to note that there could be no solution if the fractional bandwidth is larger than a threshold, which will be shown in Section \([6.4]\). If there is no solution in solving for \( \psi_{i,r} \), no codebook exists. Because there are no closed-form solutions to this equation, numerical methods must be applied. First, we find the beamwidth for \( C_t \) without beam squint, \( \Delta \psi_{C_t} \), which is the difference between the two solutions of \( C_{\text{NBS}}(\psi_F, \psi) = C_t \), where \( C_{\text{NBS}}(\psi_F, \psi) \) is defined in \([4.5]\). All beams without beam squint have the same beamwidth in virtual angle. From Fact 1, \( \psi_{i,r} \) must satisfy

\[
\psi_{i,F} \leq \psi_{i,r} < \psi_{i,F} + \frac{\Delta \psi_{C_t}}{2}. \tag{6.13}
\]

As \( \psi \) increases from \( \psi_{i,F} \) to \( \psi_{i,F} + \frac{\Delta \psi_{C_t}}{2} \), \( C_{\text{BS}}(\psi_{i,F}, \psi, b) \) monotonically decreases. Therefore, \( \psi_{i,r} \) can be found through binary search in the interval \([\psi_{i,F}, \psi_{i,F} + \frac{\Delta \psi_{C_t}}{2}]\) \([14]\).

After the right edge of beam \((\frac{M+1}{2})\), \( \psi_{\frac{M+1}{2},r} \), is obtained, the left edge of beam
\(\frac{(M+3)}{2}\) is

\[
\psi_{\frac{M+3}{2},l} = \psi_{\frac{M+1}{2},r}.
\]  

(6.14)

The next step is to find \(\psi_{\frac{M+3}{2},F}\) given \(\psi_{\frac{M+3}{2},l}\). Suppose \(\psi_{i,l}\) is known, \(\psi_{i,F}\) must follow

\[
\psi_{i,l} \leq \psi_{i,F} < \psi_{i,l} + \frac{\Delta \psi_{C_t}}{2},
\]

(6.15)

where \(C_{BS}(\psi_{i,F}, \psi_{i,l}, b)\) monotonically decreases as \(\psi_{i,F}\) increases. Thus, \(\psi_{i,F}\) can be also obtained by numerically solving (6.12) with binary search. Note that if there is no solution in solving for \(\psi_{i,F}\), no codebook exists. \(\psi_{\frac{M+3}{2},F}\) can be determined from \(\psi_{\frac{M+3}{2},l}\).

Repeat the process to continue beam alignment on the right of the broadside. Finally, beam \(M\) satisfies

\[
\begin{align*}
\psi_{M,l} &< \psi_m, \\
\psi_{M,r} &\geq \psi_m.
\end{align*}
\]

(6.16)

Beam 1 to \(\frac{M-1}{2}\) are symmetric to beam \(M\) to \(\frac{M+3}{2}\) with respect to the broadside, respectively. The design of odd-number codebook \(C_o\) is completed.

6.3.2 Even-Number Codebook

In the even-number codebook \(C_e\), \(M\) is even. The left edge of beam \(\frac{(M+2)}{2}\) is on the broadside, i.e., \(\psi_{\frac{M+2}{2},l} = 0\). \(\psi_{\frac{M+2}{2},F}\) can be obtained by numerically solving (6.12) given \(\psi_{\frac{M+2}{2},l}\), and then \(\psi_{\frac{M+2}{2},r}\) can be obtained by numerically solving (6.12) given \(\psi_{\frac{M+2}{2},F}\). \(\psi_{\frac{M+4}{2},l} = \psi_{\frac{M+2}{2},r}\). Repeat the process to continue beam alignment on the right side of the broadside. Finally, beam \(M\) must satisfy the requirement in (6.16). Beam 1 to \(\frac{M}{2}\) are symmetric to beam \(M\) to \(\frac{M+2}{2}\) with respect to the broadside, respectively.
The design of even-number codebook $C_e$ is completed.

6.3.3 Codebook Design Algorithm

Either the odd-number codebook or the even-number codebook achieves the minimum codebook size. We select the one with smaller codebook size. The minimum codebook size $|C_{cap}|_{\text{min}} = \min (|C_o|, |C_e|)$.

A codebook design algorithm is summarized in Algorithm 1. The codebook design is divided into two situations, odd codebook size $|C|_o$ and even codebook size $|C|_e$. Procedure 1 in Algorithm 1 designs the odd-number codebook $C_o$; Procedure 2 designs the even-number codebook $C_e$. Between $C_o$ and $C_e$, the one with smaller codebook size is selected as the final designed codebook. Note that for simplicity, $\psi'_{k,F}$, $\psi'_{k,r}$ and $\psi_{k,r}$, $k = 1, 2, ..., \psi_{i,l}$ are defined in Algorithm 1, which have different order from (6.5). After designing the codebook with Algorithm 1, the beams can be aligned with beam focus angles in increasing order to recover the order in (6.5).

6.4 Constraint of the Fractional Bandwidth $b$

Fix the ratio $\frac{P_B}{B\sigma^2}$. As the fractional bandwidth $b$ increases, the channel capacity decreases as shown in Chapter 4.1 $R_x(w, C_t, b)$ decreases, and the minimum codebook size increases. As $b$ increases further, $R_x(w, C_t, b)$ could be an empty set around $\phi_F = \phi_m$, that is, no solution exists in solving for $\psi_{i,r}$ given $\psi_{i,F}$ or solving for $\psi_{i,F}$ given $\psi_{i,l}$. Therefore, no codebook exists. In this case, even the maximum channel capacity $C_{BS}(\psi_m, \psi_m, b)$ is smaller than or equal to $C_t$, and no beam can cover the angle range around $\psi_m$. Therefore, $b$ is upper bounded. Denoted the upper bound as $b_{sup}$. $b_{sup}$ can be determined by numerically solving equation,

$$C_{BS}(\psi_m, \psi_m, b) = C_t(r), \quad (6.17)$$
Algorithm 1 Codebook Design with Beam Squint

1: \textbf{procedure} 1(Odd-Number Codebook)
2: \hspace{0.5em} \textbf{Let} \( k = 1 \)
3: \hspace{0.5em} \textbf{Align the first beam so that} \( \psi'_{1,F} = 0 \)
4: \hspace{0.5em} \textbf{Derive} \( \psi'_{1,r} \) \textbf{from} \( \psi'_{1,F} \)
5: \hspace{0.5em} \textbf{while} \( \psi'_{k,r} < \psi_m \) \textbf{do}
6: \hspace{1em} \textbf{Let} \( \psi'_{k+1,l} = \psi'_{k,r} \)
7: \hspace{1em} \textbf{Derive} \( \psi'_{k+1,F} \) \textbf{from} \( \psi'_{k+1,l} \) \textbf{by numerically solving} (6.15)
8: \hspace{1em} \textbf{Derive} \( \psi'_{k+1,r} \) \textbf{from} \( \psi'_{k+1,F} \) \textbf{by numerically solving} (6.15) \( \psi'_{k+2,F} = -\psi'_{k+1,F} \)
9: \hspace{1em} \textbf{Align two beams with} \( \psi'_{k+1,F} \) \textbf{and} \( \psi'_{k+2,F} \)
10: \hspace{0.5em} \textbf{end while}
11: \hspace{0.5em} \textbf{end procedure}

15: \textbf{procedure} 2(Even-Number Codebook)
16: \hspace{0.5em} \textbf{Let} \( k = 0, \psi'_{0,r} = 0 \)
17: \hspace{0.5em} \textbf{while} \( \psi'_{c,r} < \psi_m \) \textbf{do}
18: \hspace{1em} \textbf{Let} \( \psi'_{k+1,l} = \psi'_{k,r} \)
19: \hspace{1em} \textbf{Derive} \( \psi'_{k+1,F} \) \textbf{from} \( \psi'_{k+1,l} \) \textbf{by numerically solving} (6.15)
20: \hspace{1em} \textbf{Derive} \( \psi'_{k+1,r} \) \textbf{from} \( \psi'_{k+1,F} \) \textbf{by numerically solving} (6.15) \( \psi'_{k+2,F} = -\psi'_{k+1,F} \)
21: \hspace{1em} \textbf{Align two beams with} \( \psi'_{k+1,F} \) \textbf{and} \( \psi'_{k+2,F} \)
22: \hspace{0.5em} \textbf{end while}
23: \hspace{0.5em} \textbf{end procedure}

27: \( |\mathcal{C}_c|_{\text{min}} = \min (|\mathcal{C}_o|, |\mathcal{C}_e|) \)
28: \textbf{Select the Procedure that achieves} \( |\mathcal{C}_c|_{\text{min}} \)

where \( C_{BS}(\psi_F, \psi, b) \) is defined in (4.4), and \( C_t(r) \) is defined in (6.3). Note that if there is no solution to (6.17), the upper bound does not exist, which could happen if \( C_t(r) \) is too small. For example, if \( C_t = 0 \), no solution to (6.17) can be found.

As \( b \) approaches \( b_{\text{sup}} \), \( \Delta \psi_1 \to 0 \), \( \Delta \psi_M \to 0 \), and thus, \( |\mathcal{C}_c|_{\text{min}} \to \infty \). The upper bound cannot be achieved, i.e., \( b < b_{\text{sup}} \). Suppose \( b = b_{\text{sup}} \) and \( M \) is a finite number. Then, \( R_c(w_M, C_t, b) = \psi_m \), and there will be a gap between \( R_c(w_{M-1}, C_t, b) \) and \( R_c(w_M, C_t, b) \). No codebook exists in this case.

Figure 6.1 illustrates examples of \( b_{\text{sup}} \) as a function of the number of antennas.
Figure 6.1. Upper bound of the fractional bandwidth $b_{\text{sup}}$ as a function of the number of antennas $N$. Here $\psi_m = 1$, the target coverage angle range $\mathcal{R}_T = [-1, 1]$, $r = \frac{\sqrt{2}}{2}$, the number of subcarriers in an OFDM symbol $N_f = 2048$, and $\frac{P_R}{B\sigma^2} = 0$ dB.

As the number of antennas $N$ increases, $b_{\text{sup}}$ decreases. No codebook exists when $b \geq b_{\text{sup}}$ for a given $N$.

Curve fitting with multiple setups shows that with fixed $r$,

$$b_{\text{sup}} \approx \frac{a}{N}, \quad (6.18)$$

where $a$ is a constant. For example, $b_{\text{sup}} \approx 3.04/N$ for $\frac{P_R}{B\sigma^2} = 0$ dB, $r = \frac{\sqrt{2}}{2}$ and $C_t = C_{t,3\text{dB}}$.

6.5 Numerical Results

6.5.1 Improvement of Channel Capacity

Algorithm 1 can provide a codebook with minimum channel capacity $C_t(r)$. A traditional codebook design algorithm allocates a beam to cover $\mathcal{R}_e(w, C_t(r), 0)$ and
does not account for beam squint. For $\psi$ in this set, the channel capacity would be $C_{BS}(\psi_F, \psi, b)$, and therefore minimum channel capacity

$$C_{BS,\min}(\psi_F, b, r) = \min_{\psi \in \mathcal{R}_c(w, C_t(r), 0)} C_{BS}(\psi_F, \psi, b),$$

(6.19)

where is $\mathcal{R}_c(w, C_t(r), b)$ defined in (6.2). $C_{BS,\min}(\psi_F, b, r)$ is achieved if $\psi$ is the on the beam edge with larger magnitude.

Define the capacity improvement ratio as

$$I(\psi_F, b, r) = \frac{C_t(r) - C_{BS,\min}(\psi_F, b, r)}{C_{BS,\min}(\psi_F, b, r)}.$$  

(6.20)

For a given fractional bandwidth $b$, the maximum capacity improvement ratio for $\psi_F \in [-1, 1]$ is

$$I_{\max}(b, r) = \max_{\psi_F \in [-1, 1]} I(\psi_F, b, r).$$

(6.21)

$I_{\max}(b, r)$ is achieved if $\psi_F = \pm 1$.

Figure 6.2 illustrates that $I(\psi_F, b, \sqrt{\frac{2}{2}})$ increases as $N$ and $\psi_F$ increase, where $r = \sqrt{\frac{2}{2}}$. The figure also demonstrates that the effect of beam squint increases as $\psi_F$ increases. Figure 6.3 illustrates that $I_{\max}(b, \sqrt{\frac{2}{2}})$ increases as $b$ increases, that is, the improvement of the codebook design algorithm will be more significant if the bandwidth is larger. For small $N$, the improvement is negligible.

6.5.2 Codebook Size

Figure 6.4 shows examples of the minimum codebook size $|\mathcal{C}|_{\min}$ as a function of the number of antennas $N$ for $\psi_m = 1$, i.e., the maximum covered AoA/AoD of the codebook is $90^\circ$. The minimum codebook size increases as $N$ or $b$ increases. Note that the minimum codebook size also varies with $\frac{P_R}{\sigma^2}$; however, our simulations suggest
Figure 6.2. Capacity improvement ratio $I(\psi_F, b, r)$ of (6.20) in a fine beam as a function of beam focus angle $\psi_F$. Here $r = \sqrt{2}$, the number of subcarriers in an OFDM symbol $N_f = 2048$, fractional bandwidth $b = 0.0342$ ($B = 2.5$ GHz, $f_c = 73$ GHz), and $\frac{P_R}{B\sigma^2} = 0$ dB.

Figure 6.3. Maximum capacity improvement ratio $I_{\text{max}}(b, r)$ of (6.21), as a function of fractional bandwidth $b$. Here $r = \sqrt{2}$, the number of subcarriers in an OFDM symbol $N_f = 2048$, and $\frac{P_R}{B\sigma^2} = 0$ dB.
Figure 6.4. Minimum codebook size $|\mathcal{C}_{\text{cap}}|_{\min}$ in a ULA required to cover $\psi \in [-1, 1]$ as a function of the number of antennas $N$. $C_t$ is set according to (6.3). Here the number of subcarriers in an OFDM symbol $N_f = 2048$, and $\frac{P_R}{B\sigma^2} = 0$ dB. Fractional bandwidth $b = 0.0179$ corresponds to $B = 0.5$ GHz and $f_c = 28$ GHz; $b = 0.0342$ corresponds to $B = 2.5$ GHz and $f_c = 73$ GHz; $b = 0.0417$ corresponds to $B = 2.5$ GHz and $f_c = 60$ GHz; $b = 0.0714$ corresponds to $B = 2$ GHz and $f_c = 28$ GHz.

that it only has a small effect on the codebook size. No codebook exists for $N \geq 42$ with $b = 0.0714$.

6.6 Summary

To compensate for the beam squint, a beamforming codebook is designed with a channel capacity constraint. The codebook size increases as the fractional bandwidth or the number of antennas in the array increases, and appears to grow unbounded beyond certain thresholds on these parameters. Furthermore, the fractional bandwidth is upper bounded for a fixed number of antennas.
CHAPTER 7

BEAMFORMING CODEBOOK DESIGN WITH ARRAY GAIN CONSTRAINT

The beamforming codebook design with channel capacity constraint depends jointly on the transmitter, receiver, and path gain. However, in practice, it is convenient for the beamforming codebook designs to be decoupled at the transmitter and receiver. One widely used beamforming codebook design criterion is to require the array gain to be larger than a threshold $g_t$ for all subcarriers \[57\]. This requirement can be imposed at the transmitter and receiver separately, possibly with different thresholds, and therefore decouple the beamforming codebook designs.

In this chapter, beamforming codebooks are developed to compensate for the beam squint of a ULA while satisfying minimum array gain constraint.

7.1 Definition of a Beamforming Codebook with Array Gain Constraint

We consider a ULA, either at the transmitter or at the receiver, with $N$ antennas and $d = \lambda_c/2$, where $d$ is the distance between adjacent antennas, and $\lambda_c$ is the wavelength of the carrier frequency as shown in Chapter \[2\]. The array gain threshold $g_t$ is given from a high-level system design. A codebook $C$ is defined as

$$C := \{w_1, w_2, \ldots, w_M\}, \quad (7.1)$$
where there are $M$ beams, the $i$th beam has weights $w_i$ and beam focus angle $\psi_{i,F}$, and the beam focus angles are in increasing order as

$$-\psi_m \leq \psi_{1,F} < \psi_{2,F} < \cdots < \psi_{M,F} \leq \psi_m.$$  

(7.2)

The codebook size $|C| = M$.

We are only interested in the main lobe of each beam in the codebook. Without consideration of beam squint, the size of the main lobe in the virtual angle $\psi$ is $\frac{4}{N}$. To cover the (virtual) angle range $\mathcal{R}_T = [-\psi_m, \psi_m]$ utilizing only the main lobe of each beam, there is a necessary condition for the codebook size

$$M > \frac{2\psi_m}{4/N} = \frac{N\psi_m}{2}.$$  

(7.3)

For a given AoA/AoD $\psi \in \mathcal{R}_T$, the minimum array gain of beam $i$ across all subcarriers within bandwidth $B$ is defined as

$$g_{w_i, \text{min}}(\psi) = \min_{\xi \in [1-b/2, 1+b/2]} g(\xi \psi - \psi_{i,F}).$$  

(7.4)

We note that $g_{w_i, \text{min}}(\psi)$ varies for beams with different beam focus angles. For a given AoA/AoD $\psi \in \mathcal{R}_T$, among the $M$ beams, only the beam with the maximum gain is used for communication. The codebook array gain for AoA/AoD $\psi$ is therefore defined as

$$g_{c}(\psi) = \max_{i=1,2,\ldots,M} \{g_{w_i, \text{min}}(\psi)\}.$$  

(7.5)

Finally, the minimum array gain of the codebook is defined as

$$g_{c, \text{min}} = \min_{\psi \in \mathcal{R}_T} g_{c}(\psi).$$  

(7.6)
The goal of the codebook design is to maximize the minimum codebook array gain $g_{c,\text{min}}$ for a given codebook size $M$. Denote $\Omega_M$ as the set of all codebooks containing $M$ beams. There is an infinite number of codebooks in $\Omega_M$. Among them, we focus on the ones that maximize $g_{c,\text{min}}$, defined as

$$C_{\text{gain}}(M) = \underset{C \in \Omega_M}{\arg\max} g_{c,\text{min}},$$  \hspace{1cm} (7.7)

The corresponding maximum for the designed codebook is defined as optimized codebook array gain $g_d$,

$$g_d(M) = \max_{C \in \Omega_M} g_{c,\text{min}}.$$  \hspace{1cm} (7.8)

As we will see in Section 7.2, there will be only one codebook for a given $M$, $M > \frac{N\psi_m}{2}$. Among all the codebooks $C_{\text{gain}}(M)$, $M = 1, 2, \ldots$, we are mainly interested in the codebook with the minimum codebook size that has array gain larger than a minimum array gain threshold $g_t$. The minimum codebook size is

$$M_{\text{min}} = \underset{M \in \mathbb{Z}^+, M > \frac{N\psi_m}{2}}{\arg\min} g_d(M) \geq g_t,$$  \hspace{1cm} (7.9)

where $\mathbb{Z}^+$ is the set of positive integers. The corresponding desired beamforming codebook is

$$C^*_{\text{gain}} = C_{\text{gain}}(M_{\text{min}}).$$  \hspace{1cm} (7.10)

The corresponding optimized codebook array gain is

$$g^*_d = g_d(M_{\text{min}}).$$  \hspace{1cm} (7.11)

$C_{\text{gain}}^*$ is the codebook we aim to design in this chapter.
7.2 Beamforming Codebook Design

The desired codebook $C_{\text{gain}}^*$ in (7.10) has minimum codebook size and achieves optimized codebook array gain in (7.11). Without beam squint, there are two steps to design the codebook $C_{\text{gain}}^*$. First, determine the minimum codebook size to cover $R_T$ while satisfying the minimum array gain criterion $g_{C,\text{min}} \geq g_t$. Second, align beams in the codebook to maximize the minimum codebook array gain $g_{C,\text{min}}$ to achieve the optimized codebook array gain $g_d$.

However, with beam squint, it is impossible to find the minimum codebook size before beam alignment. In this chapter, we first assume the codebook size $M$ is a known parameter, and use it to align the beams in the codebook to achieve the optimized codebook array gain $g_d$, which is a function of the codebook size $M$. Then, the minimum codebook size can be determined based on the constraint $g_d \geq g_t$. The codebook design procedure is illustrated in Figure 7.1.

7.2.1 Codebook Design Given Codebook Size

The coverage of beam $i$ is defined as

$$R_g(w_i) := \{\psi \in [-1,1] : g_{w_i,\text{min}}(\psi) \geq g_{C,\text{min}}\},$$

(7.12)
where \( i = 1, 2, \ldots, M \), and \( g_{c, \text{min}} \) is the minimum codebook array gain defined in (7.6). Define left edge and right edge of beam \( i \) as

\[
\psi_{i,l} = \min \mathcal{R}_g (w_i), \quad (7.13)
\]

\[
\psi_{i,r} = \max \mathcal{R}_g (w_i), \quad (7.14)
\]

respectively. Then \( \mathcal{R}_g (w_i) = [\psi_{i,l}, \psi_{i,r}] \), and the corresponding beamwidth is defined as \( \Delta \psi_i \),

\[
\Delta \psi_i = \psi_{i,r} - \psi_{i,l}. \quad (7.15)
\]

To maximize the minimum codebook array gain \( g_{c, \text{min}} \) or to achieve the optimized codebook array gain \( g_d \), the beam focus angles of the \( M \) beams in the codebook should be aligned to minimize the overlap of any two beams. Ideally,

\[
\mathcal{R}_g (w_i) \subseteq \mathcal{R}_T, \ i = 1, 2, \ldots, M, \quad (7.16)
\]

\[
\int_{-1}^{1} 1_{\mathcal{R}_g (w_i) \cap \mathcal{R}_g (w_{i+1})} (\psi) \, d\psi = 0, \ i = 1, 2, \ldots, M - 1, \quad (7.17)
\]

where \( \mathcal{R}_T \) is the target codebook coverage, and \( 1_{\mathcal{R}} (\psi) \) is an indicator function

\[
1_{\mathcal{R}} (\psi) = \begin{cases} 
1, & \text{if } \psi \in \mathcal{R}, \\
0, & \text{if } \psi \notin \mathcal{R}.
\end{cases} \quad (7.18)
\]

Note that (7.17) can be satisfied in two ways. First, \( \mathcal{R}_g (w_i) \cap \mathcal{R}_g (w_{i+1}) = \emptyset, \ i = 1, 2, \ldots, M - 1 \). Second, \( \mathcal{R}_g (w_i) \) and \( \mathcal{R}_g (w_{i+1}) \), \( i = 1, 2, \ldots, M - 1 \), intersects at a single point. Since \( \mathcal{R}_g (w_i) \) and \( \mathcal{R}_g (w_{i+1}) \), \( i = 1, 2, \ldots, M - 1 \), are both closed set, we have the following lemma.

**Lemma 2.** For given codebook size \( M \), the optimized codebook array gain \( g_d \) is
achieved if and only if:

\[
\begin{align*}
\psi_{1,l} &= -\psi_m \\
\psi_{i+1,l} &= \psi_{i,r}, \quad i = 1, 2, \ldots, M - 1 \\
\psi_{M,r} &= \psi_m
\end{align*}
\] (7.19)

In lemma 2, (7.17) is satisfied with \( R_g (w_i) \) and \( R_g (w_{i+1}) \), \( i = 1, 2, \ldots, M - 1 \), intersecting at a single point.

If \( g_d \) is achieved, all the beams in the codebook have

\[
g_{w_i,\min} (\psi_{i,l}) = g_{w_i,\min} (\psi_{i,r}) = g_d, \quad i = 1, 2, \ldots, M.
\] (7.20)

At this point, \( g_d \) is not known. We temporally assume \( g_d \) is known, and derive the beam focus angles of all \( M \) beams in the codebook as a function of \( g_d \). Then \( g_d \) can be solved with the beam edge constraints in (7.19). The beams should be aligned symmetrically with respect to the broadside. Otherwise, \( \psi_{1,l} < -\psi_m, \psi_{M,r} > \psi_m \), or there exists at least one \( i = 1, 2, \ldots, M \) such that \( \psi_{i+1,l} < \psi_{i,r} \); the coverage of at least one beam will be wasted, and \( g_d \) is not achieved.

Our procedure for codebook design depends on whether \( M \) is odd or even. If \( M \) is odd, the \( \frac{M+1}{2} \)-th beam has zero beam focus angle, i.e.,

\[
\psi_{\frac{M+1}{2},r} = 0.
\] (7.21)

If \( M \) is even, the coverages of the \( \frac{M}{2} \)-th and \( \frac{M+2}{2} \)-th beams intercept at \( \psi = 0 \), i.e.,

\[
\psi_{\frac{M}{2},r} = \psi_{\frac{M+2}{2},l} = 0.
\] (7.22)

Due to the symmetry of the codebook, we only discuss the beams with beam
focus angles larger than 0. We have restricted all subcarriers are located in the main lobe of the beam. Therefore, for $\psi_{i,r} \geq 0$, the subcarrier with highest frequency has the smallest array gain. Therefore,

$$g_{w_i, \text{min}}(\psi_{i,r}) = g((1 + b/2) \psi_{i,r} - \psi_{i,F}) = g_d. \quad (7.23)$$

Define the generalized beam edge as

$$\psi_d = g^{-1}(g_d), \quad (7.24)$$

where $g^{-1}(y), y \in \mathbb{R}^+$ is an inverse function of $|g(x)|$ for $x \in [0, \frac{2}{N}]$, and $\psi_d > 0$.

From (7.23), we can obtain

$$\psi_{i,r} = \frac{\psi_{i,F} + g^{-1}(g_d)}{1 + b/2} = \frac{2(\psi_{i,F} + \psi_d)}{2 + b}. \quad (7.25)$$

Since all subcarriers are located in the main lobe of the beam, for $\psi_{i,l} \geq 0$, the subcarrier with smallest frequency has the smallest array gain. Therefore,

$$g_{w_j, \text{min}}(\psi_{j,l}) = g((1 - b/2) \psi_{j,l} - \psi_{j,F}) = g_d. \quad (7.26)$$

Thus, we obtain

$$\psi_{i,l} = \frac{\psi_{i,F} - g^{-1}(g_d)}{1 - b/2} = \frac{2(\psi_{i,F} - \psi_d)}{2 - b}, \quad (7.27)$$

and

$$\psi_{i,F} = \psi_d + \psi_{i,l} (1 - b/2). \quad (7.28)$$

From (7.25) and (7.28), the right edge can be derived from the left edge for beam
$$i, i > \frac{M+1}{2},$$

$$\psi_{i,r} = \frac{\psi_{i,l} (1 - b/2) + 2\psi_d}{1 + b/2} = \frac{2 - b}{2 + b}\psi_{i,l} + \frac{4}{2 + b}\psi_d. \quad (7.29)$$

Similarly, the beam focus angle of beam \(i + 1, i \geq \frac{M+1}{2}\), can be expressed as

$$\psi_{i+1,F} = \psi_d + \psi_{i+1,l} (1 - b/2) = \psi_d + \psi_{i,r} (1 - b/2) = \psi_d + (1 - b/2) \frac{\psi_{i,F} + \psi_d}{1 + b/2} = \frac{2 - b}{2 + b}\psi_{i,F} + \frac{4}{2 + b}\psi_d. \quad (7.30)$$

If \(\psi_d\) is known, all the parameters of the codebook can therefore be obtained.

**Theorem 2.** Given the number of beams \(M\), the fractional bandwidth \(b > 0\), and the target codebook coverage \(\mathcal{R}_T = [-\psi_m, \psi_m]\), the generalized beam edge satisfies

$$\psi_d(M) = \begin{cases} \frac{b\psi_m}{2 - 4\frac{b}{2\pi} \left(\frac{2-b}{2\pi}\right) \frac{M-1}{2}}, & M \text{ is odd,} \\ \frac{b\psi_m}{2 - 2\left(\frac{2-b}{2\pi}\right) \frac{M}{2}}, & M \text{ is even.} \end{cases} \quad (7.31)$$

A proof is provided in Appendix A.2.

Again, the selection of \(M\) should ensure \(g(\psi_d) > 0\), which will be discussed in Section 7.2.2. The generalized beam edge \(\psi_d\) for the case without beam squint can
be obtained by taking the limit

\[
\lim_{b \to 0} \psi_d(M) = \lim_{b \to 0} \begin{cases} 
\frac{b\psi_m}{2 - \frac{4}{\pi} \frac{2b^2}{2b^2 + b}}, & \text{M is odd,} \\
\frac{b\psi_m}{2 - 2\frac{2b^2}{2b^2 + b}} M, & \text{M is even.}
\end{cases}
\]

\[
= \frac{\psi_m}{M}. \tag{7.32}
\]

The proof can be derived based on similar steps to derive \( \psi_d \). Without beam squint, all beams in the codebook have the same beamwidth.

Given \( M \), \( \psi_d \) can be calculated from (7.31). \( \psi_{i,F}, \psi_{i,l}, \) and \( \psi_{i,F} \) for \( i = 1, 2, \ldots, M \) can be obtained for odd \( M \) from (7.21), (7.25), (7.27) and (7.30). Similarly, \( \psi_{i,F}, \psi_{i,l}, \) and \( \psi_{i,F} \) for \( i = 1, 2, \ldots, M \) can be obtained for even \( M \) through (7.22), (7.25), (7.27) and (7.30). The parameters of the codebook are summarized in Table 7.1. All the parameters are expressed in terms of \( \psi_d \).

**Theorem 3.** For given codebook size \( M \), there is one and only one codebook that achieves the optimized codebook array gain \( g_d \), and this codebook is optimal.

A proof is provided in Appendix A.3. Theorem 3 demonstrates the uniqueness of the codebook that achieves \( g_d \) for a given codebook size.

### 7.2.2 Determination of the Minimum Codebook Size

The designed codebook should satisfy

\[
g_d \geq g_t, \tag{7.33}
\]

where \( g_t \) is the minimum array gain threshold. The minimum codebook size \( M_{\min} \) is determined from \( g_t \).
### TABLE 7.1

**SUMMARY OF CODEBOOK PARAMETERS**

<table>
<thead>
<tr>
<th>M</th>
<th>$\psi_{i,F}$</th>
<th>$\psi_{i,I}$</th>
<th>$\psi_{i,R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odd</td>
<td>$i &gt; \frac{M+1}{2}$</td>
<td>$\frac{2\psi_d}{b} \left[ 1 - \left( \frac{2-b}{2+b} \right)^{i - \frac{M+1}{2}} \right]$</td>
<td>$\frac{2\psi_d}{b} \left[ 1 - \frac{2}{2+b} \left( \frac{2-b}{2+b} \right)^{i - \frac{M+1}{2}} \right]$</td>
</tr>
<tr>
<td></td>
<td>$i = \frac{M+1}{2}$</td>
<td>0</td>
<td>$-\frac{2\psi_d}{2+b}$</td>
</tr>
<tr>
<td></td>
<td>$i &lt; \frac{M+1}{2}$</td>
<td>$-\frac{2\psi_d}{b} \left[ 1 - \left( \frac{2-b}{2+b} \right)^{\frac{M+1}{2} - i} \right]$</td>
<td>$-\frac{2\psi_d}{b} \left[ 1 - \frac{2}{2+b} \left( \frac{2-b}{2+b} \right)^{\frac{M+1}{2} - i} \right]$</td>
</tr>
<tr>
<td>Even</td>
<td>$i &gt; \frac{M}{2}$</td>
<td>$\frac{2\psi_d}{b} \left[ 1 - \frac{2-b}{2} \left( \frac{2-b}{2+b} \right)^{i - \frac{M}{2} - 1} \right]$</td>
<td>$\frac{2\psi_d}{b} \left[ 1 - \left( \frac{2-b}{2+b} \right)^{i - \frac{M}{2} - 1} \right]$</td>
</tr>
<tr>
<td></td>
<td>$i \leq \frac{M}{2}$</td>
<td>$-\frac{2\psi_d}{b} \left[ 1 - \frac{2-b}{2} \left( \frac{2-b}{2+b} \right)^{\frac{M}{2} - i} \right]$</td>
<td>$-\frac{2\psi_d}{b} \left[ 1 - \left( \frac{2-b}{2+b} \right)^{\frac{M}{2} - i + 1} \right]$</td>
</tr>
</tbody>
</table>
The problem becomes to determine the minimum codebook size \( M_{\text{min}} \) given \( g_t \) such that

\[
\psi_d(M_{\text{min}}) \leq g^{-1}(g_t),
\]

(7.34)

where \( \psi_d(M_{\text{min}}) \) is calculated from (7.31) given \( M_{\text{min}} \). For simplicity, further define

\[
\psi_t = g^{-1}(g_t).
\]

(7.35)

\( M_{\text{min}} \) cannot be derived directly from (7.31) because \( M_{\text{min}} \) is an integer. Instead, \( M_{\text{min}} \) can be obtained from Theorem 4.

**Theorem 4.** For a ULA, given the number of antennas \( N \) fractional bandwidth \( b \), and the minimum array gain constraint \( g_t < \sqrt{N} \), the minimum codebook size is

\[
M_{\text{min}} = \begin{cases} 
M_e, & \text{if } M_e \text{ is even,} \\
M_o, & \text{if } M_e \text{ is odd,}
\end{cases}
\]

(7.36)

where

\[
M_o = \left\lceil \frac{2 \ln \left( \frac{2+b}{4} \left( 2 - \frac{b \psi_m}{g^{-1}(g_t)} \right) \right)}{2 - b} \ln \frac{2-b}{2+b} \right\rceil + 1,
\]

(7.37)

\[
M_e = \left\lceil \frac{2 \ln \left( 1 - \frac{b \psi_m}{2g^{-1}(g_t)} \right)}{2 - b} \ln \frac{2-b}{2+b} \right\rceil,
\]

(7.38)

and \( \lceil \cdot \rceil \) denotes the ceiling function.

\( M_o \) and \( M_e \) are derived by letting \( \psi_d = \psi_t = g^{-1}(g_t) \) in (7.31). Specifically, \( M_o \) can be obtained by replacing \( M \) with \( M_o \) in (7.31) for odd \( M \), inverting the equation and applying a ceiling function; \( M_e \) can be obtained by replacing \( M \) with \( M_e \) in (7.31) for even \( M \), inverting the equation and applying a ceiling function. Note that
Algorithm 2 Codebook Design with Beam Squint

1: Calculate $\psi_t$ based on (7.35) given $g_t$
2: Calculate the minimum codebook size $M_{\text{min}}$ using Theorem 4
3: Calculate $\psi_d$ by putting $M_{\text{min}}$ into (7.31)
4: Align the beams according to the beam parameters in Table 7.1

if $g_t > \sqrt{N}$, there is no meaning for $g^{-1}(g_t)$; if $g_t = \sqrt{N}$, $g^{-1}(g_t) = 0$ is in the denominator of (7.37) and (7.38), and there is no solution for $M_o$ and $M_e$.

A proof of Theorem 4 is provided in Appendix A.4.

Given the minimum array gain threshold $g_t$, the codebook design procedure is summarized in Algorithm 2. According to Theorem 3, the codebook designed in Algorithm 2 is unique and optimal, and all the beams are fully utilized.

7.3 Limitations on System Parameters

In this section, we discuss constraints on the generalized beam edge $\psi_d$, optimized codebook array gain $g_d$, fractional bandwidth $b$, and the number of antennas $N$. No codebook exists if any parameter is beyond its constraint.

7.3.1 Generalized Beam Edge $\psi_d$ and Optimized Codebook Array Gain $g_d$

From (7.29), we can obtain the beamwidth of beam $i$,

$$\Delta_i = \psi_{i,r} - \psi_{i,l}$$

$$= -\frac{2b}{2+b} \psi_{i,l} + \frac{4\psi_d}{2+b}$$

$$= -\frac{2b}{2-b} \psi_{i,r} + \frac{4\psi_d}{2-b}.$$  \hfill (7.39)

As the beam moves away from the broad side, i.e., having larger $|\psi_{i,l}|$ or $|\psi_{i,r}|$, the beamwidth $\Delta_i$ decreases because $\psi_d$ decreases. It is possible that $\Delta_i \leq 0$ for some $i$, indicating that no codebook can meet the minimum array gain requirement. To
guarantee the existence of a codebook, we must have

$$\Delta_M = \frac{2b}{2 - b} \psi_{M,r} + \frac{4\psi_d}{2 - b} = \frac{2b}{2 - b} \psi_m + \frac{4\psi_d}{2 - b} > 0, \quad (7.40)$$

yielding the following lemma.

**Lemma 3.** For a ULA, given the generalized beam edge $\psi_d$ is lower bounded by

$$\psi_d > \frac{b\psi_m}{2}. \quad (7.41)$$

Equivalently, given the number of antennas in the array $N$ and given fractional bandwidth $b$, the optimized codebook array gain $g_d$ is upper bounded by

$$g_d = |g(\psi_d)| < \left| g\left(\frac{b\psi_m}{2}\right) \right| = \frac{\sin\left(\frac{N\pi b\psi_m}{4}\right)}{\sqrt{N} \sin\left(\frac{\pi b\psi_m}{4}\right)}. \quad (7.42)$$

Codebook size $M \to \infty$ is the condition for $\psi_d$ and $g_d$ to converge to their bounds.

The result (7.42) is derived from (7.41) based on the fact that $g(\psi_d)$ decreases as $\psi_d$ increases for $\psi_d \in \left[0, \frac{2}{N}\right]$, or it can be obtained from (7.31). Since $\psi_d$ in (7.31) decreases as $M$ increases,

$$\psi_d > \lim_{M \to \infty} \begin{cases} \frac{b\psi_m}{2 - \frac{4}{2 + b} \left(\frac{2 - b}{2 + b}\right) \frac{M - 1}{2}} & \text{if } M \text{ is odd}, \\ \frac{b\psi_m}{2 - 2\left(\frac{2 - b}{2 + b}\right) \frac{M}{2}} & \text{if } M \text{ is even.} \end{cases} \quad (7.43)$$

The case of codebook design without beam squint is equivalent to that of $b = 0$. If there is no beam squint, i.e., $b = 0$, as the codebook size $M \to \infty$, the generalized beam edge $\psi_d \to 0$, and the optimized codebook array gain $g_d \to g_m$. It
is easy to mistakenly apply this conclusion to the case with beam squint. Lemma 3 demonstrates that there is a positive lower bound for $\psi_d$ for arbitrarily large codebook size and $g_d$ has an upper bound that is smaller than $g_m$ if there is beam squint.

7.3.2 Fractional Bandwidth $b$

Based on (7.41) in Lemma 3 we have an additional lemma.

**Lemma 4.** For a ULA, given the number of antennas $N$ and optimized codebook array gain $g_d$, the fractional bandwidth $b$ is upper bounded by

$$b < \frac{2\psi_d}{\psi_m} = \frac{2g^{-1}(g_d)}{\psi_m}.$$  \hfill (7.44)

$M \rightarrow \infty$ is required for $b$ to converge to this upper bound.

As $b$ increases, the array gains of different subcarriers tend to have larger difference, reducing the beamwidth. As $b$ grows beyond $\frac{2\psi_d}{\psi_m}$, there are always subcarriers with array gain below $g_d$, and therefore, no codebook exists.

7.3.3 Number of Antennas $N$

To study the limitation of the number of antennas in the array, we first approximate the magnitude of $g(x)$ from (3.7). Again, we consider the main lobe only, i.e., $x \in \left[ -\frac{2}{N}, \frac{2}{N} \right]$, in $g(x)$. Thus,

$$\left| \frac{\pi x}{2} \right| \leq \frac{\pi}{N}. \hfill (7.45)$$

If $N$ is larger, as shown in [4], the magnitude of $g(x)$ from (3.7) can be approximated as

$$|g(x)| = \frac{\sin \left( \frac{N\pi x}{2} \right)}{\sqrt{N \sin \left( \frac{\pi x}{2} \right)}}$$
\[ a \approx \frac{\sin \left( \frac{N\pi x}{2} \right)}{\frac{\pi x \sqrt{N}}{2}} = \sqrt{N} \text{sinc} \left( \frac{N\pi x}{2} \right), \] (7.46)

where \( x \in \left[ -\frac{2}{N}, \frac{2}{N} \right] \), (a) is due to \( \sin x \approx x \) for \( x \approx 0 \), and \( \text{sinc}(\cdot) \) is sinc function defined as

\[
\text{sinc}(x) = \begin{cases} 
1, & x = 0, \\
\frac{\sin x}{x}, & \text{otherwise}.
\end{cases}
\] (7.47)

We assume \( N \geq 4 \) when using the approximation (7.46).

To illustrate the error of this approximation, define the relative approximation error of \( g(x) \) as

\[
e_a(x, N) = \left| \frac{|g(x)| - \sqrt{N} \text{sinc} \left( \frac{N\pi x}{2} \right)}{|g(x)|} \right|, \] (7.48)

where \( x \in \left[ -\frac{2}{N}, \frac{2}{N} \right] \). Since \( g(x) \) is symmetric about \( x = 0 \), we focus on the range \( x \in \left[ 0, \frac{2}{N} \right] \). The maximum relative approximation error of \( g(x) \) for \( x \in [0, 2/N] \) is

\[
e_{a, \text{max}}(N) = \max_{x \in [0, \frac{2}{N}]} e_a(x, N), \] (7.49)

As \( x \) increases, the relative approximation error of \( \sin x \approx x \) also increases. Therefore, \( e_{a, \text{max}}(N) \) is achieved if \( x = \frac{2}{N} \), i.e., \( e_{a, \text{max}}(N) = e_a \left( \frac{2}{N}, N \right) \). As \( N \) increases, \( \frac{2}{N} \) decreases, and the maximum relative approximation error, \( e_{a, \text{max}}(N) \) decreases. The discussion here will be verified with numerical results in Section 7.4.

Define

\[
r_g = \frac{g_d}{\sqrt{N}}, \] (7.50)

116
Again, \( \sqrt{N} \) is the maximum array gain in the beam. Usually, array gain is selected related to the peak gain. For example, with 3-dB beamwidth, \( g_d \) is 3 dB below the peak gain \( \sqrt{N} \), i.e., \( r_g = 1/\sqrt{2} = 0.707 \).

Using the approximation in (7.46),

\[
\psi_d \approx \frac{2}{N\pi} \text{sinc}^{-1} \left( \frac{g_d}{\sqrt{N}} \right) = \frac{2}{N\pi} \text{sinc}^{-1} (r_g),
\]

(7.51)

where \( \text{sinc}^{-1} (\cdot) \) is the inverse function of sinc function.

Combining (7.51) with (7.44), we can obtain the following lemma.

**Lemma 5.** For a ULA, given the fractional bandwidth \( b \) and fixed \( r_g = \frac{g_d}{\sqrt{N}} \), the number of antennas in the array \( N \) has a upper bound, denoted as \( N_{\text{sup}} \). We have

\[
N < N_{\text{sup}} \approx \frac{4 \text{sinc}^{-1} (r_g)}{b\pi \psi_m}.
\]

(7.52)

As \( N \) increases, the beam becomes narrower, and thus, the array gains of different subcarriers tend to have larger difference. As \( N \) grows beyond the upper bound, there are always subcarriers with array gain below \( g_d \), and therefore, no codebook exists.

### 7.4 Numerical Results

In this section, we will show numerical examples of the designed beamforming codebook as well as the constraints on \( \psi_d \), optimized codebook array gain \( g_d \), fractional bandwidth \( b \), and the number of antennas \( N \). The diminishing returns of increasing \( N \) will also be illustrated.
Figure 7.2. Example of beamforming codebook design for the minimum codebook size subject to array gain constraint. The number of antennas $N = 16$, the minimum array gain threshold $g_t = g_m / \sqrt{2}$, and $\psi_m = 1$.

Fractional bandwidth $b = 0.0342$ corresponds to $B = 2.5$ GHz at $f_c = 73$ GHz. (a) Beam squint is not considered. The minimum codebook size is 19. (b) Beam squint is considered, and codebook design is based on Algorithm 2. The minimum codebook size is 23.
7.4.1 Examples of Codebook Design

Figure 7.2 includes examples of beamforming codebook designs without and with consideration of beam squint. In the figure, the fractional bandwidth $b$ corresponds to $B = 2.5$ GHz and $f_c = 73$ GHz. For both cases, the minimum array gain threshold $g_t = \frac{g_m}{\sqrt{2}}$, which is 3 dB below the maximum array gain; the number of antennas in the ULA is $N = 16$; the distance between two adjacent antennas $d = \frac{\lambda_c}{2}$, i.e., half of the wavelength of carrier frequency; $\psi_m = 1$ corresponds to the maximum targeted covered AoA/AoD of the codebook being $90^\circ$.

There is no beam squint in Figure 7.2 (a). The minimum codebook size is 19. Figure 7.2 (b) shows the codebook design resulting from Algorithm 2. The minimum codebook size is 23. The minimum codebook size in this example for the $N = 16$ ULA is increased by 21.1% compared to that without considering beam squint, and this percentage increase grows as $N$ increases. For example, the percentage increase is 55.3% for a $N = 32$ ULA with the same fractional bandwidth $b$. More examples will be shown later in Figure 7.8.

From Figure 7.2 (b), we can also observe that the beams are more dense for beam focus angles with higher magnitude, because of narrower effective beamwidth for larger beam focus angles as shown in (7.39). The codebook in Figure 7.2 (b) with Algorithm 2 has optimized codebook array gain $g_d = 0.7339g_m$, which is higher than the required minimum array gain threshold $g_t = \frac{g_m}{\sqrt{2}} = 0.717g_m$. All beams in Figure 7.2 (b) are fully utilized.

7.4.2 Limitations on the Generalized Beam Edge $\psi_d$ and the Optimized Codebook Array Gain $g_d$

Figure 7.3 illustrates the generalized beam edge $\psi_d$ as a function of the codebook size $M$. As $M$ increases, $\psi_d$ decreases and approaches its lower bound $b\psi_m/2$ from Lemma 3.
Figure 7.3. Generalized beam edge $\psi_d$ as a function of the codebook size $M$ for various fractional bandwidths. $\psi_m = 1$. The dash lines show the lower bounds for corresponding fractional bandwidth $b$. The case $b = 0$ corresponds to the case without beam squint; $b = 0.0179$ corresponds to $B = 0.5$ GHz and $f_c = 28$ GHz; $b = 0.0342$ corresponds to $B = 2.5$ GHz and $f_c = 73$ GHz; $b = 0.0417$ corresponds to $B = 2.5$ GHz and $f_c = 60$ GHz; $b = 0.0714$ corresponds to $B = 2$ GHz and $f_c = 28$ GHz.

From Lemma 3, the optimized codebook array gain $g_d$ increases as the codebook size increases, but it is upper bounded. Figure 7.4 and Figure 7.5 illustrate the trend of $g_d$ as a function of codebook size $M$. Figure 7.4 focuses on the comparison of different fractional bandwidths. As $b$ increases, $\psi_d$ increases for the same $M$, reducing $g_d$. Figure 7.5 focuses on the comparison of arrays with different number of antennas $N$. As $M$ increases, $g_d$ grows larger but with slower approach to the upper bound. As $N$ increases, the beamwidth becomes smaller, and therefore, more beams are required for $g_d$ to start to increase.
Figure 7.4. Optimized codebook array gain $g_d$ as a function of the codebook size $M$ for various fractional bandwidths. The number of antennas $N = 32$, and $\psi_m = 1$. The dash lines show the lower bound for corresponding fractional bandwidth $b$. The case $b = 0$ corresponds to the case without beam squint; $b = 0.0179$ corresponds to $B = 0.5$ GHz and $f_c = 28$ GHz; $b = 0.0342$ corresponds to $B = 2.5$ GHz and $f_c = 73$ GHz; $b = 0.0417$ corresponds to $B = 2.5$ GHz and $f_c = 60$ GHz; $b = 0.0714$ corresponds to $B = 2$ GHz and $f_c = 28$ GHz.
Figure 7.5. Optimized codebook array gain $g_d$ grows as codebook size $M$ increases. The growth rate keeps decreasing. $\psi_m = 1$. Fractional bandwidth $b = 0.0342$, corresponding to $B = 2.5$ GHz and $f_c = 73$ GHz. The dash lines show the upper bounds for corresponding $N$. 
Figure 7.6. Minimum codebook size $M_{\text{min}}$ required to cover virtual angles $\psi \in [-1, 1]$ as a function of fractional bandwidth $b$. The dashed lines indicate upper bounds. The ratio of optimized codebook array gain to the maximum array gain, $r_g$, is set to be 0.707, that is, $g_d$ is 3 dB smaller than $g_m$.

### 7.4.3 Limitations on the Fractional Bandwidth $b$

Figure 7.6 illustrates examples of the minimum codebook size $M_{\text{min}}$ as a function of fractional bandwidth $b$ for $\psi_m = 1$, i.e., the maximum target AoA/AoD of the codebook is $90^\circ$. The case $b = 0$ is equivalent to not considering beam squint.

For all $N$, as $b$ increases, $M_{\text{min}}$ increases, and $b$ is upper bounded for fixed $N$, confirming Lemma 4. As $b$ approaches its unreachable upper bound $\frac{2\psi_d}{\psi_m}$, $M_{\text{min}}$ increases significantly. If $b \to \frac{2\psi_d}{\psi_m}$, $\Delta_{\text{M}} = (\psi_{M,r} - \psi_{M,I}) \to 0$ for the beam with $\psi_F = \psi_m$; therefore, $M_{\text{min}} \to \infty$. If $b \geq \frac{2\psi_d}{\psi_m}$, no codebook exists.
7.4.4 Limitations on the Number of Antennas $N$

Figure 7.7 illustrates the maximum relative approximation error, $e_{a,\max}(N)$ in (7.49) as a function of $N$. As $N$ increases, the relative approximation error drops dramatically. Even when $N = 4$, $e_{a,\max}(4)$ is as low as 0.1. The numerical result shows that our approximation used in Lemma 5 is reasonable.

Figure 7.8 shows examples of the minimum codebook size $M_{\min}$ as a function of the number of antennas $N$ for $\psi_m = 1$. As expected, $M_{\min}$ increases as $N$ increases. In the example, $r_g = 0.707$, so that $N$ is upper bounded by $\left\lfloor \frac{1.772}{\psi_m b} \right\rfloor$ as described in (7.52). If $N$ goes beyond this upper bound, no codebook exists. The maximum array gain only depends on $N$. Therefore, the maximum array gain $g_m = \sqrt{N}$ is also upper bounded.

7.4.5 Diminishing Returns of Increasing Number of Antennas $N$

For fixed codebook size $M$, it turns out that the optimized codebook array gain $g_d$ is not strictly increasing in the number of antennas $N$. As illustrated in Figure 7.9
Figure 7.8. Minimum codebook size $M_{\text{min}}$ required to cover virtual angles $\psi \in [-1, 1]$ as a function of the number of antennas $N$. The dashed lines indicate upper bounds. The ratio of optimized codebook array gain to the maximum array gain, $r_g$, is set to be 0.707, that is, $g_d$ is 3 dB smaller than $g_m$. Fractional bandwidth $b = 0.0179$ corresponds to $B = 0.5$ GHz and $f_c = 28$ GHz; $b = 0.0342$ corresponds to $B = 2.5$ GHz and $f_c = 73$ GHz; $b = 0.0417$ corresponds to $B = 2.5$ GHz and $f_c = 60$ GHz; $b = 0.0714$ corresponds to $B = 2$ GHz and $f_c = 28$ GHz.
Figure 7.9. Optimized codebook array gain $g_d$ as a function of the number of antennas $N$. The codebook size $M = 30$, and $\psi_m = 1$. Fractional bandwidth $b = 0$ corresponds to the case without beam squint; $b = 0.0179$ corresponds to $B = 0.5 \text{ GHz}$ and $f_c = 28 \text{ GHz}$; $b = 0.0342$ corresponds to $B = 2.5 \text{ GHz}$ and $f_c = 73 \text{ GHz}$; $b = 0.0417$ corresponds to $B = 2.5 \text{ GHz}$ and $f_c = 60 \text{ GHz}$; $b = 0.0714$ corresponds to $B = 2 \text{ GHz}$ and $f_c = 28 \text{ GHz}$.

$g_d$ increases with $N$ for small $N$, but then decreases beyond a certain $N$. The reason is that as $N$ increases, the main lobe becomes narrower (the width of main lobe is $\frac{4}{N}$) while the generalized beam edge $\psi_d$ stays the same, so that the beam edges $\psi_{i,l}$ and $\psi_{i,r}$ approach the edges of the main lobe.

Figure 7.9 also illustrates that $g_d$ decreases as the fractional bandwidth $b$ increases, that is, wider bandwidth further decrease the array gain.
7.5 Summary

A beamforming codebook with array gain constraint was developed to compensate for beam squint in a ULA. In the codebook design, the array gains of all subcarriers and all AoA/AoD within the desired coverage range must be larger than a given minimum array gain threshold. Analysis and numerical examples showed that beam squint increases the codebook size with the minimum array gain threshold. We also demonstrated that the fractional bandwidth is upper bounded for fixed number of antennas; the number of antennas in the phased array is correspondingly upper bounded for fixed fractional bandwidth.
CHAPTER 8

CONCLUSIONS AND FUTURE WORK

8.1 Conclusions

Beam squint is a concern for millimeter wave beamforming. In this dissertation, we developed a channel model to conveniently capture wireless multipath propagation as well as beam squint at either a uniform linear array or a uniform planar array. Our study showed that beam squint can reduce channel capacity and increase channel estimation error. Therefore, system functions such as path selection or beam selection should also take beam squint into consideration. Generally speaking, the effects of beam squint on wireless system increase with the beam focus angle away from broadside, the number of antennas in the array or the fractional bandwidth.

We also compensated for beam squint in three aspects. First, we studied carrier aggregation of multiple millimeter wave bands with beam squint. Specifically, criteria were derived to determine whether or not to aggregate two symmetric bands. Second, beamforming codebooks were designed to compensate for the capacity loss due to beam squint. Third, codebook design was decoupled between the transmitter and receiver by constraining the minimum array gain for all frequencies and angles of arrival or departure. In both codebook designs, limitations on system parameters such as fractional bandwidth or number of antennas were analyzed, and the compensation for beam squint within the codebook is at the expense of increasing codebook size.
8.2 Future Work

This dissertation provides a fundamental study of beam squint in wireless communication systems. There remain a number of problems relating to beam squint that should be explored in the future, including but not limited to the following.

8.2.1 Carrier Aggregation

In this dissertation, we only study in detail the carrier aggregation problem for two symmetric bands with a ULA. Numerical optimization is proposed for arbitrary combinations of multiple bands. Therefore, some future directions could include:

- Develop simple and effective design criterion for carrier aggregation with arbitrary combinations of multiple bands instead of using numerical optimization;
- Extend the carrier aggregation problem to UPA or other widely-used array geometries;
- Consider multiple paths within the main lobe in combination with carrier aggregation;
- Explore more general beamforming architectures, e.g., hybrid beamforming, in conjunction with carrier aggregation.

8.2.2 Beamforming Codebook Design

Our treatment of beamforming codebook design subject to a channel capacity constraint ignores beam squint at the transmitter. Codebook designs for a channel capacity constraint and for a minimum array gain constraint were both targeted for ULA. Future work should explore:

- Codebook design with channel capacity constraint considering beam squint in both the transmitter and receiver;
- Extension to UPA or other widely-used array geometries;
- Generalization to hybrid beamforming architectures.
8.2.3 Channel Estimation

We have shown that beam squint increases channel estimation errors of MLE. To tackle this issue, we could, among other directions:

- Develop new channel estimation algorithms that take beam squint into consideration;
- Extend the channel estimation algorithms to UPA.

8.2.4 Beamforming in Frequency Division Duplex

In Time Division Duplex (TDD) systems, downlink and uplink share the same channels, and their beam squint issues are similar. However, in Frequency Division Duplex (FDD) systems, downlink and uplink use different bands simultaneously, increasing the system bandwidth. Problems of beam squint in FDD systems that warrant exploration include:

- Explore beam squint in FDD systems with both ULA and UPA for both analog beamforming and hybrid beamforming;
- Develop codebook designs for FDD systems to optimize the performance of both uplink and downlink;
- Determine the band separations between downlink and uplink for which TDD is better to use than FDD.
APPENDIX A

PROOFS

A.1 Proof of Theorem 1

Put (5.24) into (5.23), and we can obtain an equation of $B_{s,c}$.

\[
\log \left( 1 + \frac{\text{SNR}_r P_1 N}{P} \right) + \log \left( 1 + \frac{\text{SNR}_r P_2 \left| g \left( \frac{\psi_{B_{s,c}}}{f_o} \right) \right|^2}{P} \right) \\
= 2 \log \left( 1 + \frac{\text{SNR}_r \left| g \left( \frac{\psi_{B_{s,c}}}{2 f_o} \right) \right|^2}{2} \right). \quad (A.1)
\]

To prove Theorem 1, we need to show that the channel capacity on the left of (A.1) (focusing on the center of band 1) becomes smaller than that on the right (focusing on the center of the two bands), if SNR$_r$ increases from current equality-achieving value in (5.23). Because in that case, larger critical band separation is needed to achieve equality in (A.1) again.

Define

\[
h_1 (\text{SNR}_r) = \log \left( 1 + \frac{\text{SNR}_r P_1 N}{P} \right) + \log \left( 1 + \frac{\text{SNR}_r P_2 \left| g \left( \frac{\psi_{B_{s,c}}}{2 f_o} \right) \right|^2}{P} \right) \\
- 2 \log \left( 1 + \frac{\text{SNR}_r \left| g \left( \frac{\psi_{B_{s,c}}}{2 f_o} \right) \right|^2}{2} \right). \quad (A.2)
\]
For simplicity, we also define

\[ g_1 = \frac{P_1 N}{P}, \quad (A.3) \]

\[ g_2 = \frac{P_2 |g(\psi_{B_{s,c} f_a})|^2}{P}, \quad (A.4) \]

\[ g_3 = \frac{|g(\psi_{B_{s,c} 2f_a})|^2}{2}. \quad (A.5) \]

Then

\[ h_1 (\text{SNR}_r) = \log \frac{(1 + \text{SNR}_r g_1) (1 + \text{SNR}_r g_2)}{(1 + \text{SNR}_r g_3)^2}. \quad (A.6) \]

Suppose with \( \text{SNR}_{r,0} \) and \( B_{s,c} \), \( h_1 (\text{SNR}_{r,0}) = 0 \). To prove Theorem[I] given \( \text{SNR}_{r,1} > \text{SNR}_{r,0} \) while \( B_{s,c} \) remains the same, we must show \( h_1 (\text{SNR}_{r,1}) < 0 \), that is, given

\[ \frac{(1 + \text{SNR}_{r,0} g_3)^2}{(1 + \text{SNR}_{r,0} g_1) (1 + \text{SNR}_{r,0} g_2)} = 1, \quad (A.7) \]

we need to prove

\[ \frac{(1 + \text{SNR}_{r,1} g_3)^2}{(1 + \text{SNR}_{r,1} g_1) (1 + \text{SNR}_{r,1} g_2)} > 1, \quad (A.8) \]

for \( \text{SNR}_{r,1} > \text{SNR}_{r,0} \).

(A.7) is equivalent to a function of \( \text{SNR}_{r,0} \), \( h_2 (\text{SNR}_{r,0}) = 0 \), defined as

\[ h_2 (\text{SNR}_{r,0}) = (1 + \text{SNR}_{r,0} g_3)^2 - (1 + \text{SNR}_{r,0} g_1) (1 + \text{SNR}_{r,0} g_2) \]
\[ = (g_3^2 - g_1 g_2) \text{SNR}_{r,0}^2 - (2g_3 - g_1 - g_2) \text{SNR}_{r,0} \]
\[ = 0. \quad (A.9) \]

To prove (A.8) for \( \text{SNR}_{r,1} > \text{SNR}_{r,0} \), we need to show \( h_2 (\text{SNR}_{r,1}) > 0 \).
From (A.9),

\[ \text{SNR}_{r,0} = \frac{2g_3 - g_1 - g_2}{g_3^2 - g_1g_2}. \]  

(A.10)

Since \( \text{SNR}_{r,0} > 0 \), we have either \( 2g_3 - g_1 - g_2 > 0 \) and \( g_3^2 - g_1g_2 > 0 \), or \( 2g_3 - g_1 - g_2 < 0 \) and \( g_3^2 - g_1g_2 < 0 \). It is easy to show that both \( 2g_3 - g_1 - g_2 > 0 \) and \( g_3^2 - g_1g_2 > 0 \).

Suppose \( 2g_3 - g_1 - g_2 < 0 \). Then \( g_3 > \frac{g_1 + g_2}{2} \), and \( g_3^2 - g_1g_2 = \frac{(g_1 + g_2)^2}{4} - g_1g_2 = \frac{(g_1 - g_2)^2}{4} \geq 0 \), contradicting to the fact that \( g_3^2 - g_1g_2 < 0 \) when \( 2g_3 - g_1 - g_2 < 0 \).

Therefore, \( 2g_3 - g_1 - g_2 > 0 \) and \( g_3^2 - g_1g_2 > 0 \).

We further have

\[ h'_2(\text{SNR}_{r,0}) = 2 \left( g_3^2 - g_1g_2 \right) \text{SNR}_{r,0} - (2g_3 - g_1 - g_2) \]

\[ = 2g_3 - g_1 - g_2 > 0, \]  

(A.11)

where the superscript \((\cdot)'\) indicates derivative. Therefore, \( h_2(\text{SNR}_{r,1}) > 0 \) for \( \text{SNR}_{r,1} > \text{SNR}_{r,0} \), which completes the proof.

A.2 Proof of Theorem 2

The derivation of \( \psi_d \) can be divided into two situations, \( M \) is odd and \( M \) is even.

A.2.1 Codebook Size \( M \) is Odd

Again, if \( M \) is odd, \( \psi_{\frac{M+1}{2},F} = 0 \). From (7.25),

\[ \psi_{\frac{M+1}{2},t} = \psi_{\frac{M+1}{2},r} = \frac{2\psi_d}{2 + b}. \]  

(A.12)

Then

\[ \psi_{M,r} = \frac{2 - b}{2 + b} \psi_{M,t} + \frac{4}{2 + b} \psi_d \]
\[
\begin{align*}
\psi_d &= \frac{b\psi_m}{2 - \frac{4}{2+b} \left( \frac{2-b}{2+b} \right)^{\frac{M-1}{2}}}, \\
\text{where } 0 < b < 1.
\end{align*}
\]
Therefore,

\[ \psi_d = \frac{b \psi_m}{2 - 2 \left( \frac{2 - b}{2 + b} \right) \frac{M}{2}}. \]  \hspace{1cm} (A.16)

where \( 0 < b < 1 \).

The proof is completed.

A.3 Proof of Theorem 3

The solution of the beam focus angle according to (7.19) are denoted as \( \psi_{i,F} \), \( i = 1, 2, \cdots, M \), where \( \psi_{k,F} \leq \psi_{k+1,F} \) for \( 1 \leq k < M \).

Suppose there is a value \( g'_d \geq g_d > 0 \) which can be achieved by beam focus angle \( \psi'_{i,F}, \ i = 1, 2, \cdots, M \), where \( \psi'_{k,F} \leq \psi'_{k+1,F} \) for \( 1 \leq k < M \). The left edge and the right edge of the \( i \)th beam with focus angle \( \psi'_{i,F} \) are denoted as \( \psi'_{i,l} \) and \( \psi'_{i,r} \), with gain \( g'_d \) at the edge. Then, similar to (7.25), (7.27), and (7.28), we have

\[ \psi'_{i,l} = 2 \left( \frac{\psi'_{i,l} + \psi'_d}{2 + b} \right) \]  \hspace{1cm} (A.17)

\[ \psi'_{i,r} = 2 \left( \frac{\psi'_{i,F} - \psi'_d}{2 - b} \right) \]  \hspace{1cm} (A.18)

\[ \psi'_{i,F} = \psi'_d + \psi'_{i,l} \left( 1 - b/2 \right). \]  \hspace{1cm} (A.19)

where

\[ \psi'_d = g^{-1} \left( g'_d \right) \leq \psi_d. \]  \hspace{1cm} (A.20)
As $\psi_{k,F}' \leq \psi_{k+1,F}'$ for $1 \leq k < M$, we have $\psi_{k,l}' \leq \psi_{k+1,l}'$ and $\psi_{k,r}' \leq \psi_{k+1,r}'$ for $1 \leq k < M$.

To cover the range $[-\psi_m, \psi_m]$ with gain no smaller than $g_d'$, we have $[-\psi_m, \psi_m] \subseteq (\cup[\psi_{i,l}', \psi_{i,r}'])$. Equivalently, we have:

$$\psi_{1,l}' \leq -\psi_m = \psi_{1,l},$$
$$\psi_{M,r}' \geq \psi_m = \psi_{M,r},$$
$$\psi_{i,r}' \geq \psi_{i+1,l}'$, \quad i = 1, 2, \ldots, M - 1. \quad (A.21)$$

If $\psi_{i,F}' \leq \psi_{i,F}$ is satisfied, we have

$$\psi_{i+1,F}' - \psi_{i+1,F}$$
$$= (\psi_d' + \psi_{i+1,l}' (1 - b/2)) - (\psi_d + \psi_{i+1,l} (1 - b/2))$$
$$= (\psi_d' - \psi_d) + (\psi_{i+1,l}' - \psi_{i+1,l}) (1 - b/2)$$
$$\leq (\psi_{i,r}' - \psi_{i,r}) (1 - b/2)$$
$$= \left( \frac{2 (\psi_{i,F}' + \psi_d')}{2 + b} - \frac{2 (\psi_{i,F} + \psi_d)}{2 + b} \right) (1 - b/2)$$
$$= \frac{2 - b}{2 + b} ((\psi_{i,F}' - \psi_{i,F}) + (\psi_d' - \psi_d))$$
$$\leq 0,$$ \quad (A.22)

where equability is achieved if and only if

$$\psi_d' = \psi_d,$$
$$g_d' = g_d,$$
$$\psi_{i,r}' = \psi_{i+1,l}',$$
$$\psi_{i,F}' = \psi_{i,F}.$$ \quad (A.23)
As $\psi'_{1,l} \leq -\psi_m = \psi_{1,t}$, we have

$$\psi'_{1,F} - \psi_{1,F} = (\psi'_d + \psi'_{1,t} (1 - b/2)) - (\psi_d + \psi_{1,t} (1 - b/2))$$
$$= (\psi'_d - \psi_d) + (\psi'_d - \psi_{1,t}) (1 - b/2)$$
$$\leq 0. \quad \text{(A.24)}$$

where equability is achieved if and only if

$$\psi'_d = \psi_d,$$
$$g'_d = g_d,$$
$$\psi'_{1,t} = -\psi_m = \psi_{1,t}. \quad \text{(A.25)}$$

Therefore, $\psi'_{M,F} \leq \psi_{M,F}$. Further, we have

$$\psi'_{M,r} - \psi_m = \psi'_{M,r} - \psi_{M,r}$$
$$= \left(\frac{2 (\psi'_{M,F} + \psi'_d)}{2 + b}\right) - \left(\frac{2 (\psi'_{M,F} + \psi_d)}{2 + b}\right)$$
$$\leq 0, \quad \text{(A.26)}$$

where equability is achieved if and only if

$$\psi'_d = \psi_d,$$
$$g'_d = g_d,$$
$$\psi'_{M,F} = \psi_m = \psi_{M,F}. \quad \text{(A.27)}$$

According to \[\text{(A.21)}, we have $\psi'_{M,r} \geq \psi_m = \psi_{M,r}$. Therefore, $\psi'_{M,r} = \psi_m = \psi_{M,r}$, and all the equality conditions should be satisfied, which completes the proof.
A.4 Proof of Theorem 4

Denote the curve illustrating the relation between $\psi_d$ and $M$ in (A.14) as an odd curve, and denote the curve illustrating the relation between $\psi_d$ and $M$ in (A.16) as an even curve as shown in Figure A.1. It can be demonstrated that $\psi_d$ in (A.14) is larger than that in (A.16) for a given $M$. Therefore, $M_o \geq M_e$, where $M_o$ is defined in (7.37), and $M_e$ is defined in (7.38). In fact, the difference between $M_o$ and $M_e$ is not larger than 1, i.e., $M_o = M_e$ or $M_o = M_e + 1$. $\psi = \psi_t$ has an intersection point with the even curve and an intersection point with the odd curve. In Figure A.1 (a), both $M_e$ and $M_o$ are larger than the $M$ value of the two intersection points; in (b), $M_e$ lies between the two intersection points.

To meet the requirement of minimum array gain, the selected minimum codebook size $M_{\text{min}}$ should be either $M_o$ or $M_e$ such that the corresponding $\psi_d \leq \psi_t$. Based on whether $m$ is even or odd, the discussions can be divided into four situations:

- If $M_o = M_e$ and $m$ is even, both $M_e$ and $M_o$ are odd. Therefore, $M_{\text{min}} = M_o$.
- If $M_o = M_e$ and $m$ is odd, both $M_e$ and $M_o$ are even. Therefore, $M_{\text{min}} = M_e$.
- If $M_o = M_e + 1$ and $m$ is even, $M_e$ is even, and $M_o$ is odd. Therefore, $M_{\text{min}} = M_e$.
- If $M_o = M_e + 1$ and $m$ is odd, $M_e$ is odd, and $M_o$ is even. Therefore, $M_{\text{min}} = M_o$.

In summary, if $M_e$ is even, $M_{\text{min}} = M_e$; if $M_e$ is odd, $M_{\text{min}} = M_o$, which completes the proof.
Figure A.1. Relation between $M_o$ and $M_e$. The odd curve is the illustration of (A.14), and the even curve is the illustration of (A.16). $m$ is an integer.
(a) $M_o = M_e = m + 1$. (b) $M_o = m + 1, M_e = m$. 


42. President’s Council of Advisors on Science and Technology (PCAST). Realizing the full potential of government-held spectrum to spur economic growth. Report to the president, White House Office of Science and Technology Policy (OSTP), July 2012.


