

Copyright  
by  
Vutha Va  
2018

The Dissertation Committee for Vutha Va  
certifies that this is the approved version of the following dissertation:

## **Beam Alignment for Millimeter Wave Vehicular Communications**

Committee:

---

Robert W. Heath, Jr., Supervisor

---

Junil Choi

---

Gustavo de Veciana

---

Alexandros G. Dimakis

---

Haris Vikalo

**Beam Alignment for Millimeter Wave Vehicular  
Communications**

**by**

**Vutha Va**

**DISSERTATION**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**DOCTOR OF PHILOSOPHY**

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2018

Dedicated to my grandmother,  
Khim Lay

## Acknowledgments

First and foremost, I extend my sincere gratitude to my supervisor, Prof. Robert Heath, Jr., for his patience, guidance, and support. Besides the technical guidance, I have also learned a great deal from him including attitude toward research, writing, and presentation, which is and will be the greatest asset to my career. I will never forget the very first writing lesson from Prof. Heath on clear structures and topic sentences when I wrote my first conference paper at UT. I always appreciate his commitment to the highest quality in all our papers and presentations. This is the core value I will always strive for. I thank Prof. Junil Choi, Dr. Xinchun Zhang, Prof. Namyoon Lee, and Prof. Ahmed Alkhateeb for their guidance, both personal and technical, and their encouragement. I appreciate Prof. Choi's commitment to fast and clear feedback in all our time collaborating together. I am grateful to my fellow students at WSIL and WNCG for their friendship, kindness, and willingness to help despite their busy time. I thank my committee members, Prof. Gustavo de Veciana, Prof. Alexandros Dimakis, and Prof. Haris Vikalo for reviewing my research and serving my committee. I extend my gratitude to my collaborators at Toyota InfoTechnology Center, USA, Inc., Dr. Takayuki Shimizu and Dr. Gaurav Bansal. It has been my great pleasure working with both of them through our collaborative project and internship. I also would like to thank my master's thesis supervisors, Prof. Kei Sakaguchi and Prof. Kiyomichi Araki, whose support had led me to continue to the Ph.D. program

at UT. Lastly, I thank my parents, Huy and Mom, and my aunt Dany for always supporting me whichever turn in life I have chosen.

# **Beam Alignment for Millimeter Wave Vehicular Communications**

Publication No. \_\_\_\_\_

Vutha Va, Ph.D.

The University of Texas at Austin, 2018

Supervisor: Robert W. Heath, Jr.

Millimeter wave (mmWave) has the potential to provide vehicles with high data rate communications that will enable a whole new range of applications. Its use, however, is not straightforward due to its challenging propagation characteristics. One approach to overcome the propagation challenge is the use of directional beams, but it requires a proper alignment and presents a challenging engineering problem, especially under the high vehicular mobility.

In this dissertation, fast and efficient beam alignment solutions suitable for vehicular applications are developed. To better quantify the problem, first the impact of directional beams on the temporal variation of the channels is investigated theoretically. The proposed model includes both the Doppler effect and the pointing error due to mobility. The channel coherence time is derived, and a new concept called the beam coherence time is proposed for capturing the overhead of mmWave beam alignment.

Next, an efficient learning-based beam alignment framework is proposed. The core of this framework is the beam pair selection methods that use side information (position in this case) and past beam measurements to identify promising beam directions and eliminate unnecessary beam training. Three offline learning methods for beam pair selection are proposed: two statistics-based and one machine learning-based methods. The two statistical learning methods consist of a heuristic and an optimal selection that minimizes the misalignment probability. The third one uses a learning-to-rank approach from the recommender system literature. The proposed approach shows an order of magnitude lower overhead than existing standard (IEEE 802.11ad) enabling it to support large arrays at high speed.

Finally, an online version of the optimal statistical learning method is developed. The solution is based on the upper confidence bound algorithm with a newly introduced risk-aware feature that helps avoid severe misalignment during the learning. Along with the online beam pair selection, an online beam pair refinement is also proposed for learning to adapt the codebook to the environment to further maximize the beamforming gain. The combined solution shows a fast learning behavior that can quickly achieve positive gain over the exhaustive search on the original (and unrefined) codebook. The results show that side information can help reduce mmWave link configuration overhead.



# Table of Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>List of Tables</b>	<b>xiii</b>
<b>List of Figures</b>	<b>xiv</b>
<b>Chapter 1. Introduction</b>	<b>1</b>
1.1 High Data Rate Demand of Future V2X Applications . . . . .	2
1.2 MmWave to Support High Data Rate V2X . . . . .	5
1.3 Challenges of mmWave Propagation . . . . .	6
1.3.1 The Friis Free Space Path Loss Formula . . . . .	7
1.3.2 Blockage . . . . .	10
1.3.3 Doppler Spread . . . . .	11
1.4 MmWave Transceiver Architectures . . . . .	12
1.5 MmWave V2X: Challenges and Opportunities . . . . .	15
1.6 Thesis Statement and Summary of Contributions . . . . .	16
1.7 Organization . . . . .	19
1.8 Abbreviations . . . . .	19
<b>Chapter 2. Impact of Beamwidth in Vehicular Channels</b>	<b>21</b>
2.1 Motivation and Prior Work . . . . .	22
2.2 Contributions . . . . .	25
2.3 Models and Assumptions . . . . .	26
2.3.1 Channel Model . . . . .	26
2.3.2 Pointing Error Due to Receiver Motion . . . . .	29
2.3.3 Channel Spatial Lobe Model . . . . .	31
2.3.4 Summary of Common Parameters . . . . .	33

2.4	Channel Temporal Correlation Function . . . . .	34
2.4.1	NLOS Channel Correlation Function . . . . .	35
2.4.2	LOS Channel Correlation Function . . . . .	37
2.4.3	Numerical Verification of (2.20) and Effect of $K$ Factor . . . . .	39
2.5	Channel Coherence Time . . . . .	41
2.5.1	NLOS Channel . . . . .	42
2.5.1.1	When $ \mu_r $ is small . . . . .	43
2.5.1.2	When $ \mu_r $ is not small . . . . .	44
2.5.2	LOS Channel . . . . .	46
2.5.3	Numerical Results . . . . .	47
2.6	Beam Coherence Time . . . . .	50
2.6.1	LOS Case . . . . .	50
2.6.2	NLOS Case . . . . .	51
2.6.3	Numerical Results . . . . .	53
2.7	Implications on Beam Realignment Duration . . . . .	55
2.7.1	Lower Bound on Mutual Information . . . . .	55
2.7.2	How Often Should the Beams Be Realigned? . . . . .	62
2.8	Conclusion . . . . .	68

### **Chapter 3. Inverse Multipath Fingerprinting for V2I MmWave Beam Alignment 69**

3.1	Motivation and Prior Work . . . . .	70
3.2	Contributions . . . . .	73
3.3	System Model . . . . .	76
3.3.1	Channel Model . . . . .	76
3.3.2	Received Signal Model . . . . .	80
3.4	Beam Alignment using Subset of Beam Pairs . . . . .	83
3.4.1	Multipath Fingerprint Database . . . . .	84
3.4.2	Proposed Beam Alignment . . . . .	88
3.5	Quantifying Beam Alignment Accuracy . . . . .	90
3.6	Beam Pair Subset Selection using Statistical Learning . . . . .	91
3.6.1	Selection by Ranking Average Received Powers . . . . .	92
3.6.2	Selection by Minimizing the Misalignment Probability . . . . .	93

3.7	Numerical Evaluations of Statistical Beam Pair Subset Selection	97
3.7.1	Performance Comparison: AvgPow Versus MinMisProb	98
3.7.2	Required Training Sample Size . . . . .	100
3.7.3	Effect of Location Bin Size . . . . .	101
3.7.4	Effect of Measurement Noise . . . . .	103
3.7.5	Effect of Traffic Mismatch during Database Collection and Exploitation . . . . .	107
3.7.6	Comparison with Existing Techniques . . . . .	109
3.8	Learning-to-Rank Beam Pair Subset Selection . . . . .	116
3.8.1	Scalability Issues of Context Binning . . . . .	117
3.8.2	LtR Beam Pair Selection Method . . . . .	118
3.8.3	Numerical Evaluation and Comparison with MinMisProb	122
3.9	Conclusion . . . . .	125
<b>Chapter 4.</b>	<b>Online Beam Pair Selection and Refinement</b>	<b>126</b>
4.1	Motivation and Prior Work . . . . .	127
4.2	Contributions . . . . .	131
4.3	System Model and Some Background . . . . .	132
4.4	Proposed Two-Layer Online Learning . . . . .	135
4.5	Online Beam Pair Selection . . . . .	137
4.5.1	Problem Statement . . . . .	137
4.5.2	Greedy UCB Algorithm . . . . .	138
4.5.3	Risk-Aware Greedy UCB Algorithm . . . . .	141
4.5.4	Regret Analysis . . . . .	145
4.6	Online Beam Pair Refinement . . . . .	148
4.6.1	Problem Statement . . . . .	148
4.6.2	Modified HOO for Beam Pair Refinement . . . . .	149
4.7	Numerical Results . . . . .	155
4.7.1	Online Beam Pair Selection . . . . .	156
4.7.2	Online Beam Pair Refinement . . . . .	161
4.7.3	Integrated Online Learning Solution . . . . .	164
4.8	Conclusion . . . . .	166
4.9	Appendix . . . . .	167

4.A	Proof of Theorem 4.1 . . . . .	167
4.B	Proof of Theorem 4.2 . . . . .	169
4.C	Proof of Lemma 4.1 . . . . .	171
<b>Chapter 5.</b>	<b>Conclusion</b>	<b>174</b>
5.1	Summary . . . . .	175
5.2	Future Research Directions . . . . .	176
5.2.1	Experimental Verification of Impact of Beamwidth on Channel Temporal Variation . . . . .	176
5.2.2	Learning for Multi-User Beam Alignment . . . . .	177
5.2.3	Learning for Beam Tracking . . . . .	178
5.2.4	Online Contextual Bandit for Beam Alignment . . . . .	179
5.2.5	Extension to Hybrid Architectures . . . . .	179
5.2.6	Personalized MmWave Link Configuration . . . . .	180
<b>Bibliography</b>		<b>182</b>
<b>Vita</b>		<b>200</b>

## List of Tables

1.1	A summary of existing V2X standardized solutions. (LTE-V2X here is based on release 14 [2].) . . . . .	5
1.2	11 candidate bands for 5G identified in WRC-15. . . . .	7
2.1	Common parameters used in numerical examples. . . . .	33
3.1	An example of Type A fingerprints. For each contributing vehicles, the measurements of the top- $C$ beam pairs are stored. In each cell, the top number is the beam pair index and the bottom one is the received power. . . . .	85
3.2	An example of Type B fingerprints at a location bin. The average received power for each beam pair is recorded. . . . .	86
3.3	Common simulation parameters . . . . .	98
3.4	An example of the database used for LtR beam pair selection. The $\ell$ -th row corresponds to measurement data contributed by a car at position $\mathbf{x}_\ell$ . This is the same as Type A fingerprint database in Table 3.1 with the position tagged to each record. . . . .	117

# List of Figures

1.1	<p>Potential applications of high data rate mmWave vehicular links. With V2I links, the infrastructure can gather sensing data from the vehicles which can be used for various purposes such as traffic control, city planning, and crowd-sourced map-building [98]. If equipped with sensors such as camera or LIDAR, the infrastructure can provide a live bird-eye view of the current street or view of the crossing street that could be used for intersection automation. V2I is also crucial for infotainment applications such as broadband Internet access and cloud-based services. Vehicles can also share raw sensing data among themselves which can be used for improving sensing accuracy as well as enhancing the sensing range to cover blind spots and hidden objects [62]. Overall, sharing sensor data either with other vehicles or with the infrastructure can enhance situational awareness that could benefit advanced autonomous driving applications. High data rate V2I links provide a gateway to the backbone network that can support infotainment and cloud-based services. . . . .</p>	4
1.2	<p>Practical mmWave transmitter architectures. The receiver architectures can be constructed similarly. The Tx block is responsible for baseband processing and controlling the phase shifter and RF switches to produce the desired beamforming. Equipped with multiple RF chains (represented by DAC), the hybrid architecture is more flexible and can support spatial multiplexing. This flexibility comes at the cost of a more complex link configuration than that of the analog architecture. . . . .</p>	13
2.1	<p>Receiver displacement and change in pointing angle for the NLOS case. When the receiver moves from A to B, if the beam is not adaptive, then the set of scatterers seen at B will be different from those seen at A. This effect can be captured by the change in pointing direction <math>\Delta_\mu</math> (we called this pointing error) from the original pointing direction <math>\mu_r</math>. We ignore the change in path loss due to the displacement <math>\Delta_d</math> because for a short time duration <math>\tau</math> typically <math>\Delta_d(\tau) \ll D_r</math>. . . . .</p>	30

2.2	An illustration of spatial lobes. This figure illustrates the pattern of the incoming power arriving at the receiver. The incoming power has strong spatial dependence, and it can be observed here that there are four main directions, which can be thought of as four clusters of scatterers. At each of these directions, there is spread forming a lobe, which is termed as a spatial lobe. In this example, lobe #1 has the strongest power. . . . .	32
2.3	Comparison of the channel temporal correlations computed from the exact expression (2.17), the approximate expression (2.20), and those computed from simulation. The beamwidth is set to $\simeq 8^\circ$ ( $k_r = 50$ ), the number of paths $N = 10$ and $N = 10000$ are used, and other parameters are defined in Table 2.1. . . .	39
2.4	The effect of the Rician $K$ factor on the channel correlation function. The plots show the absolute values of the correlation coefficients for $K$ from 0 to 2 with a step of 0.2. . . . .	40
2.5	Channel coherence time for small $\mu_r$ . “Exact” refers to the coherence time obtained numerically from the exact correlation function (2.17). “Approximation” and “No pointing error” refer to (2.36) and (2.37). The result is quite sensitive to $\mu_r$ , and the approximation does not work well for small $\theta$ but still can capture the effect of the receiver motion. . . . .	47
2.6	Channel coherence time for difference $\mu_r$ . “Exact” refers to the numerical solution to the exact expression (2.17), and “Approximation” refers to (2.41). We see that the range of valid approximation increases with $\mu_r$ . . . . .	48
2.7	The beam coherence time for both the LOS and NLOS cases. The transmitter-receiver distance is set to $D = 50$ m, and the parameters from Table 2.1 are used. For the LOS case, $T_B$ increases linearly with the beamwidth $\theta$ . For the NLOS case, the increase is slower. This is because of the effect of the spatial lobe width $\beta$ on $T_B$ in the NLOS case. For both LOS and NLOS cases, $T_B$ increases fast with $\theta$ for $\mu_r \rightarrow 0^\circ$ , and slowly for $\mu_r \rightarrow 90^\circ$ . . . . .	54

2.8	Effects of different parameters on $I_{\text{low}}$ . In all cases, $\text{SNR}_s = \text{SNR}_v = 0$ dB and $D_{r,\lambda} = 1000$ are used. Except indicated otherwise in the legends, $B_c = 10$ MHz, $v = 30$ m/s, $\mu_r = 90^\circ$ , and $\theta = 10^\circ$ are used. Fig. 2.8(a) shows $I_{\text{low}}$ against pilot spacing for different $B_c$ . The optimal $\nu$ increases as $B_c$ gets larger. Fig. 2.8(b) shows $I_{\text{low}}$ for different $\mu_r$ . The optimal $\nu$ decreases as $\mu_r$ approaches $90^\circ$ . This is because the fading changes faster for $\mu_r$ near $90^\circ$ [33]. Fig. 2.8(c) shows $I_{\text{low}}$ for different speeds. As the speed increases, the time-variation increases and the pilot spacing should be decreased. Fig. 2.8(d) shows the optimal $\nu$ and the corresponding $I_{\text{low}}$ against the beamwidth $\theta$ . Both the optimal $\nu$ and $I_{\text{low}}$ decrease for a too small or too large $\theta$ . This is because $\theta$ that is too small suffers from the pointing error, and $\theta$ that is too large suffers from the Doppler spread which results in fast time variation (small $T_c$ ). . . . .	60
2.9	Comparison of the spectral efficiencies in (2.82) and (2.83) for the short- and long-term beam realignment when the beam sweeping follows the 802.15.3c method. Fig. (a) and (b) show the case when the path loss ratio of the two spatial lobes $\Gamma$ is 3 dB and 10 dB, respectively. In both cases, the long-term realignment performs better and the gap is more pronounced when $\Gamma = 10$ dB. The gap increases for larger $\Gamma$ because the sweeping is less likely to make mistake when $\Gamma$ is large so that the large overhead of the short-term realignment penalizes rather than improves the performance. . . . .	67
3.1	Ray-tracing simulation environment. The scenario is an urban street with two lanes. There are two types of vehicles: cars and trucks. The CV is a car on the left lane and its LOS path can be blocked by a truck. The antenna heights are 1.5 m at the CV (roof-mounted) and 7 m at the RSU. . . . .	76
3.2	Beam patterns in our codebook for an $8 \times 8$ array. The array is assumed to face upward in the $+z$ direction. The codebook covers the directions in the $+z$ half-space (i.e., assuming no radiation in the backplane). . . . .	82
3.3	An illustration of position-aided beam alignment in the uplink. It consists of two phases. Phase 1 is for the training request where the CV position is sent to the RSU. The RSU uses the position and its learned database to determine a list of promising beam pairs $\mathcal{S}$ . In Phase 2, the beam pairs in the list are trained, and a feedback indicated the best beam index is sent at the end. . . . .	88



3.4	Power loss probability versus the number of beam pairs trained. The 3 dB power loss probability plot for MinMisProb ends before reaching $N_b = 50$ because there is no such instance of power loss computed from the cross validation. MinMisProb outperforms AvgPow in both the misalignment and 3 dB power loss probability. . . . .	99
3.5	3 dB power loss probability of AvgPow as a function of the training sample size $N$ . Fig. 3.5(a) shows the average 3 dB power loss probability for different $N$ . Fig. 3.5(b) shows the learning curve in terms of the 3 dB power loss probability when the number of beam pairs measured is set to $N_b = 30$ . We can see from the plots that the improvement diminishes as the training sample size $N$ increases beyond about 100. . . . .	101
3.6	Average 3 dB power loss probability when using different location bin sizes. All bin sizes performs similarly when using $16 \times 16$ arrays. When using the larger 32 arrays, smaller bin sizes improve the performance. . . . .	102
3.7	CDF of the link SNR of the generated channels. The link SNR is defined as the SNR at the receiver when the transmit power is 0 dBm and both the transmitter and receiver use isotropic antennas. The average link SNR is -16.0 dB. . . . .	104
3.8	Power loss probability as a function of $N_b$ in the presence of noise. The noise impacts the misalignment probability much more than the 3 dB power loss probability. . . . .	105
3.9	Average rate of the proposed beam alignment compared to the perfect alignment case. MinMisProb consistently achieves a higher average rate than AvgPow for the same $N_b$ . The rate loss becomes negligible at $N_b$ at around 20 and 30 for MinMisProb and AvgPow, respectively. . . . .	106
3.10	Effect of the mismatch in traffic density during database collection and exploitation. Database collected in a light traffic does not work well when used in a dense traffic because the database cannot capture all the paths in the richer scattering environment of the dense traffic. On the contrary, database collected in a dense traffic works well regardless of the traffic condition when it is exploited. It only has a slightly degraded efficiency when used in a low traffic condition. . . . .	108
3.11	Required amount of beam training of the proposed method. The overhead increases roughly linearly with the number of antenna elements of the array. . . . .	111

3.12	An illustration of the beam coherence time concept. For the beam to stay aligned, the reflection point must be within the beam projection. The beam coherence time is the average time that the reflection point is within the beam projection. . . . .	112
3.13	Overhead comparison between the proposed method and that of IEEE 802.11ad. The beam training time of the proposed method only takes up to a few percent of the beam coherence time $T_B$ , while that of IEEE 802.11ad can exceed $T_B$ when using a large array leaving no time for data communications. . . . .	114
3.14	Average rate comparison with the beam alignment method using position only. The method using only the position does not need beam training but works only when the LOS path is available. It performs well in a low traffic density where the LOS path is often available but performs poorly in a dense traffic where the LOS path is often blocked. Taking advantage of the fingerprint database, the proposed method works well in both cases. . . . .	115
3.15	Top- $N_b$ recommendation process at the car's position $\mathbf{x}_\ell$ . First, the position $\mathbf{x}_\ell$ is used to produce a set of queries for the pointing directions of interest. Then, using the learned scoring function $\hat{z}(\cdot)$ , the system predicts the scores of these queries points $\mathbf{q}_{i\ell}$ . The scores are then sorted to produce the beam pair indices. $\pi(n)$ denotes the beam index with the $n$ -th highest predicted score. The top- $N_b$ indices are the output of the beam recommender system. . . . .	119
3.16	Misalignment probability with different transform function $S(\cdot)$ . A common choice used in web-recommender systems is the exponential function, but it does not work as well as the square and cubic transform. One possible explanation is the nature of the raw scores. In typical web applications, the raw scores are integer ratings (e.g., between 1 and 5), while in our case the scores are real numbers in $[0,1]$ . . . . .	123
3.17	Misalignment probability with different kernel functions. We perform feature scaling when using the linear and RBF kernels by dividing each component of the feature vector by the maximum value possible for that feature. The linear kernel does not work at all. The RBF kernel works relatively well but is outperformed by the proposed modified RBF kernel for a large $N_b$ . Overall, the modified RBF kernel provides the best performance.	124

4.1	A flowchart of the proposed online learning algorithm. The algorithm starts with a training request detection loop. When it detects a request, the algorithm decodes the user's position and input to the beam selection procedure, which then reads the learning parameters corresponding to the position and determines a subset of promising beam pairs. If beam refinement is enabled, the refinement parameters of those selected pairs are selected. The beam subset is then sent to the user and the subset of beam pairs are trained. The beam measurements are used to update the learning parameters and the algorithm returns to the training request detection loop. . . . .	136
4.2	Average 3 dB power loss probability using the proposed risk-aware greedy UCB algorithm with different training budgets $B_{\text{tr}}$ and risk thresholds $\Gamma_{\text{risk}}$ . For both $B_{\text{tr}} = 10$ and 30, the plots show similar learning behavior. A smaller training budget $B_{\text{tr}} = 10$ provides less accuracy beam alignment. The plots using different $\Gamma_{\text{risk}}$ show that the performance is not sensitive to $\Gamma_{\text{risk}}$ as long as it is not too large. . . . .	157
4.3	Performance comparison of greedy UCB with and without risk-awareness. The performance is an order of magnitude worse without risk-awareness. This is because the risk-aware greedy UCB uses the risk estimates to control the number of high-risk beam pairs selected in the subset $\mathcal{S}$ reducing the probability of severe misalignment. . . . .	158
4.4	A comparison of the accuracy of the selection set produced by the average of the proposed reward signal ( $\hat{P}_{\text{opt}}(i)$ ) and the more intuitive choice of average channel strength. The performance when using $\hat{P}_{\text{opt}}(i)$ is consistently better for all training budgets. The comparison when using the proposed practical reward signal (4.13) as opposed to the ideal reward signal (4.12) shows negligible performance loss. . . . .	160
4.5	A comparison of HOO and MAB with different $\alpha_{\text{norm}}$ and $\ell_{\text{max}}$ . Fig. (a) compares the performance when $\ell_{\text{max}} = 3$ . MAB does not use the hierarchical structure of the search tree as HOO and suffers a larger exploration penalty. The penalty is even more severe as $\alpha_{\text{norm}}$ increases. The results show that the forced exploration is not needed and $\alpha_{\text{norm}} = 0$ should be used. Fig. (b) compares the performance of HOO when using different $\ell_{\text{max}}$ . There is negligible gain for setting $\ell_{\text{max}}$ beyond 3. We also see that HOO does not have extra degradation due to exploration when we increase $\ell_{\text{max}}$ . . . . .	162

4.6	A comparison of HOO with and without the smoothness coefficient $\nu(\ell)$ (for computing the U-values). The smoothness coefficient shows negligible effect. This is likely because the refinement problem searches locally within the 3 dB beamwidth. Since this is in the vicinity of the optimal point, the smoothness bound does not help in eliminating suboptimal nodes resulting in no performance gain. . . . .	163
4.7	A comparison of average gain of the integrated solution with the three options for when to start the beam refinement. The plots show no negative impact of the beam pair refinement on the online learning for beam pair selection. It is best to start the refinement simultaneously with the online beam pair selection.	165

# Chapter 1

## Introduction

The automotive industry is at an inflection point between old and new technologies. More sensors are being incorporated into vehicles in an effort to realize safer and more efficient traffic. Communication technologies are being integrated into vehicles for safety applications such as blind spot warnings, do not pass warnings, and forward collision warnings, as well as non-safety related applications such as improving traffic efficiency, toll collections, and infotainment [1, 3, 60, 96, 100]. Although prior work on autonomous driving at present envisions their independent operation, there are many benefits to sharing rich sensor data such as LIDAR or visual camera images with other vehicles and/or with the infrastructure. Existing solutions such as the Dedicated Short-Range Communication (DSRC) or fourth generation (4G) cellular standards cannot support the data rate demands for such rich sensor data sharing. Millimeter wave (mmWave) communications could be a key solution to enable such applications that could help push the autonomous driving capability beyond what is possible with only onboard sensors.

This chapter motivates the research problem addressed in this dissertation and gives a summary of our contributions. Section 1.1 motivates the

needs for high data rate links for future advanced vehicle-to-everything (V2X) applications. In Section 1.2, we point out mmWave as a candidate solution. Section 1.3 provides some background on challenges of mmWave propagation, and Section 1.4 gives a brief summary of mmWave transceiver architectures. In Section 1.5, we highlight some challenges and opportunities for developing mmWave V2X solutions. Section 1.6 provides the thesis statement and summarizes our contributions. The chapter concludes with a summary of the structure of this dissertation and a list of abbreviations.

## **1.1 High Data Rate Demand of Future V2X Applications**

Gigabit-per-second vehicular link connections open up a whole new range of applications from safety to infotainment [28, 107]. Fig. 1.1 illustrates some of these applications. Vehicular automation relies heavily on environmental sensing, mainly via perception sensors such as camera, radar, and LIDAR that are limited to line-of-sight (LOS) sensing [52]. Sharing these sensor data among neighboring vehicles via vehicle-to-vehicle (V2V) links or with the infrastructure via vehicle-to-infrastructure (V2I) links can enhance the sensing range as well as provide redundancy in case of sensor failures, which could improve safety for driving in challenging environments such as dense urban traffics with limited sensing range. High data rate link connections can support high precision map and software updates on the fly. Up-to-date maps with current traffic conditions could be used to pick travel routes to improve

traffic efficiency. A wide range of infotainment applications are also possible such as video streaming, online gaming, and video conferencing that provides a virtual office environment on the go. These applications not only improve vehicular automation capability but also enhance passenger experience in both entertainment and productivity.

We now provides some concrete examples of the data rate demands of some of these applications. The data rate depends on the transmission frequency which is assumed to be 10 Hz (a typical value used for DSRC basic safety messages [60]). We have mentioned sharing rich sensor data such as those coming from LIDAR or camera. The data size depends on the resolution of the image and the level of compression. For example, for low-resolution LIDAR as used in [62], the data rate is relatively low at around 60 kbps;  $(180 \text{ beams} \times 32 \text{ bits} + 32 \times 8(\text{overhead})) \times 10 \text{ Hz} = 60.16 \text{ kbps}$ . For high-resolution LIDAR such as Velodyne HDL-64E [110], the required data rate is  $2083 \text{ beams} \times 24 \text{ bits} \times 64 \text{ vertical elements} \times 1.04 \text{ overhead} \times 10 \text{ Hz} = 33.27 \text{ Mbps}$ . For camera images, data rates range from around 10 Mbps for low resolution compressed images [62] up to around 0.5 Gbps for high-resolution uncompressed images (e.g., Prosilica GT with  $2048 \times 2048$  resolution [12]). Sharing a few of these perception sensors likely requires a data rate on the order of tens to hundreds of Mbps. Now, these numbers are for supporting a single link. In realistic settings, a vehicle typically will maintain multiple links to its neighbors. Therefore, combining these on top of infotainment likely requires Gbps data rate.

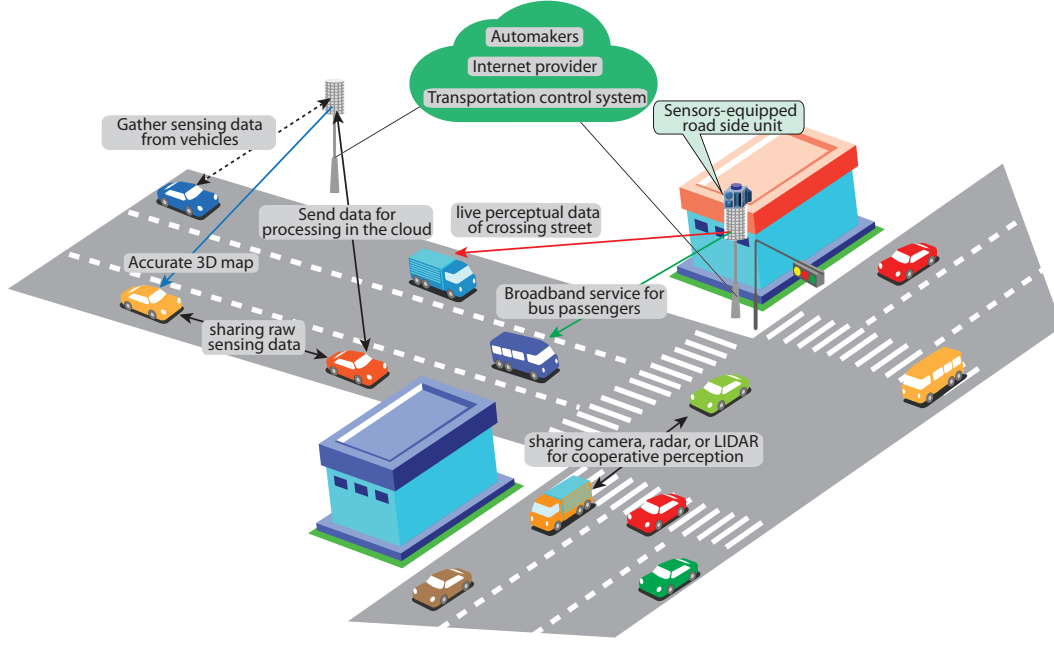


Figure 1.1: Potential applications of high data rate mmWave vehicular links. With V2I links, the infrastructure can gather sensing data from the vehicles which can be used for various purposes such as traffic control, city planning, and crowd-sourced map-building [98]. If equipped with sensors such as camera or LIDAR, the infrastructure can provide a live bird-eye view of the current street or view of the crossing street that could be used for intersection automation. V2I is also crucial for infotainment applications such as broadband Internet access and cloud-based services. Vehicles can also share raw sensing data among themselves which can be used for improving sensing accuracy as well as enhancing the sensing range to cover blind spots and hidden objects [62]. Overall, sharing sensor data either with other vehicles or with the infrastructure can enhance situational awareness that could benefit advanced autonomous driving applications. High data rate V2I links provide a gateway to the backbone network that can support infotainment and cloud-based services.



Table 1.1: A summary of existing V2X standardized solutions. (LTE-V2X here is based on release 14 [2].)

Features	DSRC	D2D LTE-V2X	Cellular LTE-V2X
Bandwidth	10 MHz	Up to 20 MHz	Up to 20 MHz
Frequency band	5.9 GHz	5.9 GHz	450 MHz-3.8 GHz
Bit rate	3-27 Mbps	Up to 44 Mbps	Up to 75 Mbps
Range	~100s m	~100s m	Up to a few km
Coverage	Ubiquitous	Ubiquitous	Inside LTE coverage
Mobility support	High speed	High speed	High speed
Comm. fee	Free	unknown	unknown

Existing solutions including DSRC and the 4G cellular standards, however, are unlikely able to support the data rate demands of these applications [14,56,60]. Table 1.1 summarizes important features of existing solutions. DSRC’s physical (PHY) layer as defined in IEEE 802.11p [49] can support a maximum data rate of 27 Mbps, but field tests suggest that practical data rates are only around 6 Mbps due to interference between neighboring vehicles [56]. The new 4G Long Term Evolution (LTE) V2X standard targets higher data rates than that of DSRC [2], but it will be challenging if not impossible to meet the data rate demands with the limited bandwidth available at sub-6 GHz bands.

## 1.2 MmWave to Support High Data Rate V2X

The huge amount of underutilized spectrum at mmWave bands makes it an attractive solution to support the high data rate demand for future V2X applications. MmWave spectrum refers to radio frequencies in the range

30-300 GHz (corresponding wavelength between 10 and 1 mm), although in practice the 20 GHz bands are sometimes also referred to as mmWave. Notable mmWave bands include the 60 GHz bands, potential 5G bands (28 GHz and 39 GHz), and automotive radar bands. The FCC has allocated a contiguous 7 GHz frequency band between 57-64 GHz since 2001 for short-range communications in an unlicensed basis [7, 119]. One remarkable feature of the 60 GHz band is its availability in most parts of the world [7, 86], although different regions have different amounts of spectrum. Also, in Europe, the 63-64 GHz has been allocated exclusively for Intelligent Transportation System (ITS) applications [97]. While 28 GHz and 39 GHz seem to be the most likely choices for 5G deployment in the US, more spectrum could be allocated for worldwide harmonization of frequency bands. In particular, at the World Radiocommunication Conference 2015 (WRC-15), 11 bands (see Table 1.2) have been identified as candidate bands for 5G and the allocation decision will be made in the WRC-19 [83]. The 77 GHz (76-77 GHz) band for long-range radars in the US is another candidate for mmWave V2X. These examples here already include tens of GHz of bandwidth. It is, however, not straightforward to unlock these bands due to the propagation challenge of mmWave, which is the topic of the next section.

### **1.3 Challenges of mmWave Propagation**

The main challenges of mmWave include the high path loss under isotropic communication (or equivalently the shrinking antenna aperture as

Table 1.2: 11 candidate bands for 5G identified in WRC-15.

Band	Bandwidth
24.25 - 27.50 GHz	3,250 MHz
31.80 - 33.40 GHz	1,600 MHz
37.00 - 40.50 GHz	3,500 MHz
40.50 - 42.50 GHz	2,000 MHz
42.50 - 43.50 GHz	1,000 MHz
45.50 - 47.00 GHz	1,500 MHz
47.00 - 47.20 GHz	200 MHz
47.20 - 50.20 GHz	3,000 MHz
50.40 - 52.60 GHz	2,200 MHz
66.00 - 76.00 GHz	10,000 MHz
81.00 - 86.00 GHz	5,000 MHz

the frequency increases) and the susceptibility to blockage [86]. Under the high mobility of V2X applications, there is also concern regarding the Doppler effect which increases with the frequency. We first describe the Friis formula for free space path loss to show the dependence on the carrier frequency and to clarify the role of antenna gain in mmWave systems. Then, we describe the blockage at mmWave which becomes more severe than at lower frequencies due to the poor diffraction and high penetration loss. Finally, we review the Doppler effect and point out that narrower beams can reduce the Doppler spread, which is a topic we will investigate in detail in Chapter 2.

### 1.3.1 The Friis Free Space Path Loss Formula

Consider a pair of transmitter and receiver antennas separated by a distance  $r$ . We want to compute the power received at the receive antenna. Assuming an isotropic transmit antenna transmitting with power  $P_t$  (thus the

power is radiated uniformly in all directions), the power density  $p$  observed at the receive antenna's location is

$$p = \frac{P_t}{4\pi r^2}. \quad (1.1)$$

Assuming that the receive antenna is in the far-field and the transmit antenna has a gain of  $G_t$  in the direction of the receiver, then the power density becomes

$$p = \frac{P_t G_t}{4\pi r^2}. \quad (1.2)$$

The power absorbed by the receive antenna can be expressed using the effective aperture concept, from which the received power  $P_r = pA_{\text{eff}}$  with  $A_{\text{eff}}$  denoting the effective aperture of the receive antenna. The effective aperture is given by

$$A_{\text{eff}} = \frac{\lambda^2}{4\pi} G, \quad (1.3)$$

where  $\lambda$  is the carrier wavelength and  $G$  is the antenna gain. The derivation of  $A_{\text{eff}}$  is a bit involved, and several approaches exist. See [30] for a detailed derivation based on thermodynamics and reciprocity principle of electromagnetic waves. Putting all these together and replacing  $G$  by  $G_r$ , the receive antenna gain, we obtain the Friis formula:

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi r)^2}. \quad (1.4)$$

Notice in (1.4) that the received power depends on the carrier wavelength  $\lambda$ . As the operating frequency is shifted to higher mmWave frequencies,  $\lambda$  becomes smaller and thus path loss will increase *provided that  $G_t$  and  $G_r$  are fixed*. For example, if  $P_t$ ,  $G_t$ , and  $G_r$  are fixed and the carrier frequency is

shifted from 2 GHz to 60 GHz, then the increase in path loss can be computed using the formula to yield  $10 \log_{10} (P_r(2)/P_r(60)) = 29.5$  dB.

We will now provide some comments on the role of directional antennas in mmWave communications. In the previous example, we saw a path loss increase of almost 30 dB when moving from 2 GHz to 60 GHz *if  $P_t$ ,  $G_t$ , and  $G_r$  are fixed*. In general, antenna gains can be closely linked with its physical size normalized in wavelength. A simple example is the uniform linear array with isotropic elements, where the ideal antenna gain can be approximated by  $G \simeq 2(L/\lambda)$  for  $L \gg d$  with  $L$  the array length and  $d$  the element spacing [18, Eq. (6-44a)]. If we keep the array length fixed, then the path loss ratio at 2 GHz and 60 GHz carrier frequency can be computed as

$$\frac{P_r(2)}{P_r(60)} = \frac{(2L/\lambda_2)^2 \lambda_2^2}{(2L/\lambda_{60})^2 \lambda_{60}^2} = 1, \quad (1.5)$$

where  $\lambda_2$  and  $\lambda_{60}$  denote the wavelength at 2 GHz and 60 GHz, respectively. Thus, we can see that if the array length is kept unchanged, then the antenna gain increases as the frequency increases and the path loss stays unchanged when moving from 2 GHz to 60 GHz. For large planar arrays, it can be shown that the gain is  $G \simeq 4\pi L_x L_y / \lambda^2$ , where  $L_x$  and  $L_y$  are the array length in the  $x$  and  $y$  direction [18, Eq. (6-103)]. Doing the same calculation as above, one can show that keeping the same array physical size at both the transmitter and the receiver, the overall path loss actually *decreases* when moving to a higher frequency. The underlying reason for this effect is the fact that with the same antenna physical size, the antenna can focus energy more when operating at higher frequencies. This is called array gain.

There is one caveat to this argument. One should note that with very directional antennas, the maximum gain as used in this example can be achieved only when the transmit and receive beams are properly aligned. Although in principle one can increase the antenna gain to compensate for the increased path loss, care must be taken to account for the overhead of aligning the beams, which increases as the directionality increases, to the communication performance.

### 1.3.2 Blockage

Two main distinguishing features of mmWave propagation are the poor diffraction capability and high penetration loss, which make blockage an important effect. For example, it was observed that there is a received power difference of more than 40 dB at 28 GHz and 73 GHz when a mobile receiver goes around a building corner [86]. Severe attenuation of diffracted signals is also observed in indoor environments [86, 121]. Penetration loss is also more severe than at lower frequencies. For example, the measurements in [121] at 28 GHz show that penetration loss of tinted glass can be as high as 40 dB and penetration through brick can cause up to 28 dB losses. Besides these common materials, human bodies can also cause attenuation. For example, the indoor measurements at 60 GHz in [77] show that human blockage can cause fading with a dynamic range of 35 dB; more loss has been reported in other work [29].

The implication here is that once the LOS path is blocked by an object,

it is unlikely that the signal can propagate past the object through diffraction or penetrate through the object (even with the directional antenna gain). Thus, there could be a sharp decrease in the received power which would require alternative paths to maintain link connection. Therefore, efficient beam alignment and tracking are required to deal with blockage situations.

### 1.3.3 Doppler Spread

The Doppler effect refers to the difference in the perceived frequency of a traveling wave from its true frequency. When the transmitter is moving toward the receiver the perceived frequency is higher than the true frequency, and the perceived frequency becomes lower if the transmitter is moving away from the receiver. The difference between the perceived frequency and the true frequency is called the Doppler shift. When there are multiple paths from the transmitter to the receiver, the Doppler shifts for all the paths make up what is called the Doppler spread. The Doppler spread provides a measure of the severity of the time-variation of the channel.

The Doppler effect, which is expected to increase as the carrier frequency increases, causes some concern that the severity could challenge the PHY layer design, especially in vehicular environments. Classical result assuming the Clarke-Jakes power angular spectrum (signals arriving uniformly in all directions) leads to a conclusion that the Doppler spread increases proportional to the carrier frequency [40, 54], i.e., the Doppler spread would increase by  $30\times$  if the carrier frequency were to change from 2 GHz to 60 GHz. This

argument, however, is inaccurate due to the use of directional beamforming in mmWave communications. In fact, it can be shown that directional beamforming can reduce the severity of the Doppler effect. To properly capture the benefit of directionality in the vehicular context, the change in pointing direction has to be considered. The channel coherence time taking into account the change in pointing direction due to mobility will be derived in Chapter 2. Another important point to note is that the Doppler spread determines the time-variation of the channel and thus will dictate the feasible packet length for communications. The large spectral channels in mmWave bands make it possible to deploy systems with GHz bandwidth which greatly reduces the symbol period as well as the packet length. This implies that the systems will become more robust to time-variation of the channels.

## 1.4 MmWave Transceiver Architectures

Conventional MIMO systems at lower frequency typically assume a fully digital architecture, where each antenna is attached to an RF chain [46]. Due to cost and power consumption, the fully digital architecture does not scale to the mmWave domain [11, 76], where large arrays are required to overcome the propagation challenges.

Fig. 1.2 shows two practical mmWave transmitter architectures (can be constructed similarly for the receiver): analog and hybrid. The analog architecture has only one RF chain (represented by a digital-to-analog converter (DAC)) and the beamforming is controlled by configuring the phases of the



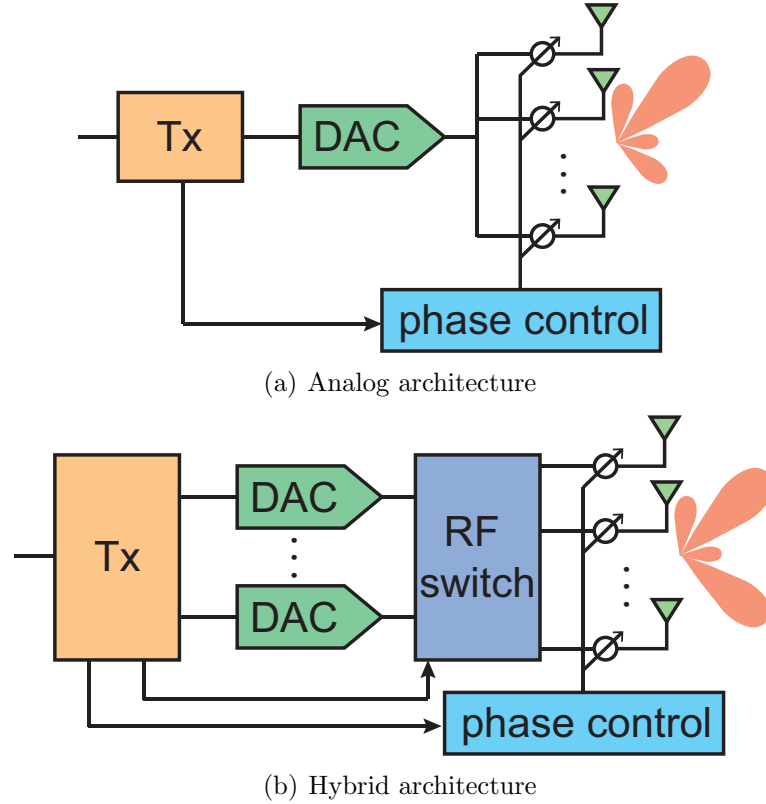


Figure 1.2: Practical mmWave transmitter architectures. The receiver architectures can be constructed similarly. The Tx block is responsible for baseband processing and controlling the phase shifter and RF switches to produce the desired beamforming. Equipped with multiple RF chains (represented by DAC), the hybrid architecture is more flexible and can support spatial multiplexing. This flexibility comes at the cost of a more complex link configuration than that of the analog architecture.

phase-shifters. Note that the phase-shifters can only modify the phases of the signals and not the amplitude, which is commonly called the constant modulus constraint. In the analog beamforming case, the link configuration reduces to finding the beam pointing direction that maximizes the received power, which is called beam alignment. While the analog beam configuration is simpler, it can only support one single stream (i.e., cannot support spatial multiplexing) because there is only one RF chain. A hybrid architecture has several RF chains, typically much less than the number of antennas, which allow it to balance the tradeoff between the fully digital and analog architectures. The multiple RF chains allow it to support spatial multiplexing, up to the number of RF chains. The beamformer/combiner consists of two parts: the digital part and the analog part which is subject to the constant modulus constraint. While more flexible, the link configuration also requires more effort.

In this dissertation, we focus on the analog architecture and develop fast learning-based beam alignment methods. Since the proposed methods provide means to identify promising beam directions, the information is also useful in reducing the overhead for hybrid beamforming. The information forms a prior belief of the channel, and thus by focusing the probing effort in those promising directions, the training overhead can be reduced. This is an interesting future research direction for the solutions proposed in this dissertation.

## 1.5 MmWave V2X: Challenges and Opportunities

The use of pencil beams and susceptibility to blockage make beam-based mmWave communications challenging, especially in a mobile setting such as the vehicular one. While directional beams are also required for indoor applications such as WLAN, in an outdoor setting the link distances are typically larger and it requires much narrower beams, likely at both the transmitter and receiver. This means that the training overhead for the alignment will increase. On top of the high overhead to perform one beam alignment procedure, the high vehicular mobility will cause frequent misalignment either due to blockage by neighboring vehicles or the pointing error due to its own displacement. Therefore, without a fast and efficient beam alignment, most of the communication time will be lost just to keep the beams aligned.

While mmWave V2X has a more stringent requirement on the beam alignment than an indoor use case, it also has several key characteristics that can be exploited. Buildings along the roadside and vehicles themselves are good reflectors that can support alternative propagation paths in blockage situations. Also, while the vehicular environment is dynamic, the road geometry is fixed and has regular patterns. Thus, if the system can identify the current state of the environment the vehicle is in, it could be possible to predict the performance of different pointing directions without any actual beam training. This is indeed possible using the information from the many sensors equipped in modern vehicles. All these point to the opportunities to develop learning algorithms to take full advantage of those sensors to reduce training overhead.

As a first step in this direction, we focus on using the position as the side information from onboard sensors in this dissertation.

## 1.6 Thesis Statement and Summary of Contributions

The thesis statement of this dissertation is

*Position-based learning can exploit the propagation characteristics of the environment to reduce mmWave link configuration overhead.*

In this dissertation, we develop new approaches for fast mmWave beam alignment taking advantages of the side information not traditionally used in communications. While the main focus is on the V2I settings, some of the results are also applicable for the V2V and cellular settings. We start by investigating the relationship between beamwidths and the temporal variation of the vehicular channel. This lays the foundation for understanding the beam training overhead in mmWave V2X. We then develop a new beam alignment method that leverages position information and past beam measurements to identify promising beam directions and reduce the beam alignment overhead. This second contribution is an offline learning approach that requires training data beforehand. In many practical situations, online learning is desirable because the overhead of the deployment can be eliminated and its performance can be improved over time. We develop the online version of our proposed offline approach in the third contribution. In addition to the online beam pair

selection, we also develop a new online beam pair refinement that allows the adaptation of the beam codebook to the environment to further maximize the beamforming gain.

We summarize our contributions as follows:

- Chapter 2: Impact of beamwidth on vehicular channels
  1. Derive closed-form expressions for the channel coherence time incorporating both the pointing error due to mobility and the Doppler effect. The obtained results show that there exists a non-zero optimal beamwidth that maximizes the channel coherence time.
  2. Introduce a new concept of beam coherence time and show that beam alignment should be conducted every beam coherence time and not every channel coherence time.
  - This work was published in [102, 105].
- Chapter 3: Inverse fingerprinting for mmWave V2I beam alignment
  1. Propose a novel and efficient beam alignment method leveraging multipath fingerprints. Fingerprints here refer to long-term spatial channels indexed by locations. The result shows how side information can be used to reduce overhead in mmWave communications.
  2. Define the power loss probability to quantify the beam alignment accuracy. This provides a mathematical framework to optimize the beam pair selection in the proposed method.

3. Propose two statistical learning methods and one machine learning (ML) based approach for offline beam pair selection. The two statistical learning methods include one heuristic and one optimal selection method derived using our developed mathematical framework. The ML-based approach uses a learning-to-rank (LtR) approach from the recommender system literature.

■ This work was published in [103, 104].

- Chapter 4: Online learning for beam pair selection and refinement

1. Develop an online learning version of the optimal beam pair selection proposed in Chapter 3. We use the multi-armed bandit (MAB) framework with a newly introduced risk-awareness component aiming to reduce the probability of severe beam misalignment events during the learning.
2. Propose an online beam pair refinement method to adapt the beam codebook to the environment. The idea is to match the pointing directions of the beams to those of the peaks of the power angular spectrum (PAS) that is environment-dependent. We formulate this problem as a continuum-armed bandit (CAB) problem and solve it using an optimistic optimization approach.

■ This work has been submitted for a possible publication [108].

## 1.7 Organization

The rest of the dissertation is organized as follows. We present the findings on the impact of beamwidth on the temporal channel variation in vehicular environments in Chapter 2. We develop offline beam pair selection methods that leverage position and past multipath information to reduce the beam alignment overhead in Chapter 3. In Chapter 4, we propose an online learning version of the optimal beam selection method with risk-awareness and also develop a beam pair refinement to adapt the beam codebook to the environment to further maximize the beamforming gain. Finally, we conclude the dissertation and describe some potential future research venues in Chapter 5.

## 1.8 Abbreviations

<b>4G</b>	Fourth generation
<b>5G</b>	Fifth generation
<b>AoA</b>	Angle of arrival
<b>AoD</b>	Angle of departure
<b>AWGN</b>	Additive white Gaussian noise
<b>BO</b>	Bayesian optimization
<b>CAB</b>	Continuum-armed bandit
<b>CDF</b>	Cumulative distribution function
<b>DSRC</b>	Dedicated Short-Range Communication
<b>EIRP</b>	Equivalent isotropic radiated power
<b>GPS</b>	Global positioning system

<b>HOO</b>	Hierarchical optimistic optimization
<b>ITS</b>	Intelligent Transportation System
<b>LOS</b>	Line-of-sight
<b>LtR</b>	Learning-to-rank
<b>MAB</b>	Multi-armed bandit
<b>ML</b>	Machine Learning
<b>NLOS</b>	Non-line-of-sight
<b>PAS</b>	Power angular spectrum
<b>PDF</b>	Probability density function
<b>RF</b>	Radio frequency
<b>RL</b>	Reinforcement learning
<b>RSU</b>	Road side unit
<b>SNR</b>	Signal-to-Noise Ratio
<b>ULA</b>	Uniform linear array
<b>UPA</b>	Uniform planar array
<b>V2I</b>	Vehicle-to-infrastructure
<b>V2V</b>	Vehicle-to-vehicle
<b>V2X</b>	Vehicle-to-everything
<b>WSSUS</b>	Wide sense stationary and uncorrelated scattering



## Chapter 2

### Impact of Beamwidth in Vehicular Channels

In this chapter, we characterize the impact of the beamwidth on the temporal variation in vehicular channels. This is done using two concepts: the channel coherence time and a newly defined beam coherence time. Closed-form expressions relating the channel coherence time and beamwidth are derived. Unlike prior work that assumes perfect beam pointing, the pointing error due to the receiver motion is incorporated to show that there exists a non-zero optimal beamwidth that maximizes the channel coherence time. While the *channel* coherence time determines how often to re-estimate the channel coefficients, we define the *beam* coherence time as an effective measure of how often to realign the beams by taking into account the pointing error due to mobility. It is shown that beam alignment in every beam coherence time performs better than the beam alignment in every channel coherence time when overheads (including beam alignment and channel estimation) are included. This work was published in [102, 105] (© IEEE).

## 2.1 Motivation and Prior Work

High data rate millimeter wave (mmWave) communications could enable a whole new range of innovative applications [28, 107], but its use in the vehicular context is often viewed with some skepticism due to concern regarding the Doppler effect. Based on the Clarke-Jakes power angular spectrum (PAS), it follows that the channel coherence time  $T_c$  is inversely proportional to the maximum Doppler frequency  $f_D$ , i.e.,  $T_c \simeq \frac{1}{f_D}$  [40]. This implies that by moving from a typical cellular frequency at around 2 GHz to a mmWave frequency at 60 GHz, one would expect a  $30\times$  decrease in the channel coherence time. This is, in fact, inaccurate for mmWave systems that use directional antennas (or beams) creating angular selectivity in the incoming signal, which effectively decreases the Doppler spread.

Using directional transmission and reception can reduce the effective channel variation at the expense of beam alignment overhead. While the received power is optimal if the beams are aligned in every channel coherence time (i.e., whenever the channel coefficients change), the overhead could be too expensive. The physical beam can be associated with a propagation path (similar to a path of a ray in the ray-tracing model in [99]) whose angle of arrival could change much slower than the fading channel coefficient. This leads to the concept of the beam coherence time. One natural question is how much is lost if the beams are realigned at this slower speed? We show in this chapter that beam coherence time is the right choice when all the overheads (for both beam alignment and channel estimation) are included.

Relevant prior work includes [25,84,95,120] that characterized the channel correlation under non-isotropic scattering, which cause signals to concentrate in the angular domain. While in this chapter, a directional receive beam is used to control the angular selectivity, both result in a similar effect. The difference is whether the selectivity is controlled by the receiver or up to the environment. Generally, there are two directions in this line of research: one is to provide a generalized framework that can be used for any scattering distribution [84,95] and the other is to constrain to a given distribution that allows tractable expressions for further analysis [25,120]. The work in [95] presented a generalized framework to compute a spatial correlation function for general 3D scattering distributions. Their result was based on the decomposition of the plane wave into an infinite sum of the spherical Bessel functions and Legendre polynomials. A similar approach was used in [84] to compute correlation functions in 2D while also taking the antenna patterns into account. For the 2D case, the plane wave is decomposed into an infinite sum of the Bessel functions. Although the approaches in [95] and [84] are general, the obtained correlation functions are intractable for further analysis. The work in [25,120] instead considered only the von Mises scattering distribution and derived closed-form correlation functions using two-ring models. Our approach follows this later path and adopts the von Mises distribution to represent the effective PAS. Different from [25,120], we also incorporate the pointing error due to the receiver motion into the correlation functions, which is an essential characteristic when using directional beams in vehicular environments.

Other related work appears in [27, 33, 81, 88]. The relationship between the channel coherence time and beamwidth was studied in [27, 33]. A general framework to compute the coherence time was derived in [33] for any PAS. The correlation was defined using the channel amplitude and the main assumption was that the channel is Rayleigh faded. We define correlation using the complex channel coefficient which considers both the amplitude and the phase. The work in [27] relates the coherence time with the number of antennas of a uniform linear array. A simple expression was derived for a special case when the pointing angle is  $90^\circ$ . The work in [81] exploits the decrease in Doppler spread due to directionality and proposed a beam partitioning method in a rich scattering environment such that each beam experiences the same amount of Doppler spread. Note that in [27, 33, 81], no pointing error was considered and their results suggest that the coherence time goes to infinity when the beamwidth approaches zero. We incorporate pointing error due to the receiver mobility and show that there exists a non-zero optimal beamwidth that maximizes the channel coherence time. To the best of our knowledge, this is the first time that both the Doppler and pointing error are incorporated to derive the channel coherence time. Recently, [88] quantified the channel coherence time considering pointing error due to wind-induced vibration for a mmWave wireless backhaul application. Note that the source of pointing error in [88] is different from the one considered in this chapter.

## 2.2 Contributions

Our main objective in this chapter is to establish the potential of the mmWave vehicular communications using directional beams in fast changing vehicular environments. Our main contributions in this chapter are summarized as follows:

- We derive a channel temporal correlation function taking into consideration both the pointing error due to the receiver motion and Doppler effect. Based on the obtained correlation function, we derive the channel coherence time and show how it connects to the receive beamwidth and the pointing direction. Our results show that there exists a non-zero optimal beamwidth that maximizes the channel coherence time, unlike prior work that assumes perfect beam pointing.
- We propose a new concept called the beam coherence time, which is used as the basis for studying the long-term beam realignment. This lays the foundation for the third contribution.
- We investigate the choice of the beam realignment duration taking into account both the beam alignment overhead and the loss due to the channel time-variation. We show that long-term beam realignment performs better and thus the beams should be realigned every *beam* coherence time, not every *channel* coherence time.

Note that although our baseband channel model is general, our focus is on

mmWave bands and accordingly all of our numerical examples use parameters from the 60 GHz band.

The rest of this chapter is organized as follows. Section 2.3 describes our models and assumptions. Using the models, we derive novel channel temporal correlation functions taking the pointing error into account for both the LOS and NLOS cases in Section 2.4. In Section 2.5, we derive the channel coherence time from the obtained correlation functions. In Section 2.6, we define a novel beam coherence time, which is tailored to the beam alignment concept. Based on these results, we investigate some implications for the beam alignment duration in Section 2.7. Finally, Section 2.8 concludes the chapter.

## **2.3 Models and Assumptions**

This section starts with the channel model and then introduces the pointing error due to the receiver motion. Next, we describe a spatial lobe model that provides a statistical description of the angular spread of the PAS. The spatial lobe model will be used in the derivation of the beam coherence time. Finally, we provide a table summarizing the common parameters we use in our numerical examples.

### **2.3.1 Channel Model**

This subsection first describes the NLOS channel, which will be later incorporated into the LOS channel model. For the NLOS channel, we assume a narrowband wide sense stationary and uncorrelated scattering (WSSUS)

model given by [40]

$$h_{\text{nlos}}(t) = \int_{-\pi}^{\pi} \sqrt{\chi \mathcal{P}'(\alpha) G(\alpha|\mu_r)} e^{j[\phi_0(\alpha) + \phi(\alpha) + 2\pi f_D t \cos(\alpha)]} d\alpha. \quad (2.1)$$

Here,  $\chi$  is a normalization constant,  $\mathcal{P}'(\alpha)$  is the PAS,  $G(\alpha|\mu_r)$  is the antenna pattern with the main lobe pointing at  $\mu_r$ ,  $\phi_0(\alpha)$  is the phase due to the distance traveled up to time 0,  $\phi(\alpha)$  is the random phase associated with the path with the angle of arrival  $\alpha$ , and  $f_D$  is the maximum Doppler frequency. Note that all angles are defined in reference to the direction of travel of the receiver (Fig. 2.1). Under the uncorrelated scattering assumption,  $\phi(\alpha)$  are uncorrelated and uniformly distributed in  $[0, 2\pi)$ . For the time scale considered, it is assumed that the scatterers are stationary. This is the wide sense stationary assumption that is reasonable for a short duration. Note that although the channel model here assumes a large number of paths, our simulation results in Fig. 2.3 show that our results also hold for small numbers of paths.

We define the effective PAS  $\mathcal{P}(\alpha|\mu_r) = \chi \mathcal{P}'(\alpha) G(\alpha|\mu_r)$ . To ensure unit power channel coefficients, i.e.  $\mathbb{E}[|h_{\text{nlos}}(t)|^2] = 1$  with  $\mathbb{E}[\cdot]$  denoting the expectation operator,  $\chi$  has to satisfy  $\int_{-\pi}^{\pi} \mathcal{P}(\alpha|\mu_r) d\alpha = 1$ . We assume  $G(\alpha|\mu_r)$  takes the shape of the von Mises probability density function (PDF). We assume that the PAS  $\mathcal{P}'(\alpha)$  has angular spread much larger than the beamwidth (more accurate for narrow beams) so that it is flat over the range of  $\alpha$  where the beam pattern has non-negligible values. This assumption means  $\chi \mathcal{P}'(\alpha) \simeq 1$  so that  $\mathcal{P}(\alpha|\mu_r) \simeq G(\alpha|\mu_r)$  which is the von Mises PDF given by

$$\mathcal{P}(\alpha|\mu_r) \simeq \frac{1}{2\pi I_0(k_r)} e^{k_r \cos(\alpha - \mu_r)}, \quad (2.2)$$

where  $I_0(\cdot)$  is the zeroth order modified Bessel function of the first kind,  $\mu_r$  is the mean, and  $k_r$  is the shape parameter. Note that the derivations of the correlation function and the channel coherence time use only the effective PAS. The beam coherence time is defined in terms of the spatial lobe and will need an explicit model of  $\mathcal{P}'(\alpha)$ . Following the model in [101],  $\mathcal{P}'(\alpha)$  is assumed to take the shape of a Gaussian PDF with variance  $\beta^2$ . Some examples of the use of Gaussian PDF in this context are its adoption as an antenna pattern in a 5G channel model in [4, Section 5.3.7.2], and as the model for angles of arrival in another 5G channel model in [101]. The von Mises PDF can be thought of as a circular version of the Gaussian PDF, and when  $k_r$  is large, it can be approximated by a Gaussian PDF with the same mean  $\mu_r$  and variance of  $1/k_r$ . We define the beamwidth  $\theta$  by  $k_r \simeq 1/\theta^2$ . The assumption in (2.2) becomes more accurate when  $\beta$  is large compared to the beamwidth. We choose the von Mises PDF for two reasons: (i) its good resemblance to a real antenna pattern and (ii) its tractability for analysis.

Next, we describe our LOS channel model. Introducing the LOS component, the channel coefficient now becomes

$$h(t) = \sqrt{\frac{K}{K+1}} h_{\text{los}}(t) + \sqrt{\frac{1}{K+1}} h_{\text{nlos}}(t), \quad (2.3)$$

where  $K$  is the Rician K factor, which determines the relative power between the LOS and NLOS components. The LOS component is modeled as

$$h_{\text{los}}(t) = \int_{-\pi}^{\pi} \sqrt{G(\alpha|\mu_r)} e^{-j\frac{2\pi}{\lambda}D} e^{j2\pi f_D t \cos(\alpha)} \delta(\alpha_{\text{los}} - \alpha) d\alpha, \quad (2.4)$$



where  $D$  is the distance between the transmitter and the receiver at time 0,  $\alpha_{\text{los}}$  is the angle of arrival of the LOS path, and  $\delta(\cdot)$  denotes the Dirac delta function [120].

### 2.3.2 Pointing Error Due to Receiver Motion

The model to be described here is based on the observation that if the receive beam is fixed, and the receiver moves, then the beam will become misaligned. Misalignment implies that the receiver sees the channel with a different lens than when properly aligned and thus the channel temporal correlation will be affected. Note that the receive beam pointing direction is  $\mu_r$ , which can be in the LOS or NLOS direction.

We use the one-ring model for the NLOS, where scatterers are distributed on a ring of radius  $D_r$  as shown in Fig. 2.1. Let the receiver be at point A at time  $t$  and move at a constant speed  $v$  along the direction of travel to reach point B at time  $t + \tau$ . The total displacement from A to B is  $\Delta_d(\tau) = v\tau$ . When the receiver moves from A to B by  $\Delta_d(\tau)$ , the receiver will see a different set of scatterers and the distances to the scatterers also change. We assume that  $\Delta_d(\tau) \ll D_r$ , so that  $\Delta_d(\tau)$  has negligible effect on the path loss and captures the receiver motion effect through the pointing error  $\Delta_\mu(\tau)$  as shown in Fig. 2.1. Note that this pointing error is the angular difference needed to correct the initial pointing direction  $\mu_r$  at A so that the beam always sees the same set of scatterers. For notational convenience,  $\Delta_\mu, \Delta_d$  are used instead of  $\Delta_\mu(\tau), \Delta_d(\tau)$ . The relationship between  $\Delta_d$  and  $\Delta_\mu$  can be obtained

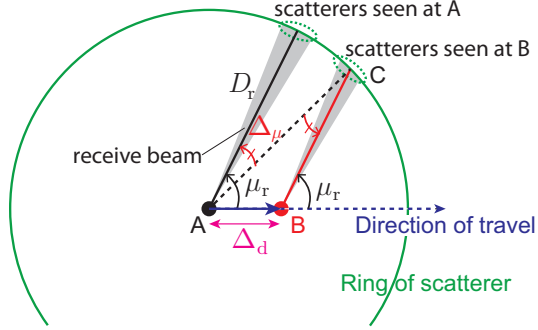


Figure 2.1: Receiver displacement and change in pointing angle for the NLOS case. When the receiver moves from A to B, if the beam is not adaptive, then the set of scatterers seen at B will be different from those seen at A. This effect can be captured by the change in pointing direction  $\Delta_\mu$  (we called this pointing error) from the original pointing direction  $\mu_r$ . We ignore the change in path loss due to the displacement  $\Delta_d$  because for a short time duration  $\tau$  typically  $\Delta_d(\tau) \ll D_r$ .

using the law of sines on the triangle ABC in Fig. 2.1 to get

$$\frac{\Delta_d}{\sin \Delta_\mu} = \frac{D_r}{\sin(\pi - \mu_r)}. \quad (2.5)$$

For small  $\Delta_\mu$ ,  $\sin \Delta_\mu \simeq \Delta_\mu$ , and since  $\sin(\pi - \mu_r) = \sin \mu_r$ , we have

$$\Delta_\mu \simeq \frac{\Delta_d}{D_r} \sin \mu_r. \quad (2.6)$$

Since  $f_D = v/\lambda$ , we have  $\Delta_d = v\tau = f_D\lambda\tau$ , where  $\lambda$  is the carrier wavelength.

Substituting this into (2.6) to get

$$\Delta_\mu \simeq \frac{f_D\tau}{D_{r,\lambda}} \sin \mu_r, \quad (2.7)$$

where  $D_{r,\lambda} = D_r/\lambda$  is the scattering radius normalized by the carrier wavelength  $\lambda$ .

The same reasoning can be applied to the LOS case by replacing the scattering radius  $D_r$  by the transmitter-receiver distance  $D$ . Let  $D_\lambda = D/\lambda$  and  $\alpha_{\text{los}}$  be the direction toward the transmitter (in reference to the direction of travel), and define the pointing error  $\Delta_\mu^{\text{los}}$  as the angular difference needed to correct the beam direction so that it always points directly at the transmitter, we have

$$\Delta_\mu^{\text{los}} \simeq \frac{f_D \tau}{D_\lambda} \sin(\alpha_{\text{los}}). \quad (2.8)$$

The approximate relations (2.7) and (2.8) will be used in later derivations in this chapter.

### 2.3.3 Channel Spatial Lobe Model

In this subsection, we explain the spatial lobe model that will be the basis for our definition of the beam coherence time. We consider only the azimuthal plane. This model provides a statistical description of the angular spread of the PAS. A signal transmitted from the transmitter propagates through different paths to arrive at the receiver. These multipaths arrive at different angles with some concentrations at certain angles which create patterns as illustrated in Fig. 2.2, which are called spatial lobes. Four spatial lobes are shown in Fig. 2.2. These lobes can be thought of as clusters of scatterers with distinct angles of arrival. The number of spatial lobes depends on the environment and ranges from 1-6 in an urban environment measurement at 28 GHz [89]. Beam alignment is the process of finding the direction of the spatial lobe with the highest power, i.e., the lobe with the highest peak (lobe #1 in

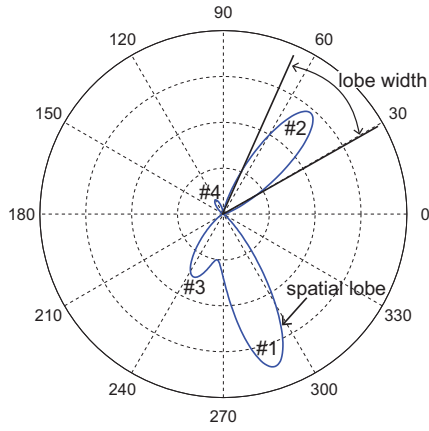


Figure 2.2: An illustration of spatial lobes. This figure illustrates the pattern of the incoming power arriving at the receiver. The incoming power has strong spatial dependence, and it can be observed here that there are four main directions, which can be thought of as four clusters of scatterers. At each of these directions, there is spread forming a lobe, which is termed as a spatial lobe. In this example, lobe #1 has the strongest power.

Fig. 2.2). The lobe width determines the difficulty in aligning the beam. The narrower the spatial lobe, the more difficult the alignment becomes, and the easier the beam gets misaligned due to the receiver motion. Thus, this lobe width plays a fundamental role in defining the beam coherence time.

The lobe width  $\beta$ , which is the standard deviation of the PAS  $\mathcal{P}'(\alpha)$ , is modeled following the empirical model proposed in [89], which uses a Gaussian distribution, i.e.,

$$\beta \sim \mathcal{N}(m_{\text{AS}}, \sigma_{\text{AS}}^2). \quad (2.9)$$

The mean  $m_{\text{AS}}$  and the standard deviation  $\sigma_{\text{AS}}$  depend on the environment. The model in [89] was based on measurements in an urban area, where  $\sigma_{\text{AS}} =$

Table 2.1: Common parameters used in numerical examples.

Parameter	Value
Carrier frequency $f_c$	60 GHz
Vehicle speed $v$	30 m/s
Scattering radius $D_{r,\lambda}$	1000 $\lambda$
Standard deviation $\sigma_{AS}$	25.7 deg.

25.7° was derived. In our numerical examples, different values of  $m_{AS}$  are used, but  $\sigma_{AS}$  is always fixed to 25.7°.

### 2.3.4 Summary of Common Parameters

This subsection summarizes the common parameters that are used in our numerical examples in this chapter. Note that except for the beamwidth and the pilot spacing (introduced in Section 2.7) which are system parameters, all others are channel parameters. These common parameters are shown in Table 2.1. When values different than Table 2.1 are used, it will be explicitly indicated. Although our result can be applied to any carrier frequency, we focus on mmWave bands and set the carrier frequency to  $f_c = 60$  GHz. We assume a highway scenario and set the vehicle speed to  $v = 30$  m/s. The scattering radius is set to 5 m, which is equivalent to 1000 wavelengths at 60 GHz. The width of the spatial lobe is modeled as Gaussian as described in the previous subsection, and we set the standard deviation to  $\sigma_{AS} = 25.7^\circ$ .

## 2.4 Channel Temporal Correlation Function

There are two possible definitions of the channel temporal correlation function. The first one is based on the amplitude of the channel coefficients [33]

$$R_{|h|}(\tau) = \frac{\mathbb{E}[g(t)g(t+\tau)] - (\mathbb{E}[g(t)])^2}{\mathbb{E}[g(t)^2] - (\mathbb{E}[g(t)])^2}, \quad (2.10)$$

where  $g(t) = |h(t)|$ . The second definition is based on the complex channel coefficients themselves [27] and is defined as

$$R_h(\tau) = \mathbb{E}[h(t)h^*(t+\tau)], \quad (2.11)$$

where  $(\cdot)^*$  denotes complex conjugate. Most modern communication systems use coherent detection, where both the amplitude and phase are important. In that respect, the definition in (2.11) is more natural and is the definition we adopt. Note that when  $h(t)$  is complex Gaussian, the two definitions are in fact equivalent [54, Pages 47-51], in the sense that there is a simple relationship between the two. In particular, it can be shown that  $R_{|h|}(\tau) = \frac{\pi}{4(4-\pi)} |R_h(\tau)|^2$ .

The channel model in (2.3) has both LOS and NLOS components. For the LOS component,  $h_{\text{los}}(t)$  depends on the pointing direction, and proper normalization is needed to be consistent with (2.11). We still define the correlation function for the LOS component  $R_{\text{los}}(\tau)$  based on the product  $h_{\text{los}}(t)h_{\text{los}}^*(t+\tau)$ , but now we introduce a normalization such that  $|R_{\text{los}}(\tau = 0)| = 1$  and  $|R_{\text{los}}(\tau \neq 0)| < 1$  in Section 2.4.2. Along with this definition, the correlation function is defined as

$$R_h(\tau) = \frac{K}{K+1} R_{\text{los}}(\tau) + \frac{1}{K+1} R_{\text{nlos}}(\tau). \quad (2.12)$$

In the following, we derive the correlation function for the NLOS channel using (2.11) in Section 2.4.1 and define the correlation function for the LOS in Section 2.4.2 that is consistent with the definition in (2.11). In both cases, the pointing error due to the receiver motion is incorporated.

### 2.4.1 NLOS Channel Correlation Function

Here, we derive the correlation function between  $h_{\text{nlos}}(t)$  and  $h_{\text{nlos}}(t+\tau)$  for the NLOS channel. The channel coefficients at time  $t$  and  $t+\tau$  are given by

$$h_{\text{nlos}}(t) = \int_{-\pi}^{\pi} \sqrt{\mathcal{P}(\alpha|\mu_r)} e^{j[\phi_0(\alpha)+\phi(\alpha)+2\pi f_D t \cos(\alpha)]} d\alpha, \quad (2.13)$$

$$h_{\text{nlos}}(t+\tau) = \int_{-\pi}^{\pi} \sqrt{\mathcal{P}(\alpha|\mu_r + \Delta_\mu)} e^{j[\phi_0(\alpha)+\phi(\alpha)+2\pi f_D (t+\tau) \cos(\alpha)]} d\alpha, \quad (2.14)$$

where we have incorporated the pointing error due to the receiver motion in the peak direction of the effective PAS, which is now  $\mu_r + \Delta_\mu$  instead of  $\mu_r$  in (2.14). Note that in (2.13) and (2.14), although  $\alpha$  is taken from  $-\pi$  to  $\pi$ , the channel is non-isotropic scattering because the incoming signals are weighted by the effective PAS  $\mathcal{P}(\alpha|\mu_r)$ , which takes the shape of the von Mises PDF. Plugging (2.13) and (2.14) into (2.11),

$$\begin{aligned} R_{\text{nlos}}(\tau) &= \mathbb{E} \left[ \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sqrt{\mathcal{P}(\alpha_1|\mu_r) \mathcal{P}(\alpha_2|\mu_r + \Delta_\mu)} \right. \\ &\quad \left. e^{j(\phi_0(\alpha_1)+\phi(\alpha_1)-\phi_0(\alpha_2)-\phi(\alpha_2)-2\pi f_D \tau \cos(\alpha_2))} d\alpha_1 d\alpha_2 \right] \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sqrt{\mathcal{P}(\alpha_1|\mu_r) \mathcal{P}(\alpha_2|\mu_r + \Delta_\mu)} \mathbb{E}[e^{j(\phi(\alpha_1)-\phi(\alpha_2))}] \\ &\quad e^{j(\phi_0(\alpha_1)-\phi_0(\alpha_2))} e^{-j2\pi f_D \tau \cos(\alpha_2)} d\alpha_1 d\alpha_2 \\ &= \int_{-\pi}^{\pi} \sqrt{\mathcal{P}(\alpha|\mu_r) \mathcal{P}(\alpha|\mu_r + \Delta_\mu)} e^{-j2\pi f_D \tau \cos(\alpha)} d\alpha, \end{aligned} \quad (2.15)$$

where (2.15) follows from the uncorrelated scattering assumption. According to this assumption,  $\mathbb{E}[e^{j(\phi(\alpha_1)-\phi(\alpha_2))}] = \mathbb{E}[e^{j\phi(\alpha_1)}]\mathbb{E}[e^{-j\phi(\alpha_2)}] = 0$  for  $\alpha_1 \neq \alpha_2$  and  $\mathbb{E}[e^{j(\phi(\alpha_1)-\phi(\alpha_2))}] = 1$  for  $\alpha_1 = \alpha_2$ , where  $\phi(\alpha)$  is uniformly distributed in  $[0, 2\pi)$ . Now, substituting the von Mises PDF to get

$$\begin{aligned}
R_{\text{nlos}}(\tau) &= \frac{1}{2\pi I_0(k_r)} \int_{-\pi}^{\pi} \sqrt{e^{k_r \cos(\alpha-\mu_r)} e^{k_r \cos(\alpha-(\mu_r+\Delta_\mu))}} e^{-j2\pi f_D \tau \cos(\alpha)} d\alpha \quad (2.16) \\
&= \frac{1}{2\pi I_0(k_r)} \int_{-\pi}^{\pi} e^{k_r \cos(\alpha-\mu_r-\frac{\Delta_\mu}{2}) \cos(\frac{\Delta_\mu}{2})} e^{-j2\pi f_D \tau \cos(\alpha)} d\alpha \\
&= \frac{1}{2\pi I_0(k_r)} \int_{-\pi}^{\pi} e^{k'_r \cos(\alpha-\mu')} e^{-j2\pi f_D \tau \cos(\alpha)} d\alpha \\
&= \frac{1}{2\pi I_0(k_r)} \int_{-\pi}^{\pi} e^{x' \cos \alpha + y' \sin \alpha} d\alpha \\
&= \frac{I_0(\sqrt{x'^2 + y'^2})}{I_0(k_r)}, \quad (2.17)
\end{aligned}$$

where,  $k'_r = k_r \cos\left(\frac{\Delta_\mu}{2}\right)$ ,  $\mu' = \mu_r + \frac{\Delta_\mu}{2}$ ,  $x' = k'_r \cos \mu' - j2\pi f_D \tau$ ,  $y' = k'_r \sin \mu'$ , and we have used the formula  $\int_{-\pi}^{\pi} e^{a \cos c + b \sin c} dc = 2\pi I_0(\sqrt{a^2 + b^2})$  [42, 3.338-4]. Despite the simple form of (2.17), it is intractable for further analysis because the argument to the Bessel function involves the cosine of  $\Delta_\mu$ , which is also a function of  $\tau$ . Fortunately, a more tractable approximated form can be obtained for large  $k_r$ , where the von Mises PDF can be approximated by the Gaussian one with the variance of  $1/k_r$  and the same mean. With this approximation, (2.16) becomes

$$R_{\text{nlos}}(\tau) \simeq \frac{1}{\sqrt{2\pi/k_r}} \int_{-\pi}^{\pi} e^{-\frac{k_r}{4}((\alpha-\mu_r)^2 + (\alpha-(\mu_r+\Delta_\mu))^2)} e^{-j2\pi f_D \tau \cos(\alpha)} d\alpha. \quad (2.18)$$

The exponent of the first term in the integral can be rewritten as  $2(\alpha - \mu')^2 + \Delta_\mu^2/2$ . Substituting this into (2.18) and approximating  $\mu' \simeq \mu$ , which is valid



for small  $\Delta_\mu$ , we have

$$R_{\text{nlos}}(\tau) \simeq \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}/k_r} e^{\frac{k_r}{2}(\alpha-\mu_r)^2} e^{-\frac{k_r \Delta_\mu^2}{8}} e^{-j2\pi f_D \tau \cos(\alpha)} d\alpha. \quad (2.19)$$

To obtain a final closed-form expression, the Gaussian PDF is approximated back to von Mises one to get

$$\begin{aligned} R_{\text{nlos}}(\tau) &\simeq \frac{e^{-\frac{k_r \Delta_\mu^2}{8}}}{2\pi I_0(k_r)} \int_{-\pi}^{\pi} e^{k_r \cos(\alpha-\mu_r)} e^{-j2\pi f_D \tau \cos(\alpha)} d\alpha \\ &= e^{-\frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2}} \frac{I_0(\sqrt{x^2 + y^2})}{I_0(k_r)} \end{aligned} \quad (2.20)$$

where,

$$x = k_r \cos \mu_r - j2\pi f_D \tau, \quad (2.21)$$

$$y = k_r \sin \mu_r. \quad (2.22)$$

In this chapter, we are interested in narrow receive beamwidths (i.e.,  $k_r$  large), and this approximation turns out to be decent enough for our purpose as will be shown in the numerical examples at the end of this section. Note that in the approximation in (2.20), the effect of the pointing error due to the receiver motion is decoupled from the effect of the Doppler spread to the channel.

#### 2.4.2 LOS Channel Correlation Function

The correlation function for the LOS channel is defined as

$$R_{\text{los}}(\tau) = \frac{h_{\text{los}}^*(t+\tau)h_{\text{los}}(t)}{\max\{|h_{\text{los}}(t)|^2, |h_{\text{los}}(t+\tau)|^2\}}, \quad (2.23)$$

where the normalization is to ensure that  $|R_{\text{los}}(\tau)| \leq 1$ . Substituting the channel in (2.4),

$$R_{\text{los}}(\tau) = \frac{\sqrt{G(\alpha_{\text{los}}|\mu_{\text{r}})G(\alpha_{\text{los}}|\mu_{\text{r}} + \Delta_{\mu}^{\text{los}})}}{\max\{G(\alpha_{\text{los}}|\mu_{\text{r}}), G(\alpha_{\text{los}}|\mu_{\text{r}} + \Delta_{\mu}^{\text{los}})\}} \times e^{j2\pi f_{\text{D}}t[\cos(\alpha_{\text{los}}) - \cos(\alpha_{\text{los}} + \Delta_{\mu}^{\text{los}})]} e^{-j2\pi f_{\text{D}}\tau \cos(\alpha_{\text{los}} + \Delta_{\mu}^{\text{los}})}, \quad (2.24)$$

where we have incorporated the receive beam pointing error due to the receiver motion over the time period  $\tau$  through  $\Delta_{\mu}^{\text{los}}$  as given in (2.8). Note that (2.24) depends on  $t$  and thus is not wide sense stationary. In the case of small  $\Delta_{\mu}^{\text{los}}$ , it can be approximated as wide sense stationary as the term  $e^{j2\pi f_{\text{D}}t[\cos(\alpha_{\text{los}}) - \cos(\alpha_{\text{los}} + \Delta_{\mu}^{\text{los}})]} \simeq 1$ . Also, note that  $|R_{\text{los}}(\tau)| = 1$  only when  $\Delta_{\mu}^{\text{los}} = 0$ .

If we assume that at time  $t$  the receive beam is pointing at  $\alpha_{\text{los}}$ , then  $G(\alpha_{\text{los}}|\mu_{\text{r}} = \alpha_{\text{los}}) = e^{k_{\text{r}}}/(2\pi I_0(k_{\text{r}}))$  and  $G(\alpha_{\text{los}}|\mu_{\text{r}} + \Delta_{\mu}^{\text{los}}) = e^{k_{\text{r}} \cos(\Delta_{\mu}^{\text{los}})}/(2\pi I_0(k_{\text{r}}))$  and we have

$$R_{\text{los}}(\tau) = \sqrt{e^{k_{\text{r}}(\cos(\Delta_{\mu}^{\text{los}}) - 1)}} e^{j2\pi f_{\text{D}}t[\cos(\alpha_{\text{los}}) - \cos(\alpha_{\text{los}} + \Delta_{\mu}^{\text{los}})]} e^{-j2\pi f_{\text{D}}\tau \cos(\alpha_{\text{los}} + \Delta_{\mu}^{\text{los}})} \quad (2.25)$$

$$\simeq e^{\frac{1}{2}k_{\text{r}}(\cos(\Delta_{\mu}^{\text{los}}) - 1)} e^{-j2\pi f_{\text{D}}\tau \cos(\alpha_{\text{los}})}, \quad (2.26)$$

where the approximation holds for small  $\Delta_{\mu}^{\text{los}}$ , which typically is the case because the transmitter-receiver distance  $D$  is large. Taking the absolute value of either (2.25) or (2.26) gives

$$|R_{\text{los}}(\tau)| = e^{\frac{1}{2}k_{\text{r}}(\cos(\Delta_{\mu}^{\text{los}}) - 1)}. \quad (2.27)$$

The expression in (2.27) means that the only factor affecting the channel correlation of the LOS channel is the pointing error.

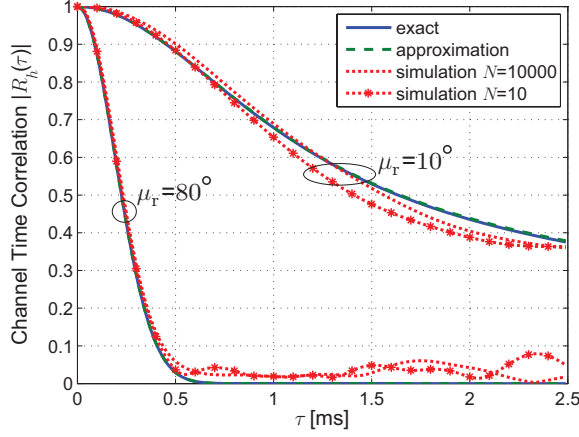


Figure 2.3: Comparison of the channel temporal correlations computed from the exact expression (2.17), the approximate expression (2.20), and those computed from simulation. The beamwidth is set to  $\simeq 8^\circ$  ( $k_r = 50$ ), the number of paths  $N = 10$  and  $N = 10000$  are used, and other parameters are defined in Table 2.1.

#### 2.4.3 Numerical Verification of (2.20) and Effect of $K$ Factor

First, we will verify our approximation for the NLOS case in (2.20) by comparing it with the exact expression given in (2.17) and the correlation computed from simulations. We set the speed  $v = 30$  m/s, the carrier frequency  $f_c = 60$  GHz, the scattering radius  $D_{r,\lambda} = 1000$  wavelengths as in Table 2.1. To simulate the channel realizations, we need the transmitter-receiver distance  $D$ , and  $D = 50$  m is used. We compute the case when  $\mu_r = 10^\circ$  and when  $\mu_r = 80^\circ$  to compare the effect of  $\mu_r$ . We fix  $k_r = 50$  (beamwidth  $\simeq 8^\circ$ ) because there was no assumption on the receive beamwidth in the derivation, and the accuracy of this approximation does not depend on it. We simulate the channel following the sum of sinusoid approach [120] using the model given in (2.1).

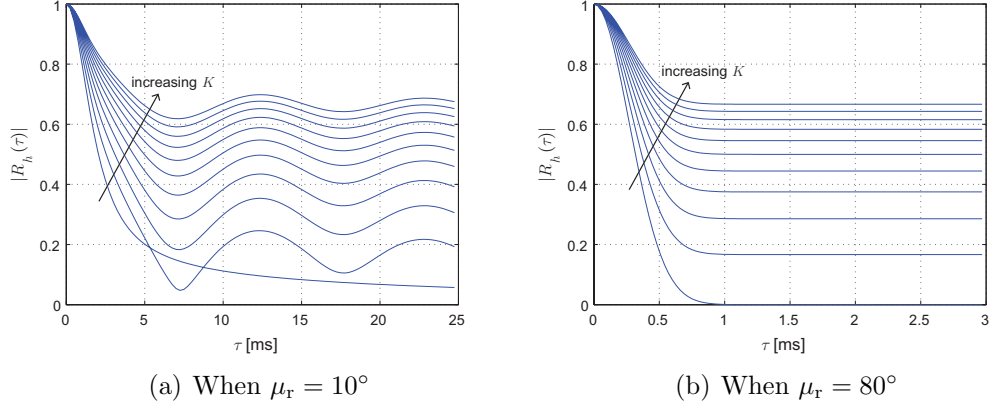


Figure 2.4: The effect of the Rician  $K$  factor on the channel correlation function. The plots show the absolute values of the correlation coefficients for  $K$  from 0 to 2 with a step of 0.2.

Note that the transmitter-receiver distance  $D$  is used only in the simulation and is not used in the exact or the approximate expression given in (2.17) and (2.20). As can be seen in Fig. 2.3, all the curves match well. Although the derived expression is based on the assumption of a large number of paths  $N$ , simulations using a small  $N$  of only 10 paths yield very close results. We conclude that our results also hold for small numbers of paths.

Since the derivation of the correlation for the LOS channel is simple, we skip its verification. Instead, we provide an example showing how the  $K$  factor affects the channel correlation as shown in Fig. 2.4. The parameters are the same as in the NLOS case. The channel correlation oscillates for  $\mu_r = 10^\circ$  but not for  $\mu_r = 80^\circ$ . The oscillation is due to the phase difference between the LOS component  $R_{\text{los}}(\tau)$  and the NLOS component  $R_{\text{nlos}}(\tau)$ . When  $\mu_r$  is close to  $90^\circ$ ,  $R_{\text{nlos}}(\tau)$  decreases fast (note that  $\mu_r = 90^\circ$  corresponds to the

fastest fading case) so that before the phase difference between  $R_{\text{los}}(\tau)$  and  $R_{\text{nlos}}(\tau)$  takes effect, the NLOS component  $R_{\text{nlos}}(\tau)$  decreases to a negligible value in relation to  $R_{\text{los}}(\tau)$  and the oscillation is not observed. The plots also show that the channel correlation increases with  $K$  regardless of  $\mu_r$ . This is because  $\Delta_\mu^{\text{los}}$  increases very slowly with  $\tau$  (because typically  $D_\lambda \gg D_{r,\lambda}$ , and specifically in this example  $D_\lambda = 10D_{r,\lambda}$ ) which leads to very slow decrease in the LOS correlation component  $|R_{\text{los}}(\tau)|$  compared to that of the NLOS component  $|R_{\text{nlos}}(\tau)|$ .

In general, for both LOS and NLOS, the temporal correlation decreases quickly for  $\mu_r$  close to  $90^\circ$  while it decreases slowly for  $\mu_r$  close to zero. The correlation of the LOS case decreases slower than the NLOS case because the pointing error caused by the mobility of the receiver is smaller due to the fact that  $D > D_r$  typically holds. That the correlation of the LOS component decreases slower than that of the NLOS component is the reason why the overall correlation decreases slower as  $K$  increases (see Fig. 2.4).

## 2.5 Channel Coherence Time

Using the channel correlation functions, we derive the coherence times in this section. The channel coherence time is defined as the time  $\tau = T_c$  at which the channel correlation decreases to  $|R_h(\tau)| = R$  for some predefined value  $R$ . Typically,  $R$  ranges from 0.3 to 0.9 [75]. Note that for a given channel, requiring a larger  $R$  will result in a smaller  $T_c$ . A general solution to  $|R_h(\tau)| = R$  is intractable because both  $R_{\text{los}}(\tau)$  and  $R_{\text{nlos}}(\tau)$  are complex

numbers and  $R_{\text{nlos}}(\tau)$  is a complicated function involving the Bessel function. Instead of dealing with this directly, we will derive the coherence time for the NLOS and LOS case separately, which serve as upper and lower bounds on the channel coherence time.

### 2.5.1 NLOS Channel

The coherence time expressions for the case when only the NLOS component exists are derived using the correlation function given in (2.20). Due to the Bessel function, the solution for a general main beam direction  $\mu_r$  and arbitrary beamwidth  $\theta$  is intractable. In the following, we assume small  $\theta$  and we will solve for two cases, namely when  $|\mu_r|$  is small and when  $|\mu_r|$  is not small. Note that the case where  $\mu_r = 0$  is when the main beam direction is parallel to the direction of travel resulting in the slowest fading case, while  $|\mu_r| = \pi/2$  is when the main beam direction is perpendicular to the direction of travel and the receiver will experience the fastest fading [33]. For most cases, our approximation is valid for beamwidth  $\theta$  up to around  $20^\circ$ . This is not a serious limitation because most likely mmWave systems will use narrow beams. For example, a prototype system developed by Samsung Electronics uses an array with  $10^\circ$  beamwidth [111] and long-range automotive radars use a beamwidth on the order of a few degrees [116].

### 2.5.1.1 When $|\mu_r|$ is small

For small  $\mu_r$ , we approximate  $x \simeq k_r - j2\pi f_D \tau$  and  $y \simeq 0$ , where  $x$  and  $y$  are defined in (2.21) and (2.22). The accuracy of this approximation depends on both  $k_r$  and  $\mu_r$ . When  $\mu_r$  is small,  $y \simeq k_r \mu_r \simeq \mu_r / \theta^2$  and roughly the approximation works for  $\theta > \sqrt{\mu_r}$ . Assuming  $\theta$  is in this range, we have

$$R_h(\tau) \simeq e^{-\frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2}} \frac{I_0(k_r - j2\pi f_D \tau)}{I_0(k_r)} \quad (2.28)$$

$$\simeq e^{-\frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2}} \frac{e^{-j2\pi f_D \tau}}{\sqrt{1 - j2\pi f_D \tau / k_r}}. \quad (2.29)$$

The last step follows by applying the asymptotic approximation of the Bessel function [6]

$$I_0(z) \simeq \frac{e^z}{\sqrt{2\pi z}}, \quad (2.30)$$

which holds for  $|z|$  large. Taking the absolute value

$$|R_h(\tau)| \simeq e^{-\frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2}} \frac{1}{|\sqrt{1 - j2\pi f_D \tau / k_r}|} \quad (2.31)$$

$$= e^{-\frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2}} \frac{1}{(1 + (2\pi f_D \tau / k_r)^2)^{1/4}}. \quad (2.32)$$

Following the definition  $|R_h(T_c)| = R$ , we solve for  $T_c$  as

$$1 + (2\pi f_D T_c / k_r)^2 = \frac{1}{R^4} e^{-\frac{k_r f_D^2 T_c^2 \sin^2 \mu_r}{2D_{r,\lambda}^2}} \quad (2.33)$$

$$\simeq \frac{1}{R^4} \left( 1 - \frac{k_r f_D^2 T_c^2 \sin^2 \mu_r}{2D_{r,\lambda}^2} \right) \quad (2.34)$$

$$\Rightarrow T_c(k_r) = \sqrt{\frac{1/R^4 - 1}{(2\pi f_D / k_r)^2 + \frac{k_r f_D^2 \sin^2 \mu_r}{2D_{r,\lambda}^2 R^4}}} \quad (2.35)$$

where we have used the approximation  $e^z \simeq 1 + z$  to eliminate the exponential term. For small beamwidth, we have  $k_r \simeq 1/\theta^2$ , and thus we can express the channel coherence time as a function of the beamwidth as

$$T_c(\theta) = \frac{1}{f_D} \sqrt{\frac{1/R^4 - 1}{(2\pi)^2\theta^4 + \frac{1}{2\theta^2 R^4} \left(\frac{\sin \mu_r}{D_{r,\lambda}}\right)^2}}. \quad (2.36)$$

When  $D_{r,\lambda} \rightarrow \infty$ , i.e., ignoring the pointing error due to the receiver movement, the coherence time simplifies to

$$T_c(\theta) = \frac{\sqrt{1/R^4 - 1}}{2\pi f_D \theta^2}. \quad (2.37)$$

This shows that the coherence time is proportional to  $1/\theta^2$ .

#### 2.5.1.2 When $|\mu_r|$ is not small

The approach here is different from the previous case. First, we compute the argument of the Bessel function, and then we apply the asymptotic approximation (2.30). Taking the log of the obtained equation, we get a polynomial equation of  $\tau$ . The exact solution is not trivial, but considering the range of values of the parameters, higher order terms are negligible and we can approximately solve a quadratic equation instead.

Defining  $c + jd = \sqrt{x^2 + y^2}$  where  $x$  and  $y$  are given in (2.21) and (2.22).

With some algebra,  $c$  and  $d$  can be derived as

$$c = \sqrt{\frac{\sqrt{a^2 + b^2} + a}{2}}, \quad d = \frac{b}{2c}, \quad (2.38)$$



where  $a = k_r^2 - (2\pi f_D)^2 \tau^2$  and  $b = -4\pi f_D k_r \cos(\mu_r) \tau$ . Substitute  $c$  and  $d$  into (2.20), then apply the asymptotic approximation (2.30), and finally take the absolute value, we have

$$R = e^{-\frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2}} \frac{e^{c-k_r}}{(1 + (b/2k_r^2)^2)^{1/4}}. \quad (2.39)$$

For large  $k_r$ , the denominator takes values close to one, and we approximate  $(1 + (b/2k_r^2)^2)^{1/4} \simeq 1$ . Taking the log on both sides and rearranging to get

$$k_r + \log R + \frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2} = \sqrt{\frac{\sqrt{a^2 + b^2} + a}{2}}.$$

Now taking the square of both sides and ignore the  $\tau^4$  term,

$$2(k_r + \log R)^2 + 4(k_r + \log R) \frac{k_r f_D^2 \tau^2 \sin^2 \mu_r}{8D_{r,\lambda}^2} - a = \sqrt{a^2 + b^2}.$$

Once again take the square of both sides, and neglect the higher order terms with respect to  $\tau$ . Then, substitute  $a$  and  $b$  we obtain (2.40), from which the approximate channel coherence time expression (2.41) can be readily derived by solving a quadratic equation in  $\tau$ .

$$\begin{aligned} &4(k_r + \log R)^4 - 4k_r^2(k_r + \log R)^2 + \\ &\left[16(k_r + \log R)^3 - 8k_r^2(k_r + \log R)\right] \frac{k_r f_D^2 \sin^2 \mu_r}{8D_{r,\lambda}^2} \tau^2 \\ &+ 4(k_r + \log R)^2 (2\pi f_D)^2 \tau^2 = (4\pi f_D k_r \cos \mu_r)^2 \tau^2 \end{aligned} \quad (2.40)$$

$$T_c^2(k_r) = \frac{(k_r^2 - (k_r + \log R)^2)/f_D^2}{\left[4(k_r + \log R) - \frac{2k_r^2}{k_r + \log R}\right] \frac{k_r \sin^2 \mu_r}{8D_{r,\lambda}^2} + (2\pi)^2 - \frac{(4\pi k_r \cos \mu_r)^2}{(k_r + \log R)^2}}. \quad (2.41)$$

Note that for a fixed  $\mu_r$ , the denominator in (2.41) can be negative leading to an invalid solution. The range of valid solutions increases with  $\mu_r$  as will

be shown in our numerical example. As evident from Fig. 2.6, if  $\mu_r$  is not too small, our result covers most of the beamwidths of interest for mmWave systems.

When  $\mu_r = 90^\circ$ , which is the fastest fading case, (2.41) can be simplified using  $4(k_r + \log R) - 2\frac{k_r^2}{k_r + \log R} \simeq 2(k_r + \log R)$ , which is valid for large  $k_r$ . Finally, substituting  $k_r = \frac{1}{\theta^2}$ , the worst case channel coherence time can be expressed as

$$T_c(\theta) = \sqrt{\frac{1 - (1 + \theta^2 \log R)^2}{\frac{1}{4}(1 + \theta^2 \log R) \left(\frac{f_D \sin \mu_r}{D_{r,\lambda}}\right)^2 + (2\pi f_D)^2 \theta^4}}. \quad (2.42)$$

When  $D_{r,\lambda} \rightarrow \infty$ , this further simplifies to

$$T_c(\theta) = \frac{1}{2\pi f_D} \sqrt{\frac{1}{\theta^2} \log \frac{1}{R^2} - (\log R)^2}. \quad (2.43)$$

Using the approximation  $\sqrt{1+z} \simeq 1 + \frac{1}{2}z$  for small  $z$ , it can be shown that  $T_c(\theta)$  increases on the order of  $1/\theta$  for small  $\theta$  at the pointing angle  $\mu_r = 90^\circ$ .

## 2.5.2 LOS Channel

When the LOS dominates,  $K/(K+1) \rightarrow 1$  and  $R_h(\tau) \simeq R_{\text{los}}(\tau)$ . Thus, we have

$$|R_h(\tau)| = e^{\frac{1}{2}k_r(\cos(\Delta_\mu^{\text{los}}) - 1)}. \quad (2.44)$$

Using (2.8) and setting  $|R_h(T_c)| = R$ , we can solve (2.44) to get

$$T_c(\theta) = \frac{D_\lambda}{f_D \sin(\alpha_{\text{los}})} \cos^{-1} \left( \frac{2}{k_r} \log R + 1 \right) \quad (2.45)$$

$$= \frac{D_\lambda}{f_D \sin(\alpha_{\text{los}})} \cos^{-1} (2\theta^2 \log R + 1). \quad (2.46)$$

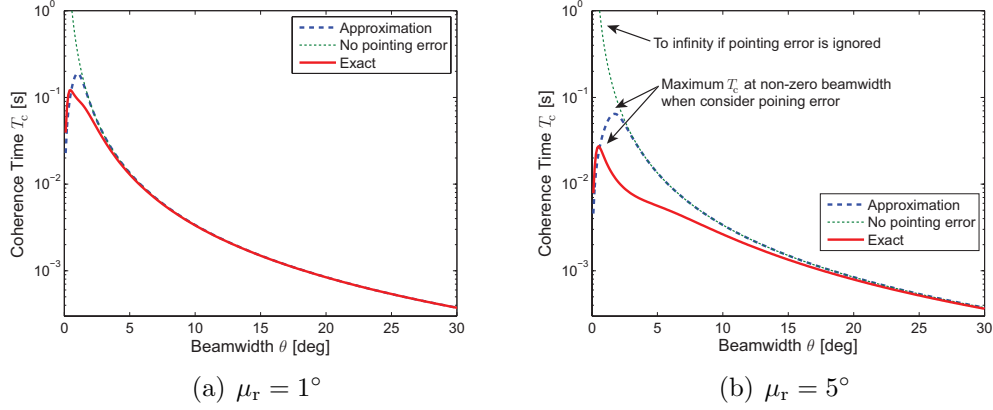


Figure 2.5: Channel coherence time for small  $\mu_r$ . “Exact” refers to the coherence time obtained numerically from the exact correlation function (2.17). “Approximation” and “No pointing error” refer to (2.36) and (2.37). The result is quite sensitive to  $\mu_r$ , and the approximation does not work well for small  $\theta$  but still can capture the effect of the receiver motion.

For this expression to be meaningful,  $\frac{2}{k_r} \log R + 1 \in [-1, 1]$  must hold. Within the typical range of  $R \in [0.3, 1]$ ,  $\frac{2}{k_r} \log R + 1 \in [-1, 1]$  is true for all  $k_r > 2$ . Since we are interested in large  $k_r$ , this constraint presents no limitation here.

### 2.5.3 Numerical Results

We will provide numerical results to verify the derivation for the NLOS case. The derivation for the LOS case does not include approximation and thus no verification is given here. The receiver speed  $v$ , the carrier frequency  $f_c$ , and the scattering radius  $D_{r,\lambda}$  in Table 2.1 are used. The target correlation is set to  $R = 0.5$ . As the approximations depend on  $\mu_r$ , we investigate their behavior for different values of  $\mu_r$  in the followings.

For the small  $|\mu_r|$  case, to see the sensitivity of the approximation in

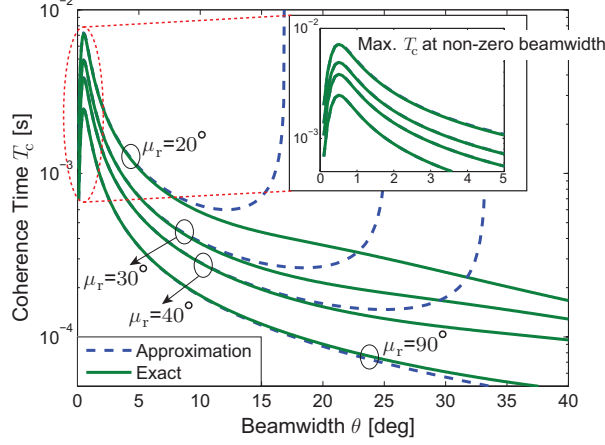


Figure 2.6: Channel coherence time for difference  $\mu_r$ . “Exact” refers to the numerical solution to the exact expression (2.17), and “Approximation” refers to (2.41). We see that the range of valid approximation increases with  $\mu_r$ .

(2.36) we plot the expression and compare it with that of the exact solution for  $\mu_r = 1^\circ$  and  $\mu_r = 5^\circ$ . The exact solution is obtained numerically using the exact correlation function (2.17). The “Approximation” and “No pointing error” refers to the expressions in (2.36) and (2.37), respectively. As mentioned in the derivation, for a given  $\mu_r$  the approximation does not work well for  $\theta$  too small. This error becomes more severe when  $|\mu_r|$  gets larger, which can be seen by comparing the plots in Fig. 2.5(a) and Fig. 2.5(b). Nevertheless, the result can still capture the effect of the pointing error.

The same study is done for the case when  $\mu_r$  is not small. Fig. 2.6 shows the results for four different values of  $\mu_r$ . For a fixed value of  $\mu_r$ , there is a point where the approximation diverges due to the singularity of the denominator in (2.41). We observe that the range of  $\theta$  for valid approximation increases

with  $\mu_r$ , and it is valid up to around  $\mu_r/2$ , i.e., half the pointing angle. Since mmWave systems require narrow beams to compensate for the high path loss, the approximation in (2.41) will be valid for most cases of interest in practice.

Having verified the correctness of our derived expressions, we now summarize the impact of relevant parameters on  $T_c$ , which include the speed  $v$ , the beamwidth  $\theta$ , and the pointing angle  $\mu_r$  (or  $\alpha_{\text{los}}$  for LOS case). As can be seen from (2.36), (2.41), and (2.46), the coherence time is inversely proportional to the maximum Doppler shift  $f_D$ . Since  $f_D$  is proportional to  $v$ ,  $T_c$  is proportional to  $1/v$ , i.e., decreases as  $v$  increases as expected. The behavior of  $T_c$  for a fixed speed is shown in Fig. 2.5 and Fig. 2.6. It can be shown that for all  $\mu_r \neq 0$ , the coherence time  $T_c$  approaches zero as the beamwidth  $\theta$  goes to zero. This is due to the effect of the pointing error. As evident from the plots, the coherence time  $T_c$  attains its maximum at some small but *non-zero*  $\theta_{\text{max}}$ . This  $\theta_{\text{max}}$  depends on  $D_{r,\lambda}$ , and it gets smaller as  $D_{r,\lambda}$  increases. For a fixed  $\theta$ , the coherence time  $T_c$  increases as  $\mu_r$  approaches zero, and it decreases as  $\mu_r$  approaches  $90^\circ$ . This agrees with the result in [33] where it is observed that fading becomes faster as  $\mu_r$  approaches  $90^\circ$ . Finally, note that the result based on the Clarke-Jakes PAS ignores the effect of the beamwidth and the results in [27, 33] suggest that  $T_c$  goes to infinity as  $\theta$  approaches zero because they do not consider the pointing error.

## 2.6 Beam Coherence Time

In this section, we first define the beam coherence time and then derive the beam coherence time expressions for the LOS and NLOS cases. The NLOS case uses our spatial lobe model described in Section 2.3.3.

We define the beam coherence time as the average time over which the beam stays aligned. We focus on only the receive beam here. For a given receive beamwidth, the beam is said to become misaligned when the received power falls below a certain ratio  $\zeta \in [0, 1]$  compared to the peak received power. Let the receive beam be pointing at the peak direction  $\mu_r$  at time  $t$ , then we can define the beam coherence time by

$$T_B = \inf_{\tau} \left\{ \tau \left| \frac{P(t + \tau)}{P(t)} < \zeta \right. \right\}. \quad (2.47)$$

Note that the power decrease here is due to the pointing error  $\Delta_{\mu}(\tau)$  as defined in Section 2.3.2. Using this definition, we derive the beam coherence times for the LOS and NLOS cases in the following.

### 2.6.1 LOS Case

Let the beam pattern be represented by the von Mises distribution as earlier. The received power is proportional to the receive beam pattern, i.e.,  $P(t) \propto G(\mu_r|\mu_r)$  and  $P(t + \tau) \propto G(\mu_r|\mu_r + \Delta_{\mu}^{\text{los}})$ . At  $\tau = T_B$ ,

$$\zeta = \frac{G(\mu_r|\mu_r + \Delta_{\mu}^{\text{los}})}{G(\mu_r|\mu_r)} \quad (2.48)$$

$$= e^{k_r(\cos(\Delta_{\mu}^{\text{los}}) - 1)}. \quad (2.49)$$

Substituting  $\Delta_\mu^{\text{los}}$  from (2.8), (2.49) can be solved to get

$$T_B(\theta) = \frac{D_\lambda}{f_D \sin \mu_r} \cos^{-1}(\theta^2 \log \zeta + 1), \quad (2.50)$$

where we have used  $k_r = 1/\theta^2$  in (2.50). Note the similarity to the channel coherence time for the LOS case. This is because the main cause for the temporal variation in the LOS case is the pointing error.

### 2.6.2 NLOS Case

First, we need to determine the pointing error caused by the receiver motion to compute the beam coherence time. In the NLOS case, the incoming power is the result of the reflection from the scatterers. Following our one-ring scatterer model, the pointing error is given by (2.7).

We now compute the received power to solve for  $T_B$ . As mentioned in Section 2.3, following the 5G channel model in [101], we model the spatial lobe by a Gaussian PDF. Recall that the spatial lobe is the PAS  $\mathcal{P}'(\alpha|\mu_r)$  before applying the receive beam pattern. The variance of the Gaussian PDF is  $\beta^2$ , defined in (2.9), and the mean is  $\mu_r$ . At time  $t$ , we assume perfect alignment so that the receive beam is pointing at the peak of  $\mathcal{P}'(\alpha|\mu_r)$ , i.e., using the beam pattern  $G(\alpha|\mu_r)$ . At time  $t + \tau$ , the beam pattern now changes to  $G(\alpha|\mu_r + \Delta_\mu)$  if no realignment is done. The received power at the pointing angle  $\mu_r + \Delta_\mu$  is

$$P(t + \tau) = \int_0^{2\pi} \mathcal{P}'(\alpha|\mu_r) G(\alpha|\mu_r + \Delta_\mu) d\alpha. \quad (2.51)$$

Note that for large  $k_r$ , the von Mises PDF in (2.2) approaches the Gaussian PDF [37, Ch. 45]. Also note that for large  $k_r$ , i.e., small variance, the dis-

tribution falls off fast and the tails at both sides beyond 0 and  $2\pi$  have little weight. These observations lead to the following approximation:

$$\begin{aligned} P(t + \tau) &\simeq \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{(\alpha - \mu_r)^2}{2\beta^2}} \frac{1}{\sqrt{2\pi\theta^2}} e^{-\frac{(\alpha - \mu_r - \Delta_\mu)^2}{2\theta^2}} d\alpha \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{u^2}{2\beta^2}} \frac{1}{\sqrt{2\pi\theta^2}} e^{-\frac{(\Delta_\mu - u)^2}{2\theta^2}} du. \end{aligned} \quad (2.52)$$

Applying a change of variable  $u = \alpha - \mu_r$ ,  $\mu_r$  can be eliminated from the first expression. The expression (2.52) is just a convolution between two Gaussian PDFs, which is well-known to result in another Gaussian PDF with mean  $\Delta_\mu$  and variance  $\beta^2 + \theta^2$  [115]. That is,

$$P(t + \tau) \simeq \frac{1}{\sqrt{2\pi(\beta^2 + \theta^2)}} e^{-\frac{\Delta_\mu^2}{2(\beta^2 + \theta^2)}}, \quad (2.53)$$

which does not depend on  $\mu_r$ . This makes sense because in the current setting it is assumed that at time  $t$  the receive beam points at  $\mu_r$  and  $P(t + \tau)$  is determined solely from the misalignment that happens at time  $t + \tau$ . This misalignment is captured by the pointing error due to the receiver motion  $\Delta_\mu$ , which is a function of  $\tau$ . We can solve for  $T_B$  directly from (2.53); however, by approximating (2.53) by a von Mises PDF, the resulting  $T_B$  is of the same form for both the LOS and NLOS cases. Using the approximation, (2.53) becomes

$$P(t + \tau) \simeq \frac{1}{2\pi I_0(1/(\beta^2 + \theta^2))} e^{\frac{\cos(\Delta_\mu)}{\beta^2 + \theta^2}}. \quad (2.54)$$

With the same steps used in the derivation in the LOS case, we get the expression for the NLOS case as

$$T_B(\theta|\beta) = \frac{D_{r,\lambda}}{f_D \sin \mu_r} \cos^{-1}((\beta^2 + \theta^2) \log \zeta + 1). \quad (2.55)$$



Note that  $\beta$  is a random variable and is modeled by the Gaussian distribution in (2.9). Thus, to get the beam coherence time we need to average over  $\beta$ :

$$T_B(\theta) = \mathbb{E}_\beta[T_B(\theta|\beta)], \quad (2.56)$$

where  $\mathbb{E}_\beta[\cdot]$  denotes the statistical expectation over  $\beta$ . The difference to the LOS case is that now  $T_B$  depends on the channel through the spatial lobe angular spread parameter  $\beta$ .

### 2.6.3 Numerical Results

We first state the simulation parameters. We set the transmitter-receiver distance  $D = 50$  m, the power ratio threshold  $\zeta = 0.5$ , and use the parameters in Table 2.1.  $\sigma_{AS}$  and  $D_{r,\lambda}$  are used only for the NLOS case.

Fig. 2.7 shows the beam coherence time for  $\mu_r = 10^\circ$  and  $\mu_r = 80^\circ$ . In the LOS case, it looks like  $T_B$  is linear with respect to  $\theta$ . It is almost linear because the argument to the  $\cos^{-1}(\cdot)$  is of the form  $1 + z^2$ , which happens to be the first order Taylor approximation of  $\cos(\cdot)$ . For the same traveled distance, the pointing direction changes less for small  $\mu_r$  which results in larger  $T_B$  for  $\mu_r = 10^\circ$ . For the NLOS case, recall that we model the angular spread of a spatial lobe as Gaussian distributed with a standard deviation  $\sigma_{AS} = 25.7^\circ$ . For the mean angle of arrival of  $80^\circ$ , we see from Fig. 2.7(b) that increasing the beamwidth does not effectively increase  $T_B$  as in the case when the mean angle of arrival is  $10^\circ$ . Comparing the results in Fig. 2.7(b) and Fig. 2.6 we see that  $T_B$  shown here is much larger than  $T_c$ . Also note that while  $T_B$  increases with the beamwidth,  $T_c$  generally decreases as the beamwidth gets larger.

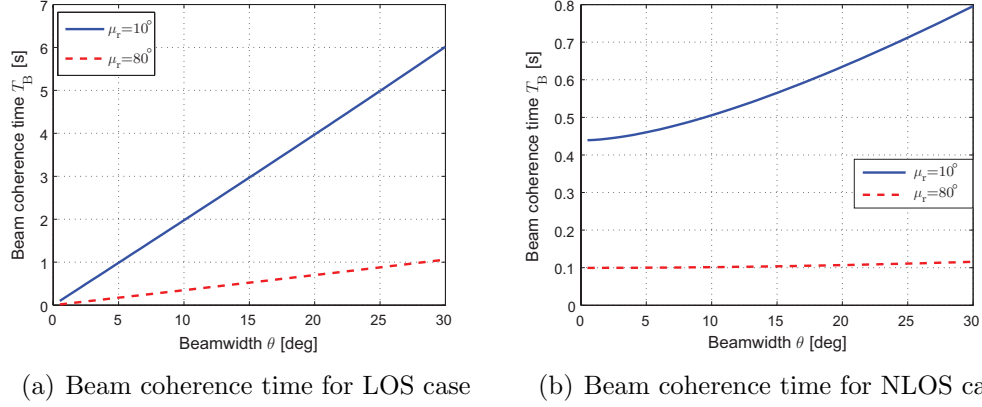


Figure 2.7: The beam coherence time for both the LOS and NLOS cases. The transmitter-receiver distance is set to  $D = 50$  m, and the parameters from Table 2.1 are used. For the LOS case,  $T_B$  increases linearly with the beamwidth  $\theta$ . For the NLOS case, the increase is slower. This is because of the effect of the spatial lobe width  $\beta$  on  $T_B$  in the NLOS case. For both LOS and NLOS cases,  $T_B$  increases fast with  $\theta$  for  $\mu_r \rightarrow 0^\circ$ , and slowly for  $\mu_r \rightarrow 90^\circ$ .

We now summarize the effect of each parameter on  $T_B$ . For both LOS and NLOS cases,  $T_B \propto 1/f_D$  as evident from (2.50) and (2.55). The impact of  $\mu_r$  is similar to that of  $T_c$ ; for a fixed  $\theta$ ,  $T_B$  increases fast for  $\mu_r$  close to zero and slower for  $\mu_r$  close to  $90^\circ$ . For the LOS case,  $T_B$  increases linearly with  $\theta$ . For the NLOS case, not just  $\theta$  but also the spatial lobe width  $\beta$  matters. Following the argument for the LOS case by treating  $z = \sqrt{\beta^2 + \theta^2}$ , we can say that  $T_B$  linearly increases with  $\sqrt{\beta^2 + \theta^2}$ . This agrees with the intuition that if the incoming energy arrives over a wide range of angles (i.e., a large  $\beta$ ) then the received energy will be the same even with some change in the pointing angle. This translates into a larger  $T_B$ .

## 2.7 Implications on Beam Realignment Duration

So far we have defined and derived two coherence times: the channel coherence time  $T_c$  and the beam coherence time  $T_B$ , relevant to scenarios where mmWave directional beams are used in vehicular environments.  $T_c$  determines how fast the channel coefficient changes in time, and thus can be used in deciding the packet length and determining the overhead for channel estimation. We explore some implications on the choice of beamwidth in this section. We show that to maximize the performance, beam realignment should be done in every beam coherence time  $T_B$  and not in every  $T_c$ .

### 2.7.1 Lower Bound on Mutual Information

We use a discrete time channel model to derive a lower bound on the mutual information. The discrete time channel model is obtained by discretizing the continuous time channel model from Section 2.3. Consider the following signal model

$$y[i] = h[i]s[i] + n[i], \quad i = 1, 2, \dots, k \quad (2.57)$$

where  $y[i]$  is the receive signal,  $h[i]$  is the channel,  $s[i]$  is the transmitted signal,  $n[i]$  is the additive white Gaussian noise (AWGN), and  $k$  is the packet length measured in the number of samples. The channel  $h[i]$  is the discrete-time version of the channel  $h(t)$  from Section 2.3. Note that  $h(t)$  is the channel seen through a receive beam  $G(\alpha|\mu_r)$  with some pointing direction  $\mu_r$ . To

discretize  $h(t)$ , we use an autoregressive model

$$h[i] = \alpha h[i-1] + \xi[i-1]. \quad (2.58)$$

Here,  $\alpha$  is the correlation coefficient and is determined from the temporal correlation function derived in Section 2.4, i.e.,

$$\alpha = R_h(\nu T) \quad (2.59)$$

with the symbol duration  $T$ ,  $\xi[i]$  is the innovation term with variance  $\sigma_\xi^2 = 1 - |R_h(\nu T)|^2$ .

For decoding, the channel has to be estimated, and the effects from both the thermal noise and the channel time-variation need to be considered. If the estimator does not have knowledge of the statistics of  $\xi[i]$  (which typically is the case), then a natural assumption is that  $\xi[i]$  is Gaussian. Because (2.58) is a Gauss-Markov channel model, following the logic used in [75], the Kalman filter provides the maximum-likelihood estimate (and also the minimum mean squared error estimate) [13]. Suppose the channel is estimated with the help of pilot symbols equally spaced in every  $\nu$  samples. The received pilot signal vector can then be written as  $\mathbf{v}_{\lfloor k/\nu \rfloor}$ , where  $\lfloor \cdot \rfloor$  denotes the floor function and  $\mathbf{a}_k$  denotes a vector of length  $k$ . Applying the Kalman filter with the pilot  $\mathbf{v}_{\lfloor k/\nu \rfloor}$  as the measurement vector, the variance of the channel estimation error at the  $\ell$ -th pilot  $\psi_\ell$  is given by the following recursive relations [75]

$$\frac{1}{\psi_1} = \frac{1}{\sigma_h^2 + \sigma_\xi^2} + \frac{\sigma_v^2}{\sigma_n^2} \quad (2.60)$$

$$\frac{1}{\psi_{\ell+1}} = \frac{1}{\alpha^2 \psi_\ell + \sigma_\xi^2} + \frac{\sigma_v^2}{\sigma_n^2}, \quad (2.61)$$

where  $\sigma_h^2$  is the channel power,  $\sigma_v^2$  is the pilot signal power assumed to be the same for all pilot symbols, and  $\sigma_n^2$  is the noise power.

To explicitly express the channel estimation error, the channel is decomposed into the known part  $\bar{h}[i]$  and the estimation error  $\tilde{h}[i]$ , i.e.,

$$h[i] = \bar{h}[i] + \tilde{h}[i], \quad (2.62)$$

The variance of the known part  $\bar{h}[i]$  can be written as

$$\sigma_{\bar{h}}^2[i] = \sigma_h^2 - \sigma_{\tilde{h}}^2[i]. \quad (2.63)$$

This notation is used in the derivation of the lower bound below. Note that the estimation error variances given in (2.60) and (2.61) are at the sampling points corresponding to the pilots. When they are used to decode the data part, the channel time variation will further degrade the estimation accuracy. This increase in estimation error is determined from the channel correlation function and the total estimation error variance at a given sampling point can be written as

$$\sigma_{\tilde{h}}^2[i] = \psi_{\lfloor i/\nu \rfloor} + (1 - |R_h((i - \lfloor i/\nu \rfloor \nu)T)|^2). \quad (2.64)$$

For a very long sequence of signals, i.e., when  $k \rightarrow \infty$ , the error variance from the Kalman filter converges to some value  $\psi$  (not dependent on the pilot index) given by [75]

$$\begin{aligned} \psi = & \frac{|R_h(\nu T)|^2 - 1 - \text{SNR}_v G_a(\theta) \sigma_\xi^2}{2 \text{SNR}_v G_a(\theta) |R_h(\nu T)|^2} \\ & + \frac{\sqrt{(|R_h(\nu T)|^2 - 1 - \text{SNR}_v G_a(\theta) \sigma_\xi^2)^2 + 4 \text{SNR}_v^2 G_a(\theta) \sigma_\xi^2 |R_h(\nu T)|^2}}{2 \text{SNR}_v G_a(\theta) |R_h(\nu T)|^2}, \end{aligned} \quad (2.65)$$

where  $\text{SNR}_v = \sigma_v^2/\sigma_n^2$  is the SNR of the pilot symbol excluding the antenna gain.  $G_a(\theta)$  is the antenna gain compared to omnidirectional antenna and is given by

$$G_a(\theta) = \frac{G(\mu_r|\mu_r)}{1/(2\pi)} = \frac{e^{1/\theta^2}}{I_0(1/\theta^2)} \quad (2.66)$$

where  $1/(2\pi)$  in the denominator is the gain of an omnidirectional antenna, and  $G(\mu_r|\mu_r)$  is the peak of the antenna pattern with the main beam pointing at  $\mu_r$ .  $G(\alpha|\mu_r)$  is assumed to have the shape of the von Mises PDF for consistency with the rest of the chapter. Note that we use the peak of the antenna pattern here because the time scale of a packet is small and there will be negligible variation in the pointing direction within one packet.

Now, consider the mutual information for only the  $i$ -th sample with channel estimate with the error given in (2.65). The worst case that the error  $\tilde{h}[i]$  can have is to act as AWGN [75]. In that case, the mutual information can be lower bounded by

$$I(s[i]; y[i] | \mathbf{v}_{[i/\nu]}) \geq \ln \left( 1 + \frac{\sigma_h^2[i] \sigma_s^2}{\sigma_h^2 \sigma_s^2 + \sigma_n^2} \right). \quad (2.67)$$

Using (2.67), and assuming the estimator does not use the decoded data for channel estimation and only use the pilot  $\mathbf{v}_{[i/\nu]}$  then it can be shown that [75]

$$I(\mathbf{s}_k; (\mathbf{y}_k, \mathbf{v}_{[k/\nu]})) \geq \sum_{i \leq k} I(s[i]; y[i] | \mathbf{v}_{[i/\nu]}). \quad (2.68)$$

Plugging in the result so far, a lower bound for the mutual information can be

written as

$$I(\mathbf{s}_k; (\mathbf{y}_k, \mathbf{v}_{\lfloor k/\nu \rfloor})) \geq \sum_{i \leq k} \ln \left( 1 + \frac{(|R_h((i - \lfloor i/\nu \rfloor)\nu)T|^2 - \psi_{\lfloor i/\nu \rfloor})\text{SNR}_s G_a(\theta)}{(\psi_{\lfloor i/\nu \rfloor} + (1 - |R_h((i - \lfloor i/\nu \rfloor)\nu)T|^2))\text{SNR}_s G_a(\theta) + 1} \right), \quad (2.69)$$

where  $\text{SNR}_s = \sigma_s^2/\sigma_n^2$  is the SNR of the data part excluding the antenna gain.

Further, assume  $k \rightarrow \infty$ , then  $\psi_{\lfloor i/k \rfloor} \rightarrow \psi$ , and we have

$$\lim_{k \rightarrow \infty} \frac{1}{k} I(\mathbf{s}_k; (\mathbf{y}_k, \mathbf{v}_{\lfloor k/\nu \rfloor})) \geq \frac{1}{\nu} \sum_{i=2}^{\nu} \ln \left( 1 + \frac{(|R_h(iT)|^2 - \psi)\text{SNR}_s G_a(\theta)}{(\psi + (1 - |R_h(iT)|^2))\text{SNR}_s G_a(\theta) + 1} \right) \quad (2.70)$$

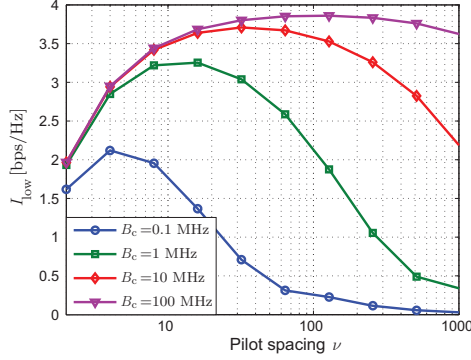
$$= I_{\text{low}}(\theta, \text{SNR}_s, \nu). \quad (2.71)$$

At high SNR or when beamwidth  $\theta$  is small (i.e., the antenna gain  $G_a(\theta)$  is large), then

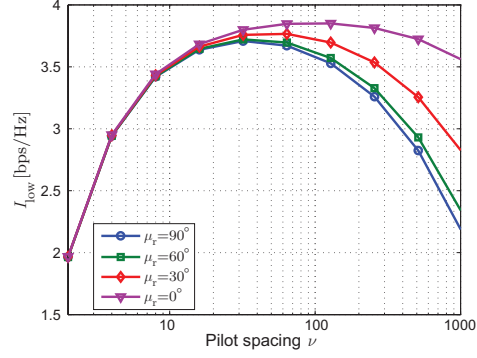
$$I_{\text{low}}(\theta, \text{SNR}_s, \nu) \simeq \frac{1}{\nu} \sum_{i=2}^{\nu} \ln \left( 1 + \frac{|R_h(iT)|^2 - \psi}{\psi + (1 - |R_h(iT)|^2)} \right), \quad (2.72)$$

which implies that the loss due to the channel time-variation acts in the same manner as the interference, and it cannot be mitigated by increasing the transmit power.

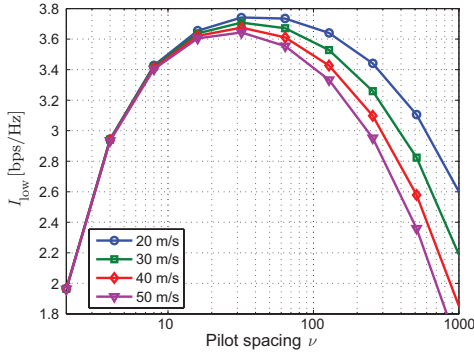
Fig. 2.8 provides some insights into the effects of different parameters on  $I_{\text{low}}$ . In all cases,  $\text{SNR}_s = \text{SNR}_v = 0$  dB and  $D_{r,\lambda} = 1000$ . Except indicated otherwise in the plots,  $B_c = 10$  MHz,  $v = 30$  m/s,  $\mu_r = 90^\circ$ , and  $\theta = 10^\circ$ . A smaller symbol duration  $T$  will result in lower overhead for estimating the time-varying channel. To ensure frequency flat channels,  $T$  is constrained by  $T \geq 1/B_c$ , with  $B_c$  the coherence bandwidth. With the choice of  $T = 1/B_c$ ,



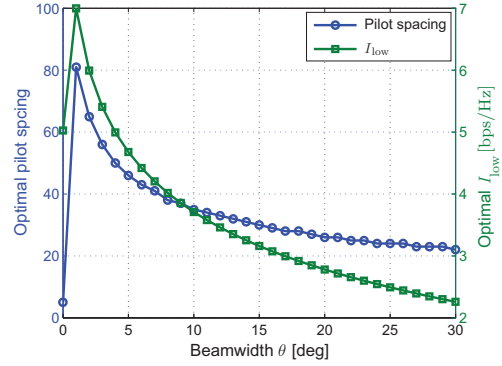
(a)  $I_{\text{low}}$  for different coherence bandwidths



(b)  $I_{\text{low}}$  for different  $\mu_r$



(c)  $I_{\text{low}}$  for different speed  $v$



(d) Optimal pilot spacing vs. beamwidth  $\theta$

Figure 2.8: Effects of different parameters on  $I_{\text{low}}$ . In all cases,  $\text{SNR}_s = \text{SNR}_v = 0$  dB and  $D_{r,\lambda} = 1000$  are used. Except indicated otherwise in the legends,  $B_c = 10$  MHz,  $v = 30$  m/s,  $\mu_r = 90^\circ$ , and  $\theta = 10^\circ$  are used. Fig. 2.8(a) shows  $I_{\text{low}}$  against pilot spacing for different  $B_c$ . The optimal  $\nu$  increases as  $B_c$  gets larger. Fig. 2.8(b) shows  $I_{\text{low}}$  for different  $\mu_r$ . The optimal  $\nu$  decreases as  $\mu_r$  approaches  $90^\circ$ . This is because the fading changes faster for  $\mu_r$  near  $90^\circ$  [33]. Fig. 2.8(c) shows  $I_{\text{low}}$  for different speeds. As the speed increases, the time-variation increases and the pilot spacing should be decreased. Fig. 2.8(d) shows the optimal  $\nu$  and the corresponding  $I_{\text{low}}$  against the beamwidth  $\theta$ . Both the optimal  $\nu$  and  $I_{\text{low}}$  decrease for a too small or too large  $\theta$ . This is because  $\theta$  that is too small suffers from the pointing error, and  $\theta$  that is too large suffers from the Doppler spread which results in fast time variation (small  $T_c$ ).



Fig. 2.8(a) shows the effect of  $B_c$  on  $I_{\text{low}}$ . We see that the optimal pilot spacing increases with  $B_c$ . For a small  $B_c$ , the symbol duration is large and the time-variation between two consecutive symbols increases. This means that the pilot spacing should be set to a smaller value for a smaller  $B_c$  to suppress channel estimation errors due to the time-variation of the channel.

Fig. 2.8(b) shows  $I_{\text{low}}$  against the pilot spacing for different  $\mu_r$ . The optimal pilot spacing increases as  $\mu_r$  decreases from  $90^\circ$  to  $0^\circ$ . This is because the channel changes faster when  $\mu_r$  approaches  $90^\circ$  and slower when  $\mu_r$  approaches  $0^\circ$  as we have seen in the discussion on the temporal correlation function in Section 2.4 and the channel coherence time in Section 2.5.

Fig. 2.8(c) shows  $I_{\text{low}}$  against the pilot spacing for different speed  $v$ . The faster the speed, the larger the channel time-variation and thus we expect the optimal pilot spacing to decrease as the speed increases. This trend can be confirmed in Fig. 2.8(c). The difference in the optimal  $I_{\text{low}}$ , however, is rather small. This suggests that for typical highway speeds, there is no need to adapt the pilot spacing to the speed of the vehicle.

Finally, Fig. 2.8(d) shows the optimal pilot spacing that maximizes  $I_{\text{low}}$  against the beamwidth  $\theta$  and the corresponding maximum  $I_{\text{low}}$ . Since a larger  $\theta$  leads to a smaller  $T_c$  as shown in Section 2.5, the optimal pilot spacing decreases with  $\theta$ . When  $\theta$  becomes too small, due to the effect of pointing error, the channel coherence time decreases and so the optimal pilot spacing also decreases. Since a smaller pilot spacing means higher overhead, the resulting  $I_{\text{low}}$  follows a similar trend. Notice that with the same optimal

pilot spacing,  $I_{\text{low}}$  for narrower beams has a higher value because of the higher antenna gains for narrower beams. The results in Fig. 2.8(d) suggest that the beam should be pointy but it should not be too pointy.

### 2.7.2 How Often Should the Beams Be Realigned?

In this subsection, we investigate the choice of the time duration between beam realignments. We consider the beam sweeping as a method to align the beams. Two possible choices for the time duration between realignments are the channel coherence time  $T_c$  (Section 2.5) and the beam coherence time  $T_B$  (Section 2.6). Assuming no error in the beam measurement during the alignment process, realignment in every  $T_c$  will ensure that the best beams, which provide the highest received power, are always chosen. If realignment is done in every  $T_B$  instead, suboptimal beams could result due to the effect of fading. The overhead is, of course, higher when realigning in every  $T_c$  than when realigning in every  $T_B$  because  $T_B \geq T_c$ . We call the realignment in every  $T_c$  the short-term realignment and the realignment in every  $T_B$  the long-term realignment. In the following, we will investigate the performance of these two cases. For the LOS channel,  $T_c$  and  $T_B$  are of comparable values (see (2.46) and (2.50)), and there is not much difference between the two. Therefore, we only study the NLOS case here.

For clarity, we consider a two-spatial-lobe channel similar to the one illustrated in Fig. 2.2 (note that four lobes are shown in the figure). Each spatial lobe corresponds to a scattering cluster that has a certain path loss

and angular spread (i.e., the lobe width). We assume that all the spatial lobes have the same fading statistic, and Rayleigh is assumed. There are two main effects of fading here. One is the probability of choosing the suboptimal spatial lobe (in the long-term beam realignment case) and the other is the calculation of the average  $I_{\text{low}}$ . The former effect is the one that could alter the conclusion of whether short- or long-term beam realignment performs better. For less severe fading, the probability of the suboptimal choice of the spatial lobe will be lower and the long-term beam alignment will perform better. Therefore, if we can show that the long-term beam alignment performs better for Rayleigh fading, the conclusion will hold for less severe fading which is expected in mmWave systems because the use of narrow beams will limit the multipath. Also, note that extension to other fading distribution is straightforward as long as the PDF of the SNR of the short-term beam alignment exists. The assumption of two spatial lobes is to simplify the analysis of the wrong choice of the spatial lobe (i.e., choosing a lobe with higher path loss) during the beam training due to fading. The two-spatial-lobe model can capture the power loss due to this wrong choice. More spatial lobes can provide more granularity of the power loss, but this can be imitated by varying the path loss ratio of the two lobes in the two-spatial-lobe model.

Denote  $\Gamma \geq 1$  the path loss ratio between the first and second spatial lobe,  $\text{PL}_i$  for  $i = \{1, 2\}$  the path losses of the two spatial lobes, then

$$\text{PL}_1 = \text{PL}_2 \Gamma, \quad (2.73)$$

where we have assumed without loss of generality that the first spatial lobe has a higher average received power. Let  $g_i = |h_i|^2$  and  $P_i$  where  $i = \{1, 2\}$  be the fading and the instantaneous received power, respectively, then we have

$$P_i = g_i \text{PL}_i. \quad (2.74)$$

Note that our channel model in (2.1) corresponds to the fading coefficient and no path loss was incorporated.

The beam sweeping algorithm will select a beam using the rule  $i^* = \arg \max_i P_i$ , explicitly

$$i^* = \begin{cases} 1 & \text{if } g_1 \geq g_2/\Gamma \\ 2 & \text{if } g_1 < g_2/\Gamma \end{cases}. \quad (2.75)$$

Let  $f_g(g)$  be the PDF of  $g_i$ , then the beam sweeping will output 1 and 2 with probabilities

$$\mathbb{P}\{i^* = 1\} = \int_0^\infty \int_{g_2/\Gamma}^\infty f_g(g_1) dg_1 f_g(g_2) dg_2, \quad (2.76)$$

$$\mathbb{P}\{i^* = 2\} = \int_0^\infty \int_0^{g_2/\Gamma} f_g(g_1) dg_1 f_g(g_2) dg_2. \quad (2.77)$$

For tractability, we assume the fading is Rayleigh so that  $g_i$  follows an exponential distribution with unit mean. When realigning in every  $T_c$ , the path yielding the highest power is always chosen, so that the received power follows the distribution of  $\max\{P_1, P_2\}$ . The SNR is proportional to the received power, and the PDF of the SNR can be derived as

$$f_{\text{short}}(\gamma) = \frac{e^{-\gamma/\bar{\gamma}_1}}{\bar{\gamma}_1} + \frac{e^{-\gamma/\bar{\gamma}_2}}{\bar{\gamma}_2} - \frac{\bar{\gamma}_1 + \bar{\gamma}_2}{\bar{\gamma}_1 \bar{\gamma}_2} e^{\frac{(\bar{\gamma}_1 + \bar{\gamma}_2)\gamma}{\bar{\gamma}_1 \bar{\gamma}_2}} \quad (2.78)$$

where  $\bar{\gamma}_i = \frac{P_i}{P_n}$  is the average SNR of the  $i$ -th spatial lobe and  $P_n$  is the noise power. When realigning in every  $T_B$ , the beam sweeping is performed at the beginning and the selected beam will be used until the next realignment. Note that  $T_B \gg T_c$  for the NLOS channels (see numerical examples in Section 2.5 and 2.6). The fading coefficient becomes uncorrelated after  $T_c$ , and thus the beam selected at the beginning could result in a suboptimal instantaneous received power at some point between beam realignment. Depending on the result of the beam sweeping, the channel experienced here follows either  $P_1$  or  $P_2$ . The PDF of the SNR in this case is

$$f_{\text{long}}(\gamma) = \frac{1}{\bar{\gamma}_i} e^{-1/\bar{\gamma}_i}. \quad (2.79)$$

So far we have derived the PDFs of the SNR for the short- and long-term realignment. Now, we will discuss the overhead of the two realignment durations. The time needed for beam sweeping is the same for both the short- and long-term realignments. Denoting this time duration by  $T_{\text{sw}}$ , then the temporal efficiencies of the short- and long-term realignments are

$$\eta_{\text{short}}(\theta) = \frac{T_c(\theta) - T_{\text{sw}}(\theta)}{T_c(\theta)} \quad (2.80)$$

$$\eta_{\text{long}}(\theta) = \frac{T_B(\theta) - T_{\text{sw}}(\theta)}{T_B(\theta)}. \quad (2.81)$$

Note that all these are functions of the beamwidth  $\theta$ .

Finally, the loss due to the channel time-variation, the temporal efficiency, and the bound on the mutual information are all considered for the

overall performance metric, i.e.,

$$C_{\text{short}}(\theta) = \eta_{\text{short}}(\theta) \mathbb{E}_{\text{short}} [I_{\text{low}}(\theta, \gamma_{\text{short}}, \nu)] \quad (2.82)$$

$$C_{\text{long}}(\theta) = \eta_{\text{long}}(\theta) \mathbb{E}_{\text{long}} [I_{\text{low}}(\theta, \gamma_{\text{long}}, \nu)] \quad (2.83)$$

$$\begin{aligned} &= \eta_{\text{long}}(\theta) (\mathbb{P}\{i^* = 1\} \mathbb{E}_{\gamma_1} [I_{\text{low}}(\theta, \gamma_1, \nu)] \\ &\quad + \mathbb{P}\{i^* = 2\} \mathbb{E}_{\gamma_2} [I_{\text{low}}(\theta, \gamma_2, \nu)]) \end{aligned} \quad (2.84)$$

where  $I_{\text{low}}(\theta, \gamma, \nu)$  is the lower bound derived in (2.71) in the previous subsection.

Now, we provide a numerical example comparing the spectral efficiencies in (2.82) and (2.83) when realignment duration is set to  $T_c$  versus  $T_B$ . To make the comparison meaningful, the pilot spacing  $\nu$  should be optimized for all  $\theta$ . This is done numerically, and the obtained optimal pilot spacings follow a similar trend to that of Fig. 2.8(d). For the beam sweeping, we consider a basic approach adopted in the IEEE 802.15.3c [53] which is based on a hierarchical beam codebook. Let  $\ell$  be the number of levels in the codebook, and the  $i$ -level have  $L_i$  beams. In this approach, at each level all the beam combination pairs are tested, so the overhead of beam training is  $L_i^2 T_{\text{TRN}}$  for the search at the  $i$ -th level.  $T_{\text{TRN}}$  is the duration needed for one beam measurement. Thus, the 802.15.3c method has overhead of  $T_{3c} = T_{\text{TRN}} \sum_{i=1}^{1/\ell} L_i^2$ . It can be shown that the optimum  $L_i$  that minimizes the number of beam training is  $L_1 = \dots = L_\ell = L = \left(\frac{\theta_0}{\theta}\right)^{1/\ell}$ , where  $\theta_0$  is the coverage and  $\theta$  is the desired beamwidth. In this case, the overhead becomes

$$T_{3c}(\theta) = \ell \left(\frac{\theta_0}{\theta}\right)^{2/\ell} T_{\text{TRN}}. \quad (2.85)$$

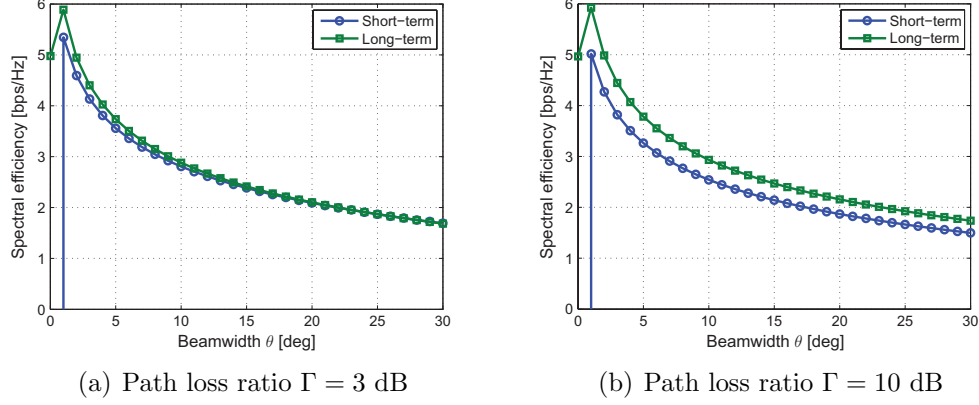


Figure 2.9: Comparison of the spectral efficiencies in (2.82) and (2.83) for the short- and long-term beam realignment when the beam sweeping follows the 802.15.3c method. Fig. (a) and (b) show the case when the path loss ratio of the two spatial lobes  $\Gamma$  is 3 dB and 10 dB, respectively. In both cases, the long-term realignment performs better and the gap is more pronounced when  $\Gamma = 10$  dB. The gap increases for larger  $\Gamma$  because the sweeping is less likely to make mistake when  $\Gamma$  is large so that the large overhead of the short-term realignment penalizes rather than improves the performance.

Note that the overhead here ignores the feedback and acknowledgment phase. Plugging in  $T_{\text{SW}}(\theta) = T_{3c}(\theta)$ , we can now compute the spectral efficiencies in (2.82) and (2.83) as a function of the beamwidth  $\theta$ . The coherence bandwidth is set to 10 MHz, pointing angle  $\mu_r = 90^\circ$  (which corresponds to the worst case),  $\theta_0 = 180^\circ$ , the training per beam  $T_{\text{TRN}} = 1 \mu\text{s}$ , and angular spread  $\sigma_{\text{AS}} = 25.7^\circ$ . Other parameters are the same as used in the previous subsection.

The result is shown in Fig. 2.9 for the case when the path loss ratio  $\Gamma$  is 3 dB and 10 dB. In both cases, the long-term realignment has higher spectral efficiency and the gap is larger for large  $\Gamma$ . This is because when  $\Gamma$  is large, the probability that beam sweeping chooses the suboptimal choice

becomes smaller so that minimal benefit can be expected from the short-term realignment. Thus, the overhead paid for the short-term realignment does not provide sufficient return and the long-term realignment performs better due to the lower required overhead.

## 2.8 Conclusion

In this chapter, we derived the channel coherence time for a wireless channel as a function of the beamwidth taking both Doppler effect and pointing error into consideration. Our results show that there exists a non-zero optimal beamwidth that maximizes the channel coherence time. If the beamwidth is too narrow, pointing error will limit the coherence time. If the beamwidth is too wide, the Doppler spread becomes the limiting factor.

We defined and computed a new quantity called the beam coherence time, which is tailored to the beam alignment context. We showed that the beam coherence time is typically an order-of-magnitude longer than the conventional channel coherence time. Incorporating both the channel estimation and beam alignment overhead, we showed that beams should be realigned every beam coherence time and not every channel coherence time.



## Chapter 3

# Inverse Multipath Fingerprinting for V2I MmWave Beam Alignment

Beam alignment is a crucial component in millimeter wave (mmWave) systems, especially in fast-changing vehicular settings. In this chapter, we propose a novel and efficient beam alignment method using multipath fingerprinting. Fingerprinting is a popular approach for localization, where the measured multipath signature is compared to the multipath fingerprint database to retrieve the most likely position. Opposite to the localization idea, here vehicle positions (e.g., available via GPS) are used to query the multipath fingerprint database, which provides prior knowledge of potential pointing directions for reliable beam alignment. We define the power loss probability to measure the beam alignment accuracy. We then propose two statistical learning methods to rank promising beam directions: one is a heuristic and the other is an optimal solution derived using the concept of power loss probability. Using the beam coherence time concept from Chapter 2, we compare the performance of the proposed method with the existing standard IEEE 802.11ad and show the superiority of the proposed method, which becomes more pronounced as the antenna array size and/or the vehicle speed increases. Finally, we present an extension using a machine learning (ML) approach to rank the beam pairs.

The ML-based approach provides a more flexible framework that does not require binning the context as in the statistical learning methods. The ML approach is more scalable for incorporating multiple types of context information. This work was published in [103,104] (© IEEE).

### 3.1 Motivation and Prior Work

Prior knowledge of the propagation environment can be used to reduce the beam alignment overhead. Here, we focus on the use of multipath fingerprints, which are the long-term multipath channel characteristics associated with locations. The term “fingerprint” originates from the localization literature [17, 65, 67], where the main premise is that channel characteristics are highly correlated with locations. In fingerprinting-based localization methods, there is a fingerprint database, which records fingerprints at different locations in the area of interest. When a terminal wants to localize itself, it first performs RF channel measurements to obtain the fingerprint at the current location. The obtained fingerprint is then matched against the fingerprints in the database and the output location is computed based on the match fingerprints in the database that are “closest” to the observed fingerprint. Since position information is readily available in the vehicular context, we propose to use this idea in inverse. Localization is an important task in driving automation, where vehicles position themselves via a suite of sensors including GPS, visual cameras, and LIDAR [114]. This position information can be used to query the fingerprint database which is indexed by location to deter-

mine beam directions that are likely to provide satisfactory link quality. Large beam training overhead reduction can be expected if potential directions are concentrated, which has been observed in measurements. For example, it was reported in [26, p. 19] that in an indoor measurement with moving pedestrians the received powers were concentrated in no more than three dominant directions.

Beam alignment is a subject of intense research because of its importance for beam-based mmWave communications. Here, we summarize relevant work in the context of analog beamforming, where both the transmitter and receiver have only one RF chain. We group existing solutions into four categories, namely, approaches using beam sweeping [58, 88, 112], angle of arrival and departure (AoA/AoD) estimation [63, 73], blackbox optimization [59, 68, 69], and side information [9, 41, 61, 80, 106, 109].

Beam sweeping are simple and robust because it makes little assumptions on the channel. It only requires that the spatial channel does not change during the sweeping time. Search efficiency, however, is poor, and often hierarchical beam codebooks are used to reduce search complexity. This approach has been adopted in existing mmWave standards such as IEEE 802.15.3c [48] and IEEE 802.11ad [50] for indoor use cases. While this approach works well for short ranges, the use of wide beams in the initial stage is the bottleneck for large link distances. To maintain the signal-to-noise ratio (SNR) of the beam training, a large spreading factor or narrow beams are required in the initial stage, which dramatically decreases its efficiency.

AoA/AoD estimation leverages the sparsity of mmWave channels to reduce the number of measurements required compared to beam sweeping. For example, compressive sensing is used in [32,63] and an approximate maximum likelihood estimator is derived using the channel structure directly in [73]. Compressive measurements have to overcome the lack of antenna gain unlike the proposed approach that uses narrow beams for the beam training and thus need much less spreading in the measurements (and thus shorter training sequences).

Another idea is to use a blackbox optimization framework to efficiently explore the beam directions. This framework is based on the premise that there is some structure (e.g., smoothness) of the objective function (i.e., the received power here) and thus one does not have to blindly search all the beam pair combinations. This approach requires a good set of initial measurements and a larger number of feedbacks than other approaches to navigate the exploration.

The final category uses side information available from sensors (including communication systems at other frequencies). Radar information is used in [41], and information from lower frequencies is used in [9,80]. More related to our work are those that use position information [5, 16, 31, 61, 106]. The work in [5, 61, 106] uses position to determine beam directions directly and does not need beam training. It, however, only works when the LOS path is available. More elaborate channel models with LOS obstruction have been investigated in [16, 31]. It is proposed in [31] to memorize successful beam configurations observed in the past but no detail is given on how to rank those

configurations in terms of their likelihood to provide a good link. Omnidirectional antennas at the users are assumed in [31], which may be impractical for mmWave communications. In [16], a heuristic is proposed for a hierarchical beam search with the help of position and multipath database. Only the search in the azimuth was considered and horn antennas were assumed. Our proposed approach is in this category, where we use position information and multipath fingerprints. Unlike [5, 61, 106], by leveraging multipath fingerprints, our approach can work in both LOS and non-LOS (NLOS) channels. Different from [16, 31], the proposed beam training uses narrow beams at both the roadside unit (RSU) and the vehicle. Also, we provide a mathematical framework to rank beam pointing directions from past measurements in the fingerprint database and both azimuth and elevation are considered.

## 3.2 Contributions

The objective of this chapter is to develop an efficient beam alignment method suitable for a vehicle-to-infrastructure (V2I) setting. In this chapter, we focus on an offline learning setting where there is a dedicated period of time for collecting the training data before they are used for efficient beam alignment. We will develop an online version of the proposed method in Chapter 4. The chapter’s contributions are summarized as follows:

- We propose a framework for fast beam alignment using subsets of beam pairs. The idea is to use context information (multipath fingerprints

in our case) to select promising beam directions and only train a small subset of beam pairs.

- We propose two types of fingerprints (Types A and B), which differ in how measurements are collected and stored. Type A assumes each contributing vehicle performs an exhaustive search over all beam pairs (so that correlation between beam pairs can be captured), while Type B collects only a fraction of the exhaustive search at a time. Type A stores the raw received power, while Type B only stores the average received power of each beam pair. This provides flexibility for actual implementations, where contributing vehicles might have time constraints.
- We introduce the power loss probability as a metric for evaluating the beam alignment accuracy. This metric leads to a mathematical framework for optimizing the candidate beam pair selection. We propose two statistical-learning-based beam pair selection methods (in the sense that the metrics used have statistical interpretations), where one is a heuristic and the other is a solution that minimizes the misalignment probability.
- We provide an extensive numerical investigation, which includes the training sample size to build the fingerprint database, beam training overhead comparison with IEEE 802.11ad, and the sensitivity to changes in vehicular traffic density. For the overhead comparison, we leverage the concept of beam coherence time from Chapter 2 to quantify the beam training cost in the vehicular context. We use realistic channels gener-

ated from a commercial ray-tracing simulator, Wireless InSite [87], in all our results.

- We propose an extension to the beam pair selection using a learning-to-rank (LtR) approach from the recommender system literature. This ML-based approach eliminates the needs to bin the context information making it a more scalable framework for incorporating multiple types of context information.

We note that while we emphasize the V2I context here, the approach can also be applied to general cellular settings. An additional challenge is in how to determine the orientation of the antenna array of the user equipment (which is needed to translate AoAs/AoDs to beam indices). This is not as important for vehicles because the array is fixed on the vehicle (e.g., the roof) and the orientation can be determined from the heading of the vehicle.

The rest of the chapter is organized as follows. We define the system model in Section 3.3. In Section 3.4, we define the multipath fingerprints and explain how the proposed beam alignment works. We provide an analytical framework for quantifying beam alignment accuracy in Section 3.5. In Section 3.6, we present two statistical learning-based beam subset selection methods using the fingerprints. We show numerical evaluations of the proposed beam alignment using the statistical selection methods in Section 3.7. We develop and evaluate a more scalable beam pair selection method based on ML in Section 3.8. Finally, we conclude the chapter in Section 2.8.

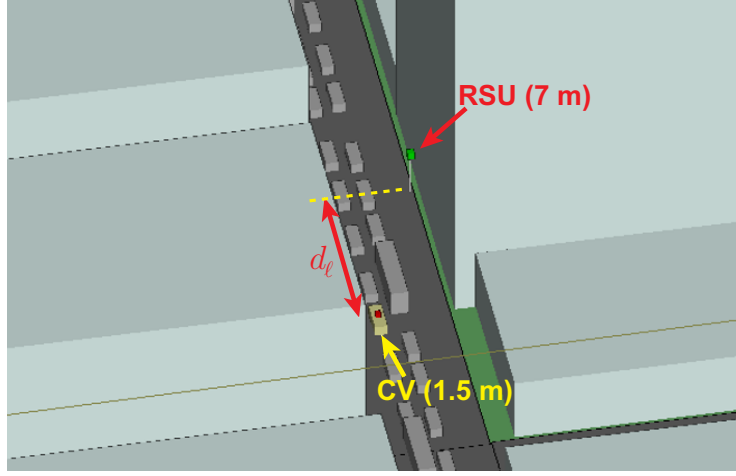


Figure 3.1: Ray-tracing simulation environment. The scenario is an urban street with two lanes. There are two types of vehicles: cars and trucks. The CV is a car on the left lane and its LOS path can be blocked by a truck. The antenna heights are 1.5 m at the CV (roof-mounted) and 7 m at the RSU.

### 3.3 System Model

We consider an urban street canyon environment with high traffic density, where LOS is often unavailable and is a challenging scenario for beam alignment. We start by describing the ray-tracing environment and how the outputs are used to obtain the channel matrices. Then, we describe the received signal model and the beam codebook.

#### 3.3.1 Channel Model

The simulation environment is shown in Fig. 3.1, which is an urban street with two lanes. All the buildings are made of concrete (relative permittivity  $\epsilon_r = 5.31$  and conductivity  $\sigma = 0.8967$  S/m [51, Table 3]), and the



road surfaces are made of asphalt ( $\epsilon_r = 3.18$  and  $\sigma = 0.3338$  S/m [70]). The surface root-mean-square roughness is set to 0.2 mm for concrete and 0.34 mm for asphalt [70, Table 1]. We allow up to two reflections and one diffraction. We simulate two types of vehicles represented by metal boxes (made of the perfect electric conductor which is predefined in Wireless InSite [87]): cars ( $1.8\text{ m} \times 5\text{ m} \times 1.5\text{ m}$ ) and trucks ( $2.5\text{ m} \times 12\text{ m} \times 3.8\text{ m}$ ). The type of each vehicle is selected randomly with 0.6 probability for a car and 0.4 probability for a truck. The average cars-to-trucks ratio is 3:2. The RSU is placed on the right side, and a car on the left lane is selected as the communicating vehicle (CV). The antenna heights are 7 m for the RSU and 1.5 m for the CV (on its roof). Because trucks are taller, they could block the LOS path between the CV and the RSU. The carrier frequency is set to 60 GHz.

To imitate the dynamic blockage environment, we simulate multiple snapshots of the scenario where vehicles are independently and randomly placed in each snapshot. The gap between vehicles (i.e., from the front bumper to the rear bumper of the heading vehicle)  $\zeta$  is assumed to be IID and follow the Erlang distribution [8] with the probability density given by

$$f_{\zeta}(\zeta) = \frac{(\kappa\mu_{\zeta})^{\kappa}}{(\kappa-1)!} \zeta^{\kappa-1} e^{-\kappa\mu_{\zeta}\zeta}, \quad (3.1)$$

where  $\kappa$  is the shape parameter and  $1/\mu_{\zeta}$  is the mean gap. Following [74],  $\kappa = 6$  and  $\mu_{\zeta} = 0.209$  are used in our simulations, which produces an average gap of 4.78 m. Since multipath fingerprints are associated with locations, we need to generate multiple channels at a given location. To do this, the CV

is placed at a longitudinal distance  $d_\ell$  from the RSU (see Fig. 3.1), where  $d_\ell$  is uniformly drawn from  $[d_0 - \sigma_d, d_0 + \sigma_d]$  for some mean distance  $d_0$  and some grid size  $2\sigma_d$ .  $d_0 = 30$  m and  $\sigma_d = 2.5$  m are used when generating the channels. When applying our method, all points within this range are treated as a location bin indexed by  $d_0$ . By discretizing the location this way, the system has some resilience to errors in positions estimated by the vehicle. The edge effect can be mitigated by defining overlapping location bins.

In this chapter and the next, no mobility is considered during the beam training. The beam training duration for the proposed method is sub-millisecond (see Section 3.7.6), so the displacement of vehicles during the beam training is negligible. For example, when  $16 \times 16$  arrays are used, the training duration is about  $150 \mu\text{s}$ , and the displacement is only 3 mm even with a speed of 20 m/s. Our proposed beam alignment can be considered as an initial link establishment, and after which a beam tracking method such as [38] could be used to maintain the link to further reduce the overhead.

We use the wideband geometric channel model parametrized by the ray-tracing output. A ray-tracing simulation outputs a number of rays, each corresponding to a propagation path. The information associated with each ray includes the received power, the delay, the phase, the AoA, and the AoD. Denoting  $(\cdot)^*$  the conjugate transpose,  $N_t$  and  $N_r$  the numbers of transmit and receive antennas,  $L_p$  the number of rays,  $\alpha_\ell$  the complex channel gain,  $\tau_\ell$  the delay,  $\theta_\ell^A$  and  $\theta_\ell^D$  the elevation AoA and AoD,  $\phi_\ell^A$  and  $\phi_\ell^D$  the azimuth AoA and AoD of the  $\ell$ -th ray,  $g(\cdot)$  the combined effect of lowpass filtering and pulse

shaping,  $B$  the system bandwidth,  $T = 1/B$  the symbol period, and  $\mathbf{a}_r(\cdot)$  and  $\mathbf{a}_t(\cdot)$  the receive and transmit steering vectors, the channel can be written as

$$\mathbf{H}[n] = \sqrt{N_r N_t} \sum_{\ell=0}^{L_p-1} \alpha_\ell g(nT - \tau_\ell) \mathbf{a}_r(\theta_\ell^A, \phi_\ell^A) \mathbf{a}_t^*(\theta_\ell^D, \phi_\ell^D). \quad (3.2)$$

The raised cosine filter with a roll-off factor of 0.1 is assumed for the pulse shaping filter. The number of rays is  $L_p = 25$ . We use the Full3D model with Shooting and Bouncing Ray (SBR) tracing mode in Wireless InSite. The ray spacing is  $0.25^\circ$ , which means the simulator shoots hundreds of rays and determines which of them form valid propagation paths and records the strongest 25 of them. The power gap among these 25 rays is more than 20 dB, and thus there is little value in keeping more rays. Note that there is no need to specify the number of clusters or rays per cluster as typical of stochastic channel models. The ray-tracing will determine all relevant propagation paths. In fact, some of the 25 rays can be thought of as belonging to the same cluster since they have similar delays and AoAs. Note that by using ray-tracing, we ensure that the channels are spatially consistent, which is a feature not available in most stochastic channel models.

Uniform planar arrays (UPA) are assumed at both the transmitter and the receiver. With a UPA, each beam is defined by its azimuth  $\phi$  and elevation  $\theta$  main beam direction. Let  $G_{\text{ant}}(\cdot)$  be the antenna element radiation pattern,  $\Omega_y = kd_y \sin(\theta) \sin(\phi)$ ,  $\Omega_x = kd_x \sin(\theta) \cos(\phi)$ ,  $k = 2\pi/\lambda$  be the wave number,  $\otimes$  denote the Kronecker product,  $N_x$  and  $N_y$  be the numbers of elements along the x- and y-axis, and  $d_x$  and  $d_y$  be the element spacing in the x- and

y-direction, a beam pointing in  $(\theta, \phi)$  direction is given by [18]

$$\mathbf{a}(\theta, \phi) = \frac{G_{\text{ant}}(\theta, \phi)}{\sqrt{N_x N_y}} \begin{bmatrix} 1 \\ e^{j\Omega_y} \\ \vdots \\ e^{j(N_y-1)\Omega_y} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ e^{j\Omega_x} \\ \vdots \\ e^{j(N_x-1)\Omega_x} \end{bmatrix}.$$

We assume  $d_x = d_y = \lambda/2$ . We assume no backplane radiation and set

$$G_{\text{ant}}(\theta, \phi) = \begin{cases} 0 & \text{if } \theta > 90^\circ \\ 1 & \text{otherwise} \end{cases}. \quad (3.3)$$

We use (3.3) for simplicity, but we can replace it with a more sophisticated one like that of a patch antenna.

### 3.3.2 Received Signal Model

We assume an analog beamforming, which uses only one RF chain. We assume the symbol timing is synchronized to the first path (shortest delay). This means that paths with larger delays are not likely synchronized to the sampling timing and the energy will leak to adjacent symbols. The leak amount is determined by the combined filter response  $g(\cdot)$ . The received signal of the  $i$ -th beam pair can be written as the time-domain convolution between the transmit signal and the effective channel seen through the  $i$ -th beam pair  $h_i[n]$ , i.e.,

$$y_i[k] = \sqrt{P_t} \sum_{n=0}^{L-1} s[k-n] \underbrace{\sum_{\ell=0}^{L_p-1} g(nT + \tau_0 - \tau_\ell) \mathbf{w}_{r(i)}^* \mathbf{H}_\ell \mathbf{f}_{t(i)}}_{h_i[n]} + v_i[k], \quad (3.4)$$

where  $\mathbf{H}_\ell = \sqrt{N_r N_t} \alpha_\ell \mathbf{a}_r(\theta_\ell^A, \phi_\ell^A) \mathbf{a}_t^*(\theta_\ell^D, \phi_\ell^D)$  the channel matrix corresponding to the  $\ell$ -th path,  $L$  is the channel length,  $P_t$  is the transmit power,

$s[k]$  is the known training signal,  $r(i)$  and  $t(i)$  denote the mapping of the beam pair index  $i$  to the combiner  $\mathbf{w}$  and beamformer  $\mathbf{f}$  vector indices, and  $v_i[k]$  is the zero mean complex Gaussian noise  $\mathcal{CN}(0, \sigma_v^2)$ . Let  $K \geq L$  be the training sequence length and  $\mathbf{y}_i = [y_i[0] \ y_i[1] \ \dots \ y_i[K-1]]^T$ ,  $\mathbf{h}_i = [h_i[0] \ h_i[1] \ \dots \ h_i[L-1]]^T$  and  $\mathbf{v}_i = [v_i[0] \ v_i[1] \ \dots \ v_i[K-1]]^T$ . Assuming a long enough cyclic prefix, the received signal after discarding the cyclic prefix can be rewritten in a matrix form as

$$\mathbf{y}_i = \mathbf{S}\mathbf{h}_i + \mathbf{v}_i, \quad (3.5)$$

where  $\mathbf{S}$  is the  $K \times L$  circularly shifted training sequence. The channel can be estimated using a least-square approach as

$$\hat{\mathbf{h}}_i = \mathbf{S}^\dagger \mathbf{y}_i \quad (3.6)$$

$$= \mathbf{h}_i + \mathbf{S}^\dagger \mathbf{v}_i \quad (3.7)$$

where  $\mathbf{S}^\dagger = (\mathbf{S}^* \mathbf{S})^{-1} \mathbf{S}^*$  is the pseudo-inverse. Using a training sequence with good correlation properties like Zadoff-Chu or Golay sequences and a proper power scaling,  $\mathbf{S}^* \mathbf{S} = \mathbf{I}$  and the estimation error  $\tilde{\mathbf{v}}_i = \mathbf{S}^\dagger \mathbf{v}_i$  can be modeled as  $\mathcal{CN}(\mathbf{0}, \frac{\sigma_v^2}{K} \mathbf{I})$  [46]. We refer to [46, Chapter 5] for more details. Note that the database collection should use the same  $P_t$  (or scaled appropriately if different  $P_t$ 's are used). In our simulations, following IEEE 802.11ad, we set the system bandwidth  $B = 1760$  MHz and  $K = 512$ . The Channel Estimation Field (CEF) of an IEEE 802.11ad frame consists of two training sequences of length 512 which can be used to perform two independent channel estimations [50].

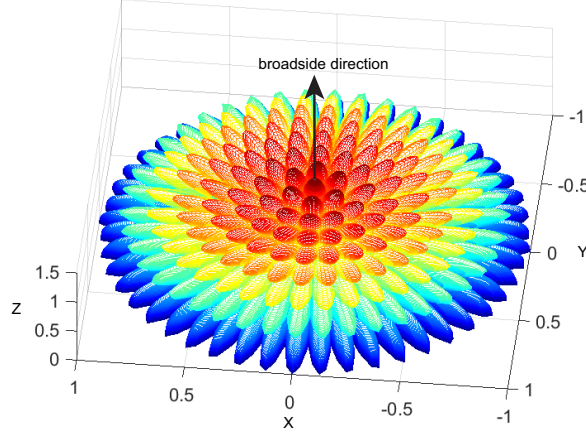


Figure 3.2: Beam patterns in our codebook for an  $8 \times 8$  array. The array is assumed to face upward in the  $+z$  direction. The codebook covers the directions in the  $+z$  half-space (i.e., assuming no radiation in the backplane).

In our setting, the delay spread is larger than the indoor scenario, and thus we assume the whole CEF is used for only one channel estimation. The actual channel length  $L$  varies for different snapshots of the ray-tracing simulation and can be larger than 512. Since the powers of those paths with large delays are observed to be negligible compared to paths with short delays, we truncate the channel to get  $L = 512$ .

The vector  $\mathbf{w}_{r(i)}$  and  $\mathbf{f}_{t(i)}$  are selected from the receiver codebook  $\mathcal{W}$  and transmitter codebook  $\mathcal{F}$ . We assume UPAs are used at both the CV and the RSU. The beams in the codebook are generated using progressive phase shift [82] between antenna elements. Fig. 3.2 shows the beams for an  $8 \times 8$  array. We numerically determine the 3 dB beamwidths of the beams (which depends on the main beam direction) and select the beams such that they

are separated by the 3 dB beamwidth starting from the broadside direction. This ensures that the array gain fluctuates less than 3 dB over the entire field of view of the antenna array. We note that our proposed approach does not depend on this specific codebook, and another codebook can replace the one used here. Codebooks adapted to the environment, however, can be expected to provide higher beamforming gains. This idea is pursued in Chapter 4 where we propose an online beam pair refinement algorithm.

### **3.4 Beam Alignment using Subset of Beam Pairs**

The main idea of the proposed approach is to leverage prior knowledge to identify promising beam directions and only train those directions. The prior knowledge is obtained from past observations in the database, and there is no guarantee that all paths seen so far are present in the current channel. For example, some paths do not appear due to blockage. Therefore, beam training among the beam directions identified from the database is still required. The beam training here, however, has a much lower overhead than conventional methods because a large number of unlikely directions have already been eliminated using the database. In this section, we will define fingerprints and explain how the database is constructed and the cost involved. Then, we will describe the proposed approach for beam alignment.

### 3.4.1 Multipath Fingerprint Database

In general, a fingerprint refers to some characteristics of the channel at a given location. These characteristics could be the received signal strengths from different access points [67] or the multipath signature of the channel from an access point [65]. In this dissertation, a fingerprint refers to a set of received powers of different pairs of transmit and receive beams at a given location bin  $[d_0 - \sigma_d, d_0 + \sigma_d]$ .

We define two types of fingerprints, which differ by how measurement data are collected and stored. The first type, called Type A, requires that the contributing vehicle perform full exhaustive beam measurements over all beam pairs. This ensures that the measurements over the different pairs happen within a beam coherence time so that the spatial channel does not change. This way, the fingerprint captures the correlation between the different beam pairs, i.e., whether they tend to have similar received powers or not. For Type A fingerprints, the raw measurement samples are stored. One could store all the measurements of all the beam pairs from each contributing vehicle, but this is not necessary. Only the measurements of the top- $C$  beam pairs (ranked by the received power) can be stored. This is because most of the beam pairs do not point along any propagation paths and have negligible received powers, and thus there is negligible information gain in keeping all beam pairs. In our simulation, we use  $C = 100$ . The average power ratio between the strongest beam pair and the 100th strongest over the channel samples used is 22.2 dB. An example of Type A fingerprints is shown in Table 3.1. In this example,



Table 3.1: An example of Type A fingerprints. For each contributing vehicles, the measurements of the top- $C$  beam pairs are stored. In each cell, the top number is the beam pair index and the bottom one is the received power.

Observation No.	Best	2nd best	...	$C$ -th best
1	5 -64.5 dBm	159 -69.2 dBm	...	346 -95.8 dBm
2	159 -70.4 dBm	263 -72.6 dBm	...	354 -97.1 dBm
...	...	...	...	...
$N$	5 -66.4 dBm	258 -68.1 dBm	...	2 -82.6 dBm

there are  $N$  observations collected by  $N$  contributing vehicles.

Type B fingerprints do not require that the measurement of all the beam pair combinations be completed within a beam coherence time. This less restrictive data collection reduces the burden on individual vehicles contributing to building the database; they do not need to do a full exhaustive search and could contribute as many beam measurements as their time allows. A simple method is to collect the data in a round-robin manner. For example, assume that each vehicle can do only  $1/4$  of the full exhaustive search, the RSU can divide the set of all beam pairs into four disjoint sets and assign these sets sequentially to subsequent contributing vehicles to collect the measurement data. The disadvantage is that now the correlation between the beam pairs cannot be easily captured. Thus, there is no benefit to store the raw samples, and only the average received powers (computed in linear scale) of the beam pairs are stored. An example of Type B fingerprints is shown in Table 3.2. We note that the sample averages can be computed recursively,

Table 3.2: An example of Type B fingerprints at a location bin. The average received power for each beam pair is recorded.

Beam pair index	1	2	...
Average received power	-92.3 dBm	-73.5 dBm	...

and thus there is no need to temporally store the raw samples when collecting Type B fingerprints.

We focus on an offline learning setting to build the database in this chapter. An online extension is the subject of Chapter 4. In the proposed approach, the RSU builds and stores the database. By offline learning, we mean there is a dedicated period of time for collecting the fingerprint database before it is exploited for efficient beam alignment. During this period, each contributing vehicle conducts beam training with the RSU. By having the vehicle be the transmitter during the beam training, there is no feedback needed to report the measured received powers. If each contributing vehicle can perform a full exhaustive search over all beam pair combinations in the codebook, we obtain Type A fingerprints. Otherwise, we get Type B fingerprints.

We now discuss the cost for building and storing the database. The database can be collected in the initial stage of the system deployment. The RSU can request vehicles passing by its coverage to conduct beam training. Most modern vehicles are GPS equipped either for navigation or for safety message (position, speed, heading etc.) dissemination such as in DSRC. Thus, it is fair to assume that any vehicles equipped with mmWave communication also have positioning capability, and all mmWave communication capable ve-

hicles can contribute to building the database. As will be shown in Section 3.7.2, around 250 full exhaustive beam measurements are enough to get a fully functioning database. These measurements can be collected from 250 vehicles if we assume Type A fingerprints. In a dense urban traffic setting, this could be done within an hour if not less (e.g., the average number of vehicles passing through an urban road segment was around 400 per lane per hour in the NGSIM Lankershim dataset [36]). We note that if the exhaustive search is used as the beam alignment method when the database is not available [71], the data collection is essentially free since the vehicles will need to conduct the exhaustive search to establish the link during this stage.

Once the database is built, the updating cost is low. The main premise of the proposed method is to learn the long-term multipath information from the database, which is the propagation directions that depend on the geometry of the environment such as the road structure and buildings. The change in the traffic density can affect the relative importance of different paths, but as shown in Section 3.7.5, if the database is collected in a dense traffic, it will also work well in light traffic conditions. Thus, we expect the database collected in high traffic density to be valid for a long period of time such as weeks or even months if there is no construction in the surrounding buildings and road structure. Of course, once a functioning database has been established, it can be reinforced by having idle vehicles passing by conduct beam measurements and replace the older data with the newly collected ones. This update can be done at a slow pace since we expect the database to change slowly.

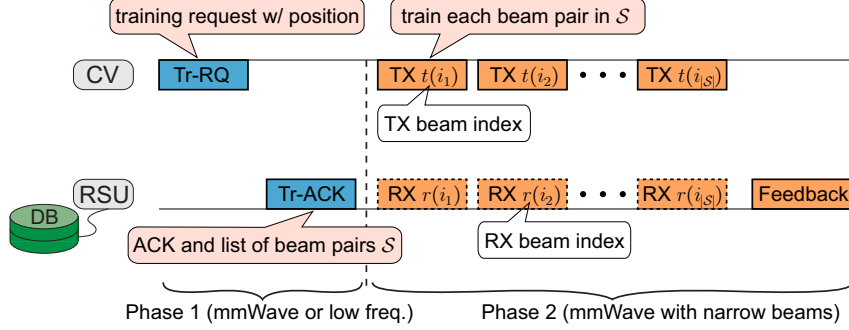


Figure 3.3: An illustration of position-aided beam alignment in the uplink. It consists of two phases. Phase 1 is for the training request where the CV position is sent to the RSU. The RSU uses the position and its learned database to determine a list of promising beam pairs  $S$ . In Phase 2, the beam pairs in the list are trained, and a feedback indicated the best beam index is sent at the end.

Finally, we consider the storage requirement. For Type A fingerprints, using 4 bytes for one received power and 2 bytes for one beam index, the total storage of Table 3.1 is  $6NC$  bytes. Assuming  $N = 250$ ,  $C = 100$  and 200 location bins per RSU coverage (1 m bin size and 200 m RSU coverage), it requires about 30 MB. Type B requires even less data storage. This amount of data can be easily stored in any modern devices. Therefore, we conclude that storage is not at all a problem.

### 3.4.2 Proposed Beam Alignment

Fig. 3.3 illustrates the position-aided beam alignment, which consists of two phases. We start with the uplink. In Phase 1, the CV sends a training request along with its context information to the RSU. In this dissertation, we use position as the context. The RSU uses the position and the database it

maintains to determine a subset of promising beam directions, denoted by  $\mathcal{S}$ . The size of the set  $\mathcal{S}$  is a system parameter that is chosen to balance the training overhead and the alignment accuracy. Beam pair selection methods will be described in Section 3.6. The RSU then responds with an acknowledgment and the beam pair subset  $\mathcal{S}$  to the CV. Since the beams are not aligned in this phase, a lower frequency control channel or mmWave with a large spreading factor can be used. In Phase 2, the beam pairs in  $\mathcal{S}$  are trained and the best beam index is fed back at the end. MmWave with narrow beams is used during this phase. In the proposed methods, the learning happens at the RSU.

In the downlink, Phase 1 changes slightly. The process starts with the RSU sending a training request to the CV, which then responds with an acknowledgment including its position. The RSU, then, sends the list of promising beam pairs  $\mathcal{S}$ . The beam training in Phase 2 is kept the same. This is possible because of the reciprocity in the AoA/AoDs, where the AoAs become the AoDs and vice versa when reversing the transmitter and receiver role. This AoA/AoD reciprocity only depends on the reciprocity property of electromagnetic waves, which holds when they propagate in passive medium like wireless channels (excluding the device's circuits) [18].

Several remarks on the proposed method now follow.

*Remark 1:* The position information here does not have to be highly accurate. It only needs to be accurate enough to identify the location bin index of the fingerprints. In our simulation, this bin size is 5 m. Edge effects can be mitigated by having overlapping location bins.

*Remark 2:* The proposed method allows graceful degradation as the number of beam pairs trained  $N_b$ , which is the size of  $\mathcal{S}$ , decreases because the alignment accuracy decreases probabilistically with  $N_b$  and there is no hard threshold on  $N_b$ . See Fig. 3.9 for an example of how the average rate changes with  $N_b$ . This is a desirable feature that allows the tradeoff between latency in link establishment and accuracy of the beam alignment.

*Remark 3:* Our method performs the beam training using only narrow beams, which has several advantages. Narrow beams provide high antenna gain and are more resilient to Doppler spread [102]. Also, methods employing wide quasi-omni beams can suffer from antenna gain fluctuation because it is challenging to produce wide beamwidths with low gain fluctuation [58].

*Remark 4:* By having the CV transmit and the RSU receive during the beam training, the RSU obtains beam measurements for free, i.e., without any feedback from the CV. These measurements are useful for updating the database in an online setting.

### 3.5 Quantifying Beam Alignment Accuracy

In this section, we define a metric for measuring the beam alignment accuracy, which allows us to compare different candidate beam pair selection methods. We will use this metric for optimizing the method to select beam pairs for training in Section 3.6. The definition assumes the measurement noise is negligible. In the presence of measurement noise, the metrics computed using (3.8) and (3.9) will be less accurate and can affect the beam pair selection

and degrade the beam alignment accuracy. We investigate the effect of noise numerically in Section 3.7.4 and show that the degradation due to noise is negligible if the transmit power is not too low.

The power loss is defined as the ratio between the received power of the optimal beam pair and the pair selected by the beam alignment method indexed by  $s$ . Let  $\mathcal{B}$  be the set of all beam pair combinations, and  $\gamma_\ell = \|\mathbf{h}_\ell\|^2$  be the received power of the  $\ell$ -th beam pair. The power loss can be written as

$$\xi = \frac{\max_{k \in \mathcal{B}} \gamma_k}{\gamma_s}. \quad (3.8)$$

If noise is negligible, the strongest beam pair will be selected after the beam training so that  $\gamma_s = \max_{i \in \mathcal{S}} \gamma_i$ , where  $\mathcal{S} \subset \mathcal{B}$  is the set of candidate beam pairs selected for beam training. The power loss probability is defined as the probability that  $\xi > c$  for some  $c \geq 1$ , i.e.,

$$P_{\text{pl}}(c, \mathcal{S}) = \mathbb{P}[\xi > c] \quad (3.9)$$

$$= \mathbb{P} \left[ \max_{k \in \mathcal{B}} \gamma_k > c \max_{i \in \mathcal{S}} \gamma_i \right]. \quad (3.10)$$

We note that since  $\mathcal{S}$  is a subset of  $\mathcal{B}$ , the definition in (3.10) is always well-defined.

### 3.6 Beam Pair Subset Selection using Statistical Learning

In this section, we propose two methods to select candidate beam pairs for beam training using the information in the fingerprint database. The

methods are based on statistical learning in the sense that the metrics used have clear statistical interpretations. The objective of the selection methods is to maximize the received power of the finally selected beam pair for a given beam training budget of  $N_b$ . The first approach is a heuristic, while the second one minimizes the misalignment probability defined in Section 3.5. The heuristic is intended to be used with Type B fingerprints, and the other method is to be used with Type A fingerprints.

### 3.6.1 Selection by Ranking Average Received Powers

This method is based on the simple intuition that we should choose candidate beam pairs with the highest expected received power. The proposed approach is to first rank the beam pairs by their average received powers in descending order and select the highest  $N_b$  pairs for beam training. Note that this metric can balance the selection of opportunistic paths that occasionally have high received power. An intuitive explanation follows like this. If the occurrence of the opportunistic path is high enough, its average received power will be larger than a path that always has moderate received power and the opportunistic path is selected. If the occurrence is rare, the average received power of the opportunistic path is low and the path with always moderate received power is selected by this method. Thus, this metric can balance the risk and gain to some extent. We call this method AvgPow.



---

**Algorithm 3.1** Greedy candidate beam pair selection

---

```
1:  $\mathcal{S}_0 \leftarrow \emptyset$ 
2: for  $n = 1 : N_b$  do
3:    $\mathcal{S}_n \leftarrow \mathcal{S}_{n-1} \cup \arg \min_{i \in \mathcal{B} \setminus \mathcal{S}_{n-1}} P_{\text{pl}}(1, \mathcal{S}_{n-1} \cup \{i\})$ 
4: end for
```

---

### 3.6.2 Selection by Minimizing the Misalignment Probability

Since the objective of beam alignment is to maximize the received power, an indirect way to achieve that is to choose the set of beam pairs to minimize the misalignment probability, which is the power loss probability  $P_{\text{pl}}(c = 1, \mathcal{S})$ . For a given training budget of  $N_b$ , the problem can be formulated as a subset selection problem given by

$$\begin{aligned} & \underset{\mathcal{S} \subset \mathcal{B}}{\text{minimize}} && P_{\text{pl}}(1, \mathcal{S}) \\ & \text{subject to} && |\mathcal{S}| = N_b. \end{aligned} \tag{3.11}$$

Here,  $|\mathcal{S}|$  denotes the cardinality of the set  $\mathcal{S}$ . This is a subset selection problem, which is combinatoric in nature and is difficult to solve in general, especially when  $|\mathcal{B}|$  is large. Fortunately, the structure of  $P_{\text{pl}}(1, \mathcal{S})$  allows an efficient solution. Note that the problem (3.11) is equivalent to maximizing  $P_{\text{opt}}(\mathcal{S}) \equiv 1 - P_{\text{pl}}(1, \mathcal{S})$ , which we call the probability of being optimal. Since  $P_{\text{opt}}(\mathcal{S})$  is a modular function, the greedy solution given in Algorithm 3.1 is optimal [34, Theorem 7]. This is a well-known result that has been reported in different forms in the literature (see [22] and references therein).

**Proposition 3.1.**  $P_{\text{opt}}(\mathcal{S})$  is modular.

*Proof.* Using the definition of power loss probability in (3.10) with  $c = 1$ ,

$$P_{\text{pl}}(1, \mathcal{S}) = \mathbb{P} \left[ \max_{k \in \mathcal{B}} \gamma_k > \max_{i \in \mathcal{S}} \gamma_i \right] \quad (3.12)$$

$$= \sum_{\ell \in \mathcal{B}} \mathbb{P} \left[ \gamma_\ell > \max_{i \in \mathcal{S}} \gamma_i \mid \gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k \right] \mathbb{P} \left[ \gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k \right] \quad (3.13)$$

$$= \sum_{\ell \in \mathcal{B} \setminus \mathcal{S}} \mathbb{P} \left[ \gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k \right] \quad (3.14)$$

$$= 1 - \sum_{\ell \in \mathcal{S}} \mathbb{P} \left[ \gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k \right], \quad (3.15)$$

where (3.13) is the application of the law of total probability on the event  $\{\gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k\}$ , and (3.14) follows because if  $\ell \in \mathcal{S}$  then  $\mathbb{P}[\gamma_\ell > \max_{i \in \mathcal{S}} \gamma_i \mid \gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k] = 0$  and if  $\ell \in \mathcal{B} \setminus \mathcal{S}$  then  $\mathbb{P}[\gamma_\ell > \max_{i \in \mathcal{S}} \gamma_i \mid \gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k] = 1$ . Using the fact that  $\sum_{\ell \in \mathcal{B}} \mathbb{P}[\gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k] = 1$ , we obtain (3.15). By definition  $P_{\text{opt}}(\mathcal{S}) \equiv 1 - P_{\text{pl}}(1, \mathcal{S})$ , we have

$$P_{\text{opt}}(\mathcal{S}) = \sum_{\ell \in \mathcal{S}} \mathbb{P} \left[ \gamma_\ell = \max_{k \in \mathcal{B}} \gamma_k \right]. \quad (3.16)$$

Thus, for any  $\mathcal{S} \subset \mathcal{T} \subset \mathcal{B}$  and  $\forall n \in \mathcal{B} \setminus \mathcal{T}$ , we have  $P_{\text{opt}}(\mathcal{S} \cup \{n\}) - P_{\text{opt}}(\mathcal{S}) = P_{\text{opt}}(\mathcal{T} \cup \{n\}) - P_{\text{opt}}(\mathcal{T}) = \mathbb{P}[\gamma_n = \max_{k \in \mathcal{B}} \gamma_k]$ , which is the definition of modular functions [22, 79].  $\square$

While the solution in Algorithm 3.1 is intuitive, using a brute force search to solve the minimization problem at each selection step is not efficient. At each step, we need to evaluate the power loss probability  $|\mathcal{B} \setminus \mathcal{S}_{n-1}| \leq$

$|\mathcal{B}| - N_b$  times. Since  $|\mathcal{B}|$  typically is much larger than  $N_b$ , this means that the total number of evaluations is  $\mathcal{O}(N_b|\mathcal{B}|)$ . Now, if we use the probability of being optimal,

$$P_{\text{opt}}(i) = \mathbb{P} \left[ \gamma_i = \max_{k \in \mathcal{B}} \gamma_k \right] \quad (3.17)$$

$$= \mathbb{P} [\gamma_i \geq \gamma_k, \forall k \in \mathcal{B} \setminus \{i\}], \quad (3.18)$$

the proof of Proposition 3.1 suggests a more efficient solution. From (3.15), we see that minimizing  $P_{\text{pl}}(1, \mathcal{S}_{n-1} \cup \{i\})$  over  $i \in \mathcal{B} \setminus \mathcal{S}_{n-1}$  is equivalent to solving

$$k = \arg \max_{i \in \mathcal{B} \setminus \mathcal{S}_{n-1}} P_{\text{opt}}(i). \quad (3.19)$$

This means that Algorithm 3.1 is equivalent to selecting the beam pairs by ranking their probability of being optimal in descending order. This solution requires to compute the probability of being optimal  $\mathcal{O}(|\mathcal{B}|)$  times.

We now present how to compute  $P_{\text{opt}}(i)$  from the database. Note that Type B cannot be used to compute  $P_{\text{opt}}(i)$  because it only stores the average received powers. Denote  $\gamma_{nk}$  the received power observed at the  $k$ -th beam pair in the  $n$ -th observation,  $P_{\text{opt}}(i)$  is estimated using Type A fingerprints by

$$P_{\text{opt}}(i) \simeq \frac{1}{N} \sum_{n=1}^N \mathbf{1}(\gamma_{ni} > \gamma_{nk}, \forall k \in \mathcal{B} \setminus \{i\}), \quad (3.20)$$

where  $N$  is the number of observations in the fingerprint database (number of rows of Table 3.1), and  $\mathbf{1}(E)$  is the indicator function which outputs 1 if  $E$  is true and 0 otherwise. Note that if we choose to keep  $C < |\mathcal{B}|$  measurements per

contributing vehicle, not all  $\gamma_{nk}$  for  $k = 1, 2, \dots, |\mathcal{B}|$  are recorded. We assume those  $\gamma_{nk}$  that are not recorded to be zero in (3.20). This is a reasonable approximation because these  $\gamma_{nk}$  are much smaller than the received power of the top ranked beam pairs that are recorded in the database. The expression in (3.20) is equivalent to counting how often the  $i$ -th beam pair is observed to be the strongest. Therefore,  $P_{\text{opt}}(i)$  is estimated to be 0 for all beam pairs that have not yet been seen to be the strongest in the database collected. This means less important beam pairs that rarely provide the strongest received power are difficult to rank using (3.20) with a reasonable  $N$ . In fact, in our simulation in Section 3.7.1 with  $N = 450$ , there are about 30 distinct beam pairs that are observed to be the best at least once in the database. This means that using (3.20), we can produce a ranked list of length up to around 30. If the allowable training budget  $N_b$  is larger than 30, we need to produce a longer list that ranks the less important beam pairs while using the same database.

To help select those less frequent beam pairs without using a too large  $N$ , we propose to rank these less important beam pairs by the same metric but computed while ignoring the correlation in the fingerprints. By assuming the independence between the pairs, we have

$$P_{\text{opt}}(i) = \mathbb{E}_{\gamma_i} \left[ \prod_{k \in \mathcal{B} \setminus \{i\}} \mathbb{P}[\gamma_i > \gamma_k | \gamma_i] \right] \quad (3.21)$$

$$\simeq \frac{1}{N} \sum_{n=1}^N \prod_{k \in \mathcal{B} \setminus \{i\}} \frac{1}{N} \sum_{m=1}^N \mathbf{1}(\gamma_{ni} > \gamma_{mk}). \quad (3.22)$$

To summarize, this beam pair selection method uses both (3.20) and (3.22) to produce a ranked list of beam pairs. Let  $\mathcal{B}_{\text{nz}}$  be the set of all beam pairs with nonzero  $P_{\text{opt}}(i)$  according to (3.20), then the top- $|\mathcal{B}_{\text{nz}}|$  in the ranked list are obtained using (3.20), and the rest of the beam pairs  $\mathcal{B} \setminus \mathcal{B}_{\text{nz}}$  are ranked using (3.22). We call this beam pair selection method MinMisProb.

### 3.7 Numerical Evaluations of Statistical Beam Pair Subset Selection

This section provides numerical evaluations of the proposed beam alignment using the two statistical beam pair subset selection methods. All evaluations here use a dataset of 500 channel samples generated from the ray-tracing simulator (see Fig. 3.1) and assume  $16 \times 16$  UPAs at both the CV and RSU unless stated otherwise. The codebook for the  $16 \times 16$  array has 271 beams. We conduct  $K$ -fold cross validation, with  $K = 10$  as recommended in [45]. Specifically, the dataset of 500 channel samples is divided into 10 subsets (or folds) of size 50 each. Then, nine of them are used as the training set to build the database, and the remaining one is used as the test set to evaluate the proposed beam alignment. This is repeated 10 times, where each time a different subset is selected as the test data. When a different evaluation method is used, it will be stated explicitly. Common simulation parameters are summarized in Table 3.3.

The two types of database are built in the following manners in the simulations. For Type A fingerprints, each contributing vehicle conducts a full

Table 3.3: Common simulation parameters

Parameters	Value
Carrier frequency	60 GHz
Bandwidth	1760 MHz
Antenna array	16×16 UPA
Mean vehicle gap	4.78 m

exhaustive beam measurement and the top-100 beam pairs are recorded as explained in Section 3.4.1. For a fair comparison, the two types of database should be built using the same number of measurement data. Thus, the database for Type B is obtained by summarizing Type A database. Specifically, instead of keeping all the raw received powers, only the average received power is recorded for each beam pair. We note that in actual implementations of Type B data collection, a full exhaustive search can be collected by a number of vehicles depending on the time budget the vehicles have as explained in Section 3.4.1.

### 3.7.1 Performance Comparison: AvgPow Versus MinMisProb

This subsection presents a performance comparison of the two statistical beam pair selection methods when the measurement noise is negligible. The impact of noise will be shown in Section 3.7.4. Fig. 3.4 shows the power loss probability of the two selection methods. Two different levels of power loss severity are shown: the misalignment probability  $P_{\text{pl}}(0 \text{ dB}, \mathcal{S})$  and the probability that the power loss is less than 3 dB  $P_{\text{pl}}(3 \text{ dB}, \mathcal{S})$  (called the 3 dB power loss probability). MinMisProb dominates AvgPow at both levels

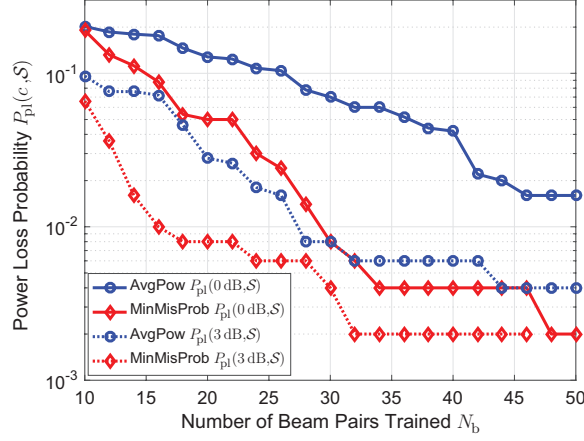


Figure 3.4: Power loss probability versus the number of beam pairs trained. The 3 dB power loss probability plot for MinMisProb ends before reaching  $N_b = 50$  because there is no such instance of power loss computed from the cross validation. MinMisProb outperforms AvgPow in both the misalignment and 3 dB power loss probability.

of misalignment. This is expected because MinMisProb is optimal (in terms of the misalignment probability) by its definition that exploits the correlation between the different beam pairs available in Type A fingerprints. We note that when computing the probability of being optimal  $P_{\text{opt}}(\cdot)$  using (3.20), the number of beam pairs with nonzero  $P_{\text{opt}}(\cdot)$  is around 30 (the exact number depends on the chosen training data). The plots of MinMisProb become flat at  $N_b$  of around 30. This means that the complementary selection using (3.22) does not perform as well as when using (3.20) that exploits the correlation information. It, however, can still identify relevant beam pairs without additional training data. We conclude from Fig. 3.4 that if Type A fingerprints are available, MinMisProb is the choice; otherwise, the AvgPow should be used.

### 3.7.2 Required Training Sample Size

This subsection provides an empirical evaluation to estimate the training sample size to build the fingerprint database. By sample size, we mean the number of exhaustive beam measurements conducted to collect the data (i.e., the number of rows  $N$  of Table 3.1). We start with the description of the evaluation method. We still use the 10-fold cross validation as before, but now instead of using all the nine folds (450 samples) for training, we only use a subset of  $N < 450$  of these samples. To average out the dependence on the sampling of the subset, we repeat the evaluation of the test set 50 times, where in each time we randomly choose  $N$  samples out of the available training set of 450 samples to build the database. Both AvgPow and MinMisProb show a similar trend, and we show only the results for AvgPow here.

We evaluate the quality of the fingerprint obtained using the training set of size  $N$  by the average of the power loss probabilities estimated by the 50 cross validations as described earlier. Fig. 3.5 shows the average 3 dB power loss probabilities for the training sample size  $N$  ranging from 50 to 250. We can see a large improvement when increasing  $N$  from 50 to 90. Subsequent increases in  $N$ , however, provide diminishing improvement. To see this effect more clearly, we plot in Fig. 3.5(b) the 3 dB power loss probability when the number of beam pairs trained is fixed at  $N_b = 30$ . This plot is typically known as the learning curve [45], which quantifies the improvement as the training sample size (i.e., the learning effort) increases. The figure shows the mean and the region of one standard deviation from the mean. We see a sharp



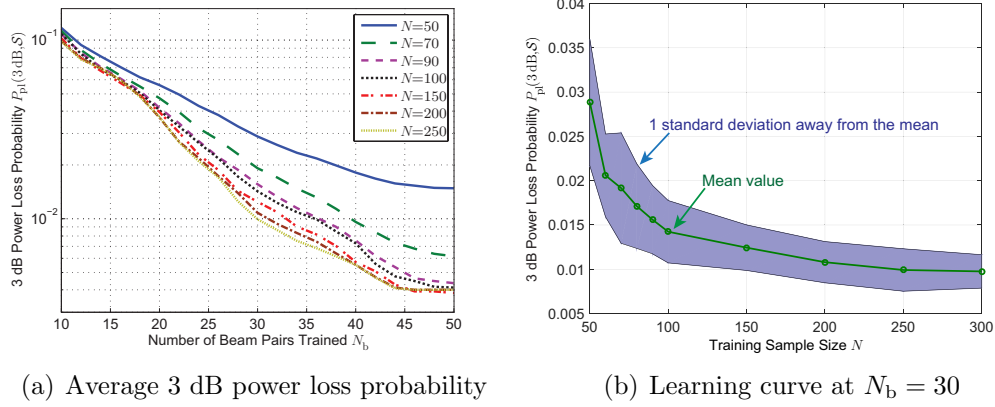


Figure 3.5: 3 dB power loss probability of AvgPow as a function of the training sample size  $N$ . Fig. 3.5(a) shows the average 3 dB power loss probability for different  $N$ . Fig. 3.5(b) shows the learning curve in terms of the 3 dB power loss probability when the number of beam pairs measured is set to  $N_b = 30$ . We can see from the plots that the improvement diminishes as the training sample size  $N$  increases beyond about 100.

improvement up to around  $N = 100$ , and a slower improvement beyond that. We conclude that a training sample size of around 250 seems good enough.

### 3.7.3 Effect of Location Bin Size

We start with the description of the evaluation method. We use 10-fold cross validations on 300 channel samples where the CV is in the location bin. Recall from Section 3.3.1 that the channel samples are generated with the center of the location bin at  $d_0 = 30$  m and the CV is randomly placed in the range  $[d_0 - 2.5, d_0 + 2.5]$ . For example, when evaluating the bin size of 2 m, we only use the channel samples where the CV's center position (the position of its antenna) is within  $[d_0 - 1, d_0 + 1]$ . Since each cross validation

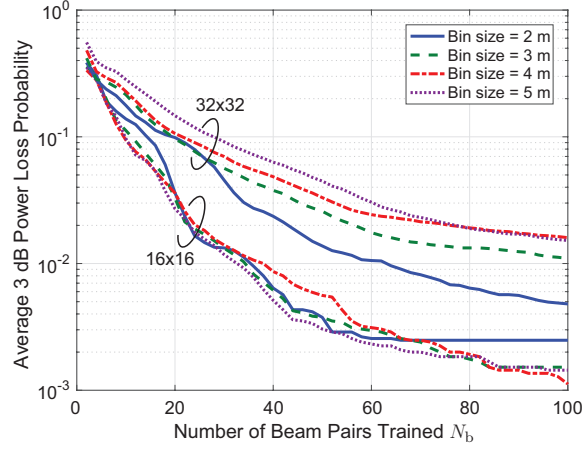


Figure 3.6: Average 3 dB power loss probability when using different location bin sizes. All bin sizes performs similarly when using  $16 \times 16$  arrays. When using the larger 32 arrays, smaller bin sizes improve the performance.

is computed on a different dataset, we need to perform averaging to eliminate dependency on the dataset. To do this, we repeat the cross validation 50 times, where each time 300 channel samples are randomly selected from the set of the channel samples where the CV is within the location bin. To be able to evaluate small bin sizes, we generated more channel samples to a total of 1000 samples. The choice of the sample size of 300 is to ensure that we have enough channel samples for building the database while not too large so that each set of 300 samples selected for each cross validation is random enough. For the comparison metric, we use the average 3 dB power loss probability, which is obtained by averaging over the results from the 50 cross validations. AvgPow is used as the beam selection method.

Fig. 3.6 shows the average 3 dB power loss probability for location bin

sizes of 2, 3, 4, and 5 m when using UPA  $16 \times 16$  and  $32 \times 32$ . The variation of the bin size from 2 m to 5 m has little impact when using the  $16 \times 16$  arrays. We note that the plots are in log-scale and the gap at  $N_b = 100$  is small (it is less than 0.002). When using a larger array of  $32 \times 32$ , which requires higher location precision, we can see that a smaller bin size provides better performance. The physical reason why the performance is not that sensitive to the location bin size might be that NLOS paths are less affected by the position accuracy than a LOS path. This is because NLOS paths have nonzero angular spread which makes it easier for beam alignment. Instead of having to align to a single direction as in the LOS path, the beam can be aligned to within the range of the angular spread. From these results, we can conclude that location bin size of 5 m is good enough when using UPA  $16 \times 16$ . When using a large array such as  $32 \times 32$ , smaller bin sizes provide better performance. Finally, note that while smaller bin sizes are preferred for beam alignment performance, it has to be large enough to account for the level of position accuracy available to the vehicles.

#### 3.7.4 Effect of Measurement Noise

In this subsection, we study the effect of measurement noise on the beam alignment accuracy. The results are shown in terms of the Equivalent Isotropic Radiated Power (EIRP), which is the transmit power plus the transmit antenna gain (in dB scale). EIRP is used instead of the transmit power because it is regulated [107]. To provide the context of the operating SNR,

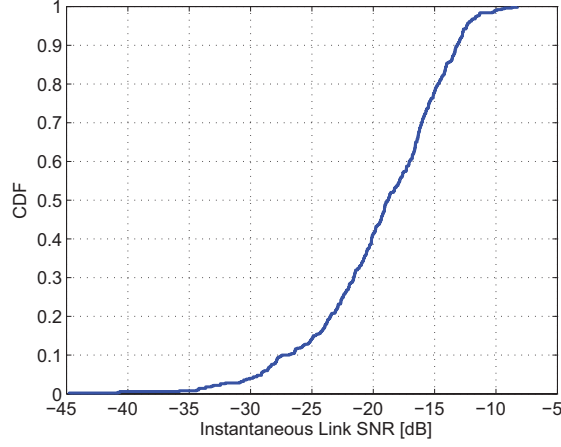


Figure 3.7: CDF of the link SNR of the generated channels. The link SNR is defined as the SNR at the receiver when the transmit power is 0 dBm and both the transmitter and receiver use isotropic antennas. The average link SNR is -16.0 dB.

we start with a description of the link SNR of the channel samples generated from the ray-tracing simulation. We assume the noise power is given by  $P_n = -174 + 10 \log_{10} B$  dBm, where  $B = 1760$  MHz is the sampling rate defined in IEEE 802.11ad for single carrier PHY [50]. Denote  $P_0$  the received power when isotropic antennas (0 dBi antenna gain) are used at both the transmitter and the receiver with 0 dBm transmit power, the link SNR is defined as

$$\text{SNR} = \frac{P_0}{P_n}. \quad (3.23)$$

Fig. 3.7 shows the CDF of the link SNR computed from the received powers output from the ray-tracing simulation. We note that with an EIRP of 0 dBm, the SNR at the receiver (before the receive antenna gain) is the link SNR. The average link SNR is -16.0 dB, and thus the average SNR at the receiver is

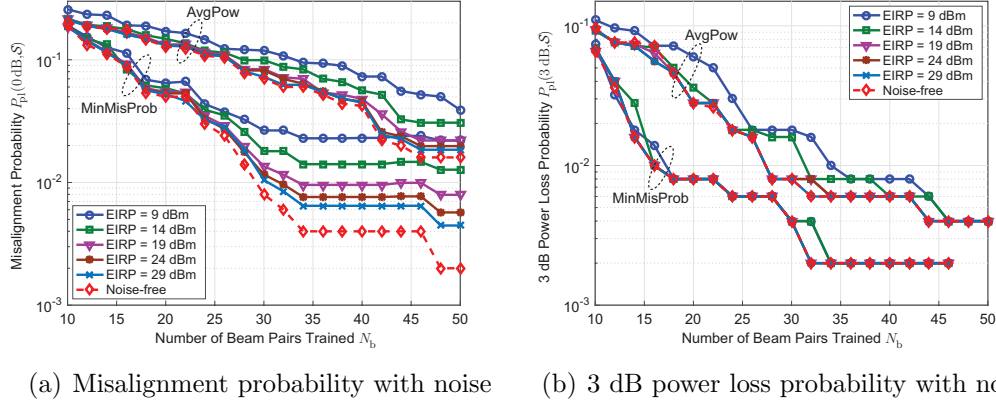


Figure 3.8: Power loss probability as a function of  $N_b$  in the presence of noise. The noise impacts the misalignment probability much more than the 3 dB power loss probability.

around 0 dB when using an EIRP of 16 dBm.

Fig. 3.8 shows a comparison of the power loss probability with and without noise for the two beam selection methods. It is assumed that the same EIRP is used during the database collection and when doing beam measurements for beam alignment. We can see that the misalignment probability is much more affected than the 3 dB power loss probability. This is because for a 3 dB power loss event to happen the noise must be large enough to flip the order of the optimal pair with a beam pair that has the power of at least 3 dB below the optimal beam pair, which occurs much less frequent than the misalignment event (i.e., any nonzero power loss event). We note that the plots are in log-scale, and the gaps in the misalignment probability between the EIRP= 9 dBm case and the noise-free case are 0.03 and 0.02 at  $N_b = 50$  for AvgPow and MinMisProb, respectively. Thus, overall we can conclude that

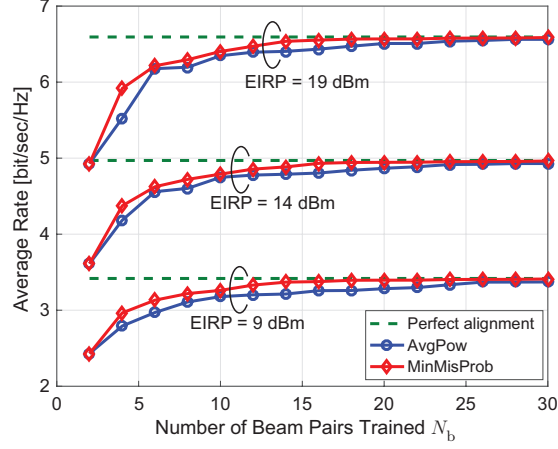


Figure 3.9: Average rate of the proposed beam alignment compared to the perfect alignment case. MinMisProb consistently achieves a higher average rate than AvgPow for the same  $N_b$ . The rate loss becomes negligible at  $N_b$  at around 20 and 30 for MinMisProb and AvgPow, respectively.

MinMisProb is less affected by noise than AvgPow.

We show the average rates when using the proposed beam alignment in Fig. 3.9. The instantaneous rate is computed using

$$R_{\text{ins}} = \log_2 \left( 1 + \frac{P_t \|\mathbf{h}_s\|^2}{P_n} \right), \quad (3.24)$$

where  $\mathbf{h}_s$  is the effective channel of the beam pair selected after the beam training. Fig. 3.9 shows the average rate as a function of the number of beam pairs trained  $N_b$  for EIRP of 9, 14, and 19 dBm. Increasing the training overhead  $N_b$  improves the alignment quality leading to higher average rates. The rate loss compared to the perfect alignment case becomes negligible at around  $N_b = 20$  for MinMisProb and at around  $N_b = 30$  for AvgPow. The gaps to the perfect alignment at small  $N_b$  do not improve with increasing

EIRP. This is due to the larger power loss probability when using a small  $N_b$  and so cannot be eliminated by increasing the transmit power.

### 3.7.5 Effect of Traffic Mismatch during Database Collection and Exploitation

In this subsection, we provide some simulation results to show the effect of the mismatch of the traffic density during the database collection and exploitation. We expect the effect to be more pronounced when the difference in traffic density is large. We, thus, study the high and low traffic density cases. To do this, we generate another dataset of 500 channel samples using the ray-tracing simulation with the same setting as described in Section 3.3.1 but with a lower vehicular traffic using  $\mu_\zeta = 0.0536$  (average vehicle gap of 18.66 m) and the average cars-to-trucks ratio of 9:1. We use all combinations of these two datasets as the training and test set to evaluate the performance of the proposed beam alignment, namely the four combinations of training and test set of (low,low), (high,low), (low,high), and (high,high). We note that (high,high) is what is used so far. AvgPow is used as the selection method.

Fig. 3.10 shows the performance in terms of the 3 dB power loss probability and the average rate normalized by the perfect beam alignment case when using EIRP of 24 dBm. When exploiting in the low traffic setting, training (i.e., building the fingerprint database) in either the low or high traffic condition yields good performance while training in the low traffic density (i.e., matched traffic condition) is slightly more efficient. On the contrary,

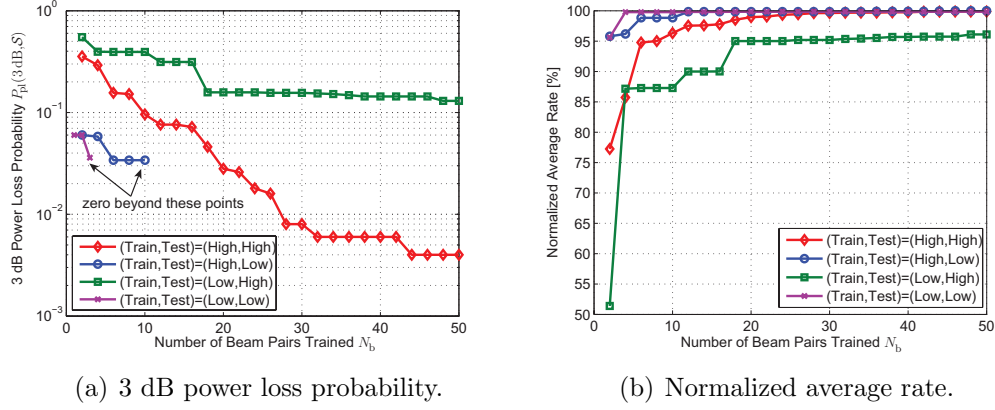


Figure 3.10: Effect of the mismatch in traffic density during database collection and exploitation. Database collected in a light traffic does not work well when used in a dense traffic because the database cannot capture all the paths in the richer scattering environment of the dense traffic. On the contrary, database collected in a dense traffic works well regardless of the traffic condition when it is exploited. It only has a slightly degraded efficiency when used in a low traffic condition.

when exploiting in the high traffic condition, the performance loss due to the mismatch in the traffic conditions during database collection and exploitation is higher. Intuitively, this is because the database collected under a low traffic density cannot adequately capture paths in the richer scattering environment of the high traffic condition. The same trend can be confirmed in the rate plot in Fig. 3.10(b). We note that the largest loss observed is around 5% when using the database collected in the low traffic condition and used in the high traffic case. These results show that building the database in the same traffic condition as when the database is used provides the best performance. If adaptation to the traffic condition is not possible or costly, then the database should be collected in high traffic conditions.



### 3.7.6 Comparison with Existing Techniques

In this subsection, we compare the performance of the proposed method with two existing solutions. The first one is a hierarchical beam search adopted in IEEE 802.11ad, and the second one uses only the position to determine the beam pointing direction.

We start with a comparison with the IEEE 802.11ad method. IEEE 802.11ad beam alignment is a beam sweeping method using a hierarchical beam codebook structure to reduce the amount of beam training [50, 58]. It is required by the standard that the antenna gain of the quasi-omni pattern (the widest level in the codebook) be at most 15 dB ( $\simeq 32$  in linear scale) lower than a directional pattern [50, Section 21.10.1]. Because of this constraint, we consider a two-level beam codebook: the quasi-omni and the sector level. We further assume for simplicity that the number of codewords at the sector level is equal to  $N_a$ , the number of elements of the array (e.g., when using a 2D DFT codebook). Since the gain in the main beam direction of an array is  $N_a$ , we have  $N_{\text{sec}} = N_a$  and  $N_{\text{QO}} = N_{\text{sec}}/32$ , where  $N_{\text{sec}}$  is the number of sector beams and  $N_{\text{QO}}$  is the number of quasi-omni patterns. Since the quasi-omni patterns are the widest in the codebook, an exhaustive search is needed at the quasi-omni level to determine the best quasi-omni pair. Once the best quasi-omni pair has been identified, IEEE 802.11ad uses a low complexity single-sided search to find the best receive and transmit sectors. A single-sided search is conducted by having the transmitter use the best transmit quasi-omni pattern while the receiver sweeps over the sectors whose pointing directions are within

the best receive quasi-omni pattern. The same procedure is used to find the best transmit sector. The total beam training time (excluding feedbacks) is given by

$$T_{11\text{ad}} = N_{\text{QO}}^2 T_{\text{QO}} + 2 \frac{N_{\text{sec}}}{N_{\text{QO}}} T_{\text{sec}}, \quad (3.25)$$

where  $T_{\text{QO}}$  and  $T_{\text{sec}}$  are the duration of a training frame at the quasi-omni and the sector level, respectively. The second term in (3.25) is based on the assumption that each quasi-omni pattern covers the same number of sectors. The factor two is because the single-sided search has to be conducted for both the transmitter and the receiver.

We note that quasi-omni patterns have low antenna gain and thus require a large spreading factor to compensate for the lack of antenna gain. IEEE 802.11ad uses  $32\times$  spreading for this. A beam training at the quasi-omni level is done by sending an SSW (sector sweep) frame of length  $26.8 \mu\text{s}$ , which consists of  $4.3 \mu\text{s}$  for the preamble and  $22.5 \mu\text{s}$  for the header and information in the SSW frame. Since the preamble might be needed for synchronization, we assume that it is unchanged and set  $T_{\text{sec}} = 4.3 + 22.5/32 = 5.0 \mu\text{s}$ .

We now compute the overhead of the proposed approach. As discussed in Section 3.4.1, the fingerprint database is expected to be valid for a long period of time and thus the cost per usage after the database is collected will be negligible. We, therefore, consider only the beam training overhead here. We define the overhead as the smallest number of beam pairs trained  $N_{\text{fp}}$  needed to achieve  $P_{\text{pl}}(3\text{dB}, \mathcal{S}) \leq 1\%$ . Since our approach does not use wide

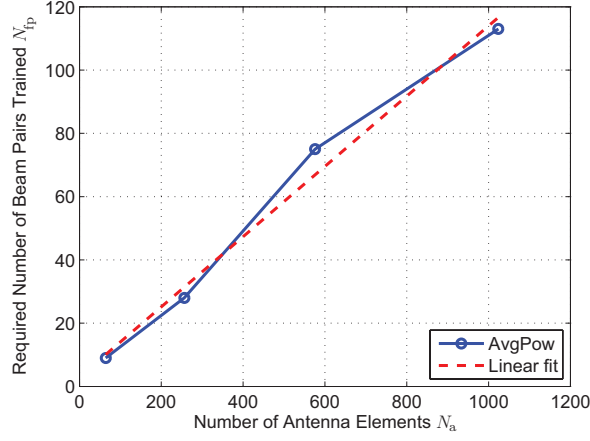


Figure 3.11: Required amount of beam training of the proposed method. The overhead increases roughly linearly with the number of antenna elements of the array.

beams for the beam training, we assume the training duration to be  $T_{sec}$ , and the total training time is

$$T_{fp} = N_{fp} T_{sec}. \quad (3.26)$$

Fig. 3.11 shows the required amount of beam training  $N_{fp}$  for UPAs of sizes  $8 \times 8$ ,  $16 \times 16$ ,  $24 \times 24$ , and  $32 \times 32$  when using AvgPow as the beam pair selection method. The codebook sizes are 87, 271, 641, and 1047, respectively. The plot shows  $N_{fp}$  as a function of the number of elements  $N_a$ , which shows a roughly linear trend in  $N_a$ .

To understand this overhead in the mobility context, we leverage the concept of beam coherence time from Chapter 2, which is the duration before beam realignment is required. The beam coherence time is the duration that the pointing error due to mobility causes the received power to drop by some

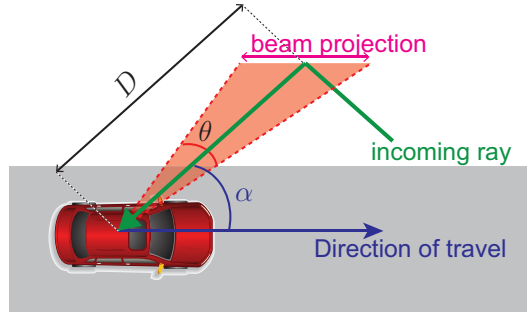


Figure 3.12: An illustration of the beam coherence time concept. For the beam to stay aligned, the reflection point must be within the beam projection. The beam coherence time is the average time that the reflection point is within the beam projection.

threshold from the peak. Since the codebook quantizes the angular domain by the 3 dB beamwidth, the initial pointing error ranges in  $[-\Theta/2, \Theta/2]$ , where  $\Theta$  is the 3 dB beamwidth. Assuming the initial pointing error to be uniform in  $[-\Theta/2, \Theta/2]$ , then the beam coherence time is the duration that the pointing error becomes larger than half the 3 dB beamwidth. Using the pointing error derived in Chapter 2, the beam coherence time  $T_B$  can be written as

$$T_B = \frac{D}{v \sin \alpha} \frac{\Theta}{2}, \quad (3.27)$$

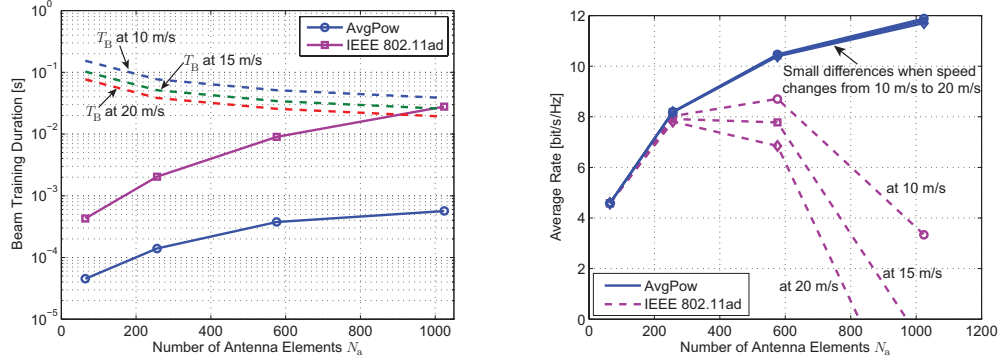
where  $D$  is the distance to the reflector/scatter,  $\alpha$  is the main beam direction with respect to the direction of travel, and  $v$  is the speed of the CV (see Fig. 3.12). We note that  $D$  refers to the distance from the receiver to the nearest reflector/scatter, and this concept can be applied to channels with high orders of reflections. While this result is based on a 2D model in the azimuth plane, it is applicable here because there is negligible change in the elevation angle as the vehicle moves. Since we are using a square array with a total

number of elements of  $N_a$ , the azimuth beamwidth  $\Theta$  can be approximated by the beamwidth of a uniform linear array of size  $\sqrt{N_a}$  given by  $\Theta \simeq 0.886 \frac{2}{\sqrt{N_a}}$  when the antenna spacing is set to half the wavelength [82, p. 885]. Using this approximation, we have

$$T_B(N_a) = \frac{0.886D}{v\sqrt{N_a} \sin \alpha}. \quad (3.28)$$

From the geometry, we set  $\alpha = 60^\circ$  and  $D = 12$  m, which represents a typical reflection path off the buildings in Fig. 3.1. These parameters are chosen as a representative worst-case example with a reasonably small distance to the nearest reflector (a building in this case) and a reasonably large  $\alpha$ . The worst case value for  $\alpha$  is  $90^\circ$ , but this path direction is unlikely in the geometry shown in Fig. 3.1.

Fig. 3.13 compares the overhead of the proposed beam alignment using AvgPow selection method with that of IEEE 802.11ad. Fig. 3.13(a) shows beam training durations as a function of the array size  $N_a$ . We recall that this beam training duration does not include the initial training request and feedback, which do not depend on the array size. We also plot the beam coherence times  $T_B$  when the vehicle speed is 10, 15, and 20 m/s. The training duration of the proposed method is at most a few percents of  $T_B$ , while that of IEEE 802.11ad can exceed  $T_B$  when the array becomes large enough. This means that IEEE 802.11ad cannot finish the training before realignment is required. Fig. 3.13(b) compares the average rate when taking the beam training duration and  $T_B$  into account assuming a transmit power of 0 dBm



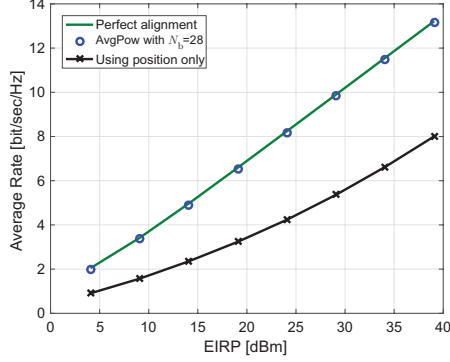
(a) Beam training duration versus the number of antenna elements. (b) Average rate considering beam training overhead (transmit power of 0 dBm).

Figure 3.13: Overhead comparison between the proposed method and that of IEEE 802.11ad. The beam training time of the proposed method only takes up to a few percent of the beam coherence time  $T_B$ , while that of IEEE 802.11ad can exceed  $T_B$  when using a large array leaving no time for data communications.

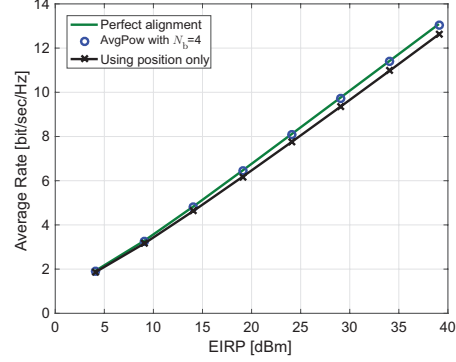
(corresponding to 24 dBm EIRP when using a  $16 \times 16$  array). The average rate here is defined as

$$R_{\text{avg}} = \frac{T_B - T_{\text{trn}}}{T_B} R_{\text{trn}}, \quad (3.29)$$

where  $T_{\text{trn}}$  is the beam training duration and  $R_{\text{trn}}$  is the average rate after the alignment. For the proposed method,  $R_{\text{trn}}$  is determined from rate plots when using the different array sizes (Fig. 3.9 shows rate plots when using  $16 \times 16$  arrays). We assume optimistically that IEEE 802.11ad achieves the perfect alignment rate. While the average rate of the proposed beam alignment keeps increasing as the array size increases, that of IEEE 802.11ad increases slowly at 10 m/s or starts to decrease at speed beyond 15 m/s when the array becomes larger than  $16 \times 16$ . We also observe that the average rate of IEEE 802.11ad



(a) High traffic density.



(b) Low traffic density.

Figure 3.14: Average rate comparison with the beam alignment method using position only. The method using only the position does not need beam training but works only when the LOS path is available. It performs well in a low traffic density where the LOS path is often available but performs poorly in a dense traffic where the LOS path is often blocked. Taking advantage of the fingerprint database, the proposed method works well in both cases.

becomes zero at  $N_a$  around 800 and 1000 when the speed is 15 and 20 m/s, respectively. This is because the training duration becomes larger than  $T_B$ , and there is no time left for data communication.

We now compare the performance of the proposed beam alignment with the method that uses position only. Fig. 3.14(a) and 3.14(b) show the average rate of the proposed method and the beam alignment using position only in high and low traffic conditions. The datasets used are the same as described in Section 3.7.5. UPA  $16 \times 16$  and 24 dBm EIRP are assumed. AvgPow is used and the number of beam pairs trained is the smallest  $N_b$  such that  $P_{pl}(3 \text{ dB}, \mathcal{S}) \leq 1\%$ . The average rates accounting for the beam training overhead and the beam coherence time  $T_B$  are computed using (3.29) assuming

a speed of 20 m/s. By using only the position, the beam training can be eliminated, but the method only works when the LOS path is available. This means that it will perform poorly when the LOS path is frequently blocked. We can confirm this effect in Fig. 3.14. In the high traffic case, there are many trucks on the street which often block the LOS directions leading to a large gap compared to the proposed method. The gap becomes smaller in the low traffic case because of the less blockage. Leveraging the fingerprints database, the proposed method works in all traffic conditions. The benefit of fingerprints increases with the traffic density, or more generally the blockage probability of the LOS path.

### 3.8 Learning-to-Rank Beam Pair Subset Selection

In this section, we develop a new beam pair selection method using a ML approach. Specifically, we use the LtR approach from the recommender system literature [72]. We start by motivating the potential benefit of an ML-based approach as compared to the context binning approach used in our statistical learning methods discussed so far. Then, we will describe the LtR method for beam pair selection, and conclude this section with a numerical evaluation comparing it with the MinMisProb beam pair selection method.

Before moving on, we need to make a clarification to the change of the training data. Our goal is to use the context information directly without the need to first quantize or bin it. To allow this, each record in the training data is tagged with the position of the contributing vehicle. An example is shown in



Table 3.4: An example of the database used for LtR beam pair selection. The  $\ell$ -th row corresponds to measurement data contributed by a car at position  $\mathbf{x}_\ell$ . This is the same as Type A fingerprint database in Table 3.1 with the position tagged to each record.

Position	Best	2nd best	...	$k$ -th best
$\mathbf{x}_1$	5 -64.5 dBm	159 -69.2 dBm	... ...	346 -95.8 dBm
$\mathbf{x}_2$	159 -70.4 dBm	263 -72.6 dBm	... ...	354 -97.1 dBm
...	...	...	...	...
$\mathbf{x}_N$	5 -66.4 dBm	258 -68.1 dBm	... ...	2 -82.6 dBm

Table 3.4, which is the same as the Type A fingerprints except for the position information. Here, we assume there is no error in the position information. This is not a realistic assumption at the moment, but future autonomous vehicles likely could have access to centimeter-level accuracy with negligible error (e.g., using LIDAR).

### 3.8.1 Scalability Issues of Context Binning

While our statistical approach provides a clear interpretation, it does not scale well when different types of context information, such as time of the day (to infer traffic density) and antenna position, are included in the learning. Assuming there are  $c$  contexts and each context is quantized into  $b$  bins, then there are  $b^c$  models to construct and maintain, which is an exponential function. Also, if precise context information (e.g., centimeter-level accurate position information) is available, to make full use of that accurate

information, the context bins must be small, i.e.,  $b$  has to be large. This is another source of scalability issue. Also note that by treating each context bin independently, we ignore the relationship between context bins and thus effectively throw away some information. A more scalable approach is to use the context information directly without binning. We propose a LtR approach as a solution in this direction.

### 3.8.2 LtR Beam Pair Selection Method

The idea of this approach is to learn a scoring function  $\hat{z}(\cdot)$  that can predict the scores of beam pairs and provide a means to rank them. For this purpose, we need to define a feature vector, which is the input to the scoring function, that distinguishes pointing directions. In particular, the feature vector corresponding to the  $i$ -th beam pair at position  $\mathbf{x}_\ell$  is defined as

$$\mathbf{q}_{i\ell} = [(\boldsymbol{\Theta}_i^{\text{rx}})^T \quad (\boldsymbol{\Theta}_i^{\text{tx}})^T \quad \mathbf{x}_\ell^T]^T, \quad (3.30)$$

where  $\boldsymbol{\Theta}_i^{\text{rx}} = [\theta_i^{\text{rx}}, \phi_i^{\text{rx}}]^T$  and  $\boldsymbol{\Theta}_i^{\text{tx}} = [\theta_i^{\text{tx}}, \phi_i^{\text{tx}}]^T$  are the azimuth and elevation of the main beam direction of the  $i$ -th receive and transmit beam pair. *Note that should other context information become available, it can be appended to the feature vector in (3.30).* Assuming that  $\hat{z}(\cdot)$  is already learned, Fig. 3.15 shows how the beam recommendation works. In the rest of the subsection, we describe how to learn  $\hat{z}$ .

We start with the description of our model for the scoring function  $\hat{z}(\cdot)$ . Because our feature vector is short compared to typical learning settings, we

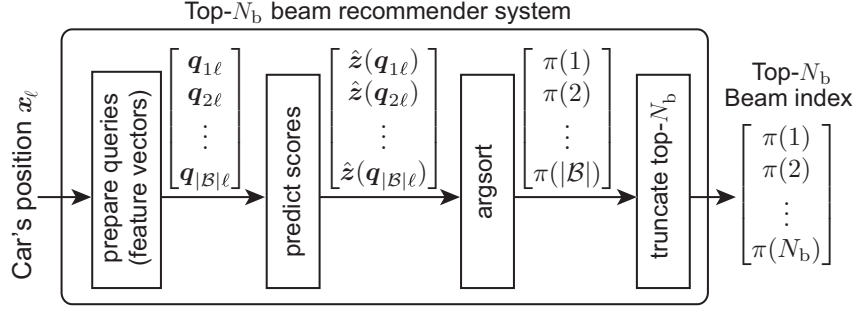


Figure 3.15: Top- $N_b$  recommendation process at the car's position  $\mathbf{x}_\ell$ . First, the position  $\mathbf{x}_\ell$  is used to produce a set of queries for the pointing directions of interest. Then, using the learned scoring function  $\hat{z}(\cdot)$ , the system predicts the scores of these queries points  $\mathbf{q}_{i\ell}$ . The scores are then sorted to produce the beam pair indices.  $\pi(n)$  denotes the beam index with the  $n$ -th highest predicted score. The top- $N_b$  indices are the output of the beam recommender system.

use a kernel-based model which can provide higher model expressibility. Specifically, the scoring function is modeled as

$$\hat{z}(\mathbf{q}) = \sum_{\ell=1}^N \sum_{i \in \mathcal{J}_\ell} w_{i\ell} \kappa(\mathbf{q}_{i\ell}, \mathbf{q}), \quad (3.31)$$

where  $\mathcal{J}_\ell$  denotes the set of beam pairs measured at location  $\mathbf{x}_\ell$  in the training set,  $\kappa(\cdot)$  is the kernel function, and  $w_{i\ell}$  are the parameters to be learned. For notational convenience, we introduce an index mapping function  $u(\cdot)$  that maps  $(i, \ell)$  uniquely to the set  $\{1, 2, \dots, N_{\text{tot}}\}$ , where  $N_{\text{tot}} = \sum_{\ell=1}^N |\mathcal{J}_\ell|$  is the total number of beam measurements in the database. Now let  $\alpha_{u(i, \ell)} = w_{i\ell}$  and  $\tilde{\mathbf{q}}_{u(i, \ell)} = \mathbf{q}_{i\ell}$ , we can rewrite  $\hat{z}(\mathbf{q})$  as

$$\hat{z}(\mathbf{q}) = \sum_{n=1}^{N_{\text{tot}}} \alpha_n \kappa(\tilde{\mathbf{q}}_n, \mathbf{q}). \quad (3.32)$$

We denote  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_{N_{\text{tot}}}]^T$ . Here, we use a modified radial basis function (RBF) as the kernel defined as

$$\kappa(\mathbf{q}_{i\ell}, \mathbf{q}) = \exp\left(-\frac{d_{\text{gc}}^2(\boldsymbol{\Theta}_i^{\text{rx}}, \boldsymbol{\Theta}^{\text{rx}})}{\sigma_{\text{rx}}^2}\right) \exp\left(-\frac{d_{\text{gc}}^2(\boldsymbol{\Theta}_i^{\text{tx}}, \boldsymbol{\Theta}^{\text{tx}})}{\sigma_{\text{tx}}^2}\right) \exp\left(-\frac{\|\mathbf{x}_\ell - \mathbf{x}\|^2}{\sigma_{\text{x}}^2}\right), \quad (3.33)$$

where  $\sigma_{\text{rx}}, \sigma_{\text{tx}}$ , and  $\sigma_{\text{x}}$  are the kernel parameters, and  $d_{\text{gc}}(\cdot)$  denotes the great circle distance on a unit sphere. We use  $d_{\text{gc}}(\cdot)$  because it properly handles the cyclic property of angles. Note that as in any learning methods, the hyperparameters such as the kernel parameters will need to be tuned. The intuition for selecting this kernel is that each component of the feature vector is a different type of physical feature, and thus they should be scaled by different kernel parameters.

We now describe the objective function that will be used for learning the scoring function. Inspired by the discounted cumulative gain (DCG) [72], a popular metric for evaluating ranked lists, we measure the quality of the predicted ranked list at position  $\mathbf{x}_\ell$  by

$$G_\ell(\boldsymbol{\alpha}) = \frac{1}{|\mathcal{J}_\ell|} \sum_{i \in \mathcal{J}_\ell} \sum_{j \in \mathcal{J}_\ell} S(\delta_{j\ell}) \mathbf{1}[R_{j\ell}(\boldsymbol{\alpha}) \leq R_{i\ell}(\boldsymbol{\alpha})], \quad (3.34)$$

where  $R_{i\ell}(\boldsymbol{\alpha})$  is the predicted rank of the  $i$ -th beam pair,  $\delta_{j\ell}$  is some measure of “goodness” of the  $j$ -th beam pair at this location,  $S(\cdot)$  is some function to transform the raw goodness metric. The inner sum in (3.34) can be thought of as the overall goodness of the beam pairs with predicted rank from 1 to  $R_{i\ell}(\boldsymbol{\alpha})$ . Following the concept of power loss introduced earlier, we define  $\delta_{j\ell}$  as

$$\delta_{j\ell} = \frac{\gamma_{j\ell}}{\gamma_{\text{max},\ell}}, \quad (3.35)$$

---

**Algorithm 3.2** SGD on  $G_{\text{tot}}(\boldsymbol{\alpha})$ 

---

```
1: while not converge or maximum number of iterations not yet reached do
2:   for each randomly selected training located at index  $\ell$  do
3:      $\boldsymbol{\alpha}^{t+1} \leftarrow \boldsymbol{\alpha}^t + \eta_t(\nabla_{\boldsymbol{\alpha}} G_{\ell}(\boldsymbol{\alpha}^t) - \lambda \|\boldsymbol{\alpha}^t\|)$ 
4:      $t \leftarrow t + 1$ 
5:   end for
6: end while
```

---

where  $\gamma_{\max, \ell}$  and  $\gamma_{j\ell}$  are the channel strength (linear scale) of the optimal beam pair and the  $j$ -th beam pair at position  $\mathbf{x}_{\ell}$ . With this definition,  $\delta_{j\ell}$  is close to 1 if the  $j$ -th beam pair is “good” and close to 0 if it is “bad.” Note that  $\delta_{j\ell}$  is not a function of the learning model parameter  $\boldsymbol{\alpha}$  and is obtained from the training data. The goodness of fit evaluating using the metric in (3.34) on the training data is the sum over all the positions in the training data, i.e.,

$$G_{\text{tot}}(\boldsymbol{\alpha}) = \sum_{\ell=1}^N G_{\ell}(\boldsymbol{\alpha}). \quad (3.36)$$

The model parameters  $\boldsymbol{\alpha}$  can be learned by maximizing  $G_{\text{tot}}(\boldsymbol{\alpha})$  over  $\boldsymbol{\alpha}$ . To prevent overfitting, we introduce the 2-norm regularization on  $\boldsymbol{\alpha}$  and obtain

$$\max_{\boldsymbol{\alpha} \in \mathbb{R}^{N_{\text{tot}}}} G_{\text{tot}}(\boldsymbol{\alpha}) - \frac{\lambda}{2} \|\boldsymbol{\alpha}\|^2, \quad (3.37)$$

where  $\lambda$  is the regularization parameter.

The next step is to solve (3.37). Since  $G_{\text{tot}}(\boldsymbol{\alpha})$  is a sum of  $G_{\ell}(\boldsymbol{\alpha})$ , we can apply a stochastic gradient descent (SGD) method over the location index  $\ell$ . We choose SGD because it is a computationally tractable algorithm [45]. Denoting  $\eta_t$  the learning rate at update round  $t$ , the algorithm is shown in Algorithm 3.2. To use Algorithm 3.2, we need to compute the gradient of

$G_\ell(\boldsymbol{\alpha})$  with respect to  $\boldsymbol{\alpha}$ . Unfortunately, the indicator function in (3.34) is not continuous and some relaxation is required. Following [117], we approximate the indicator function by

$$\mathbf{1}[R_{j\ell}(\boldsymbol{\alpha}) \leq R_{i\ell}(\boldsymbol{\alpha})] \simeq g(\hat{z}(\mathbf{q}_{j\ell}) - \hat{z}(\mathbf{q}_{i\ell})), \quad (3.38)$$

where  $g(t) = \frac{1}{1+e^{-t}}$  is the logistic function. With this relaxation, the gradient  $\nabla_{\boldsymbol{\alpha}} G_\ell(\boldsymbol{\alpha})$  can be computed as

$$\nabla_{\boldsymbol{\alpha}} G_\ell(\boldsymbol{\alpha}) \simeq \frac{1}{|\mathcal{J}_\ell|} \sum_{i \in \mathcal{J}_\ell} \sum_{j \in \mathcal{J}_\ell} S(\delta_{j\ell}) g'(\hat{z}(\mathbf{q}_{j\ell}) - \hat{z}(\mathbf{q}_{i\ell})) (\boldsymbol{\kappa}_{j\ell} - \boldsymbol{\kappa}_{i\ell}), \quad (3.39)$$

where  $g'(t) = e^{-t}/(1 + e^{-t})^2$  and  $\boldsymbol{\kappa}_{i\ell} = [\kappa(\tilde{\mathbf{q}}_1, \mathbf{q}_{i\ell}) \ \kappa(\tilde{\mathbf{q}}_2, \mathbf{q}_{i\ell}) \ \dots \ \kappa(\tilde{\mathbf{q}}_{N_{\text{tot}}}, \mathbf{q}_{i\ell})]^\text{T}$ .

### 3.8.3 Numerical Evaluation and Comparison with MinMisProb

We perform a five-fold cross validation on 500 channel samples. Note that the MinMisProb method uses a location bin of size 5 m, while in the LtR approach uses the actual position.  $16 \times 16$  UPAs are assumed at both the CV and the RSU.

We now present the performance results. We evaluate the method using the misalignment probability. The figures show the misalignment probability as a function of the number of beam pairs trained  $N_b$ . All the hyperparameters, including those of the SGD and kernel parameters, are tuned manually. We use a constant learning rate  $\eta = 0.01$  and regularization parameter  $\lambda = 0.001$ , and SGD is run for 20 epochs. The parameters used for our modified RBF kernels are  $\sigma_x = 2$  and  $\sigma_{\text{rx}} = \sigma_{\text{tx}} = 0.1$ . Fig. 3.16 shows the results when using

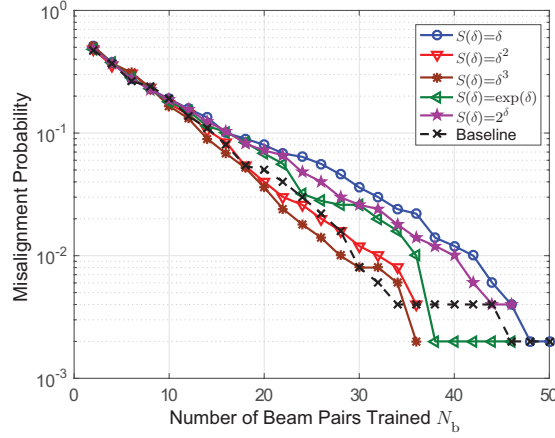


Figure 3.16: Misalignment probability with different transform function  $S(\cdot)$ . A common choice used in web-recommender systems is the exponential function, but it does not work as well as the square and cubic transform. One possible explanation is the nature of the raw scores. In typical web applications, the raw scores are integer ratings (e.g., between 1 and 5), while in our case the scores are real numbers in  $[0,1]$ .

the modified RBF kernels with different choices of  $S(\cdot)$ . The choice of  $S(\cdot)$  is critical, and here  $S(\delta) = \delta^3$  provides the best result. The common choice used in the recommendation context is the exponential function [72]. This difference might be because our raw scores  $\delta_{j\ell}$  are real numbers in  $[0, 1]$  while typical ratings are integers from 1 to 5. Comparing with the MinMisProb, the LtR approach is comparable for  $N_b$  up to around 34 and outperforms the MinMisProb for larger  $N_b$ . In fact, there is no misalignment after  $N_b = 36$ , while the MinMisProb method struggles. We can see that the LtR method can reduce the number of beam pairs trained by up to about 20% for low target misalignment probability such as 0.1%.

Fig. 3.17 shows the results when using different kernel functions. In

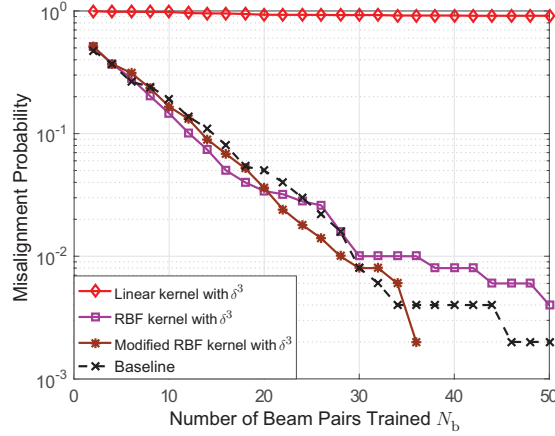


Figure 3.17: Misalignment probability with different kernel functions. We perform feature scaling when using the linear and RBF kernels by dividing each component of the feature vector by the maximum value possible for that feature. The linear kernel does not work at all. The RBF kernel works relatively well but is outperformed by the proposed modified RBF kernel for a large  $N_b$ . Overall, the modified RBF kernel provides the best performance.

all cases,  $S(\delta) = \delta^3$  is used. We compare our modified RBF kernel with the popular choices of linear and RBF kernels. We perform feature scaling by dividing each component of the feature vector by the maximum possible values that the component can take. Note that feature scaling is not necessary for the modified RBF kernel because it can be absorbed into the kernel parameters. There is no parameter for the linear kernel and there is one length scale parameter for the RBF kernels. After manual tuning, this parameter is set to 0.01. We can see that the linear kernel does not work at all. This is likely due to the nonlinearity of the problem. The RBF kernel performs slightly better than the modified RBF kernel at small  $N_b$  but becomes worse at larger  $N_b$ . Overall the proposed modified RBF kernel performs best.



### 3.9 Conclusion

We proposed an efficient beam alignment framework for mmWave V2I communications leveraging position information and multipath fingerprints. We developed three beam pair subset selection methods: two based on statistical learning and one based on ML. Numerical evaluations of the two statistical selection methods show that when Type A fingerprints are available, MinMisProb should be used. If Type A is not available, then our proposed heuristic beam selection, AvgPow, should be used with Type B fingerprints, which only store the average received power. The proposed methods require training less than 30 beam pairs when  $16 \times 16$  arrays are used and the overhead increases roughly linearly with the number of antenna elements. This low overhead allows the use of large arrays such as  $32 \times 32$  at high speed, while existing solution such as that of IEEE 802.11ad cannot handle. Because of the context binning, the number of models in a statistical learning method increases exponentially with the number of contexts while the LtR approach only needs to increase the length of the feature vector. This provides a promising framework to incorporate other contexts beyond position information in future research. The results in this chapter demonstrate that side information can be exploited to improve the efficiency of mmWave communications, which is not only desirable but also necessary in vehicular settings.

## Chapter 4

### Online Beam Pair Selection and Refinement

Accurate beam alignment is essential for beam-based millimeter wave (mmWave) communications. Conventional beam sweeping solutions often have large overhead unacceptable for mobile applications like vehicle-to-everything (V2X). In Chapter 3, we developed learning-based solutions that leverage position and past multipath information to identify good beam directions. While the results are promising for reducing the overhead, the approaches, however, are offline and require collecting the training data beforehand. In this chapter, using the multi-armed bandit (MAB) framework, we develop online learning algorithms for beam pair selection and refinement. The beam pair selection learns coarse beam directions in a predefined beam codebook, and the beam pair refinement fine-tunes the identified coarse directions to adopt the codebook to the environment. The beam pair selection uses the upper confidence bound (UCB) with a newly proposed risk-aware feature, and the beam refinement uses a modified optimistic optimization algorithm. The proposed algorithms show a fast learning behavior. When using  $16 \times 16$  arrays at both the transmitter and receiver, it can achieve on average 1 dB gain over the exhaustive search (over  $271 \times 271$  beam pairs) on the unrefined codebook within 100 time steps with a training budget of only 30 beam pairs. This work

has been submitted to the IEEE for possible publication [108].

## 4.1 Motivation and Prior Work

Fast and efficient beam alignment is crucial in enabling mmWave communications, which is a promising solution to support high data rate demands expected in the fifth generation (5G) cellular networks. One promising approach is to leverage side information from onboard sensors to reduce the alignment overhead. In Chapter 3, we proposed the inverse fingerprinting method that leverages position and past multipath information to identify promising directions among a predefined finite beam codebook using a database collected at a location bin. Although efficient, it has some limitations. First, the approach is offline which requires collecting the database before the deployment. Second, being offline, the performance depends on the accuracy of the database and cannot evolve over time. Online approaches keep collecting new observations during operation, making it possible to improve the database. Third, without any knowledge of the power angular spectrum (PAS), the codebook has to be uniformly spread over the antenna array’s field of view (e.g., in discrete angles separated by the 3 dB beamwidths). At a given location, depending on the scatterers in the environment, the PAS will have peaks at some specific angles. By adapting the beams such that their main beam directions match those peaks, we can expect gains beyond the generic good-for-all-cases codebook. That is, the position-based learning opens up an opportunity to adapt the beam codebook to the environment. In this chapter, we propose an

online beam pair selection and refinement algorithm to address these points.

Beam alignment has been intensively investigated in the literature. Several directions have been pursued such as those based on beam sweeping [58, 112], angle of arrival and departure (AoA/AoD) estimation [32, 73], black-box function optimization [59, 69], and the use of side-information [16, 31, 80, 103]. The last category is most related to our approach, especially those that use position information [16, 31, 103]. We refer to Section 3.1 for a summary of the differences to these different approaches for beam alignment. In this chapter, we develop an online version of the method proposed in Chapter 3 for beam pair selection and proposes a beam pair refinement method to adapt the beam codebook to the environment to further maximize the beamforming gain.

The recent remarkable progress in machine learning (ML) has revived interest in applying ML techniques to communications [23, 55, 118]. By learning from the data, ML can be used to design a system without the need for an explicit model (e.g., [35]). It also opens up opportunities to customize communication systems to the user or environment. Related work that applies ML to beam alignment includes [10, 104, 113]. While promising, these solutions use supervised learning techniques, which assume an offline learning setting. The proposed solutions in this chapter use the MAB framework, which is a special class of reinforcement learning (RL). Recent applications of RL/MAB for beam training include [90, 91] which uses a partially observable Markov decision process (POMDP) framework, and [44] which uses an MAB framework. The work in [90, 91] addresses tracking problems where the POMDP with known state

transition models provides a means to predict the state of the channel enabling an informed choice of the probing beams for good performance. In [44], the beam alignment problem is solved using an MAB framework with the assumption that the success probability (the received power is larger than some threshold) is a unimodal function of the pointing direction. The efficiency of their solution depends on this unimodal property, which cannot be guaranteed in our setting with random blockage.

MAB is a useful tool for modeling sequential decision-making problems with a wide range of applications [19, 21]. The most common form of MAB is the single-play MAB with a finite number of arms, where only one arm is selected in each time step. In the proposed method, multiple beam pairs (up to the training budget) are trained in each beam alignment attempt. Thus, our beam pair selection problem can be cast as a multiple-play MAB problem (also known as combinatorial bandit) [24], where multiple arms may be tried in each round or time step. In sequential decision-making problems, in each round, the player must decide between using the knowledge obtained so far to select the best arm or explore lesser-known arms, which is called the explore-exploit dilemma. The optimism in the face of uncertainty is a core design idea for balancing the explore-exploit tradeoff, based on which the UCB has been developed and results in a widely successful family of algorithms. Our solution employs the UCB in a multiple-play setting. Most related to our solution is the cascading bandit [66], which performs the same selection procedure as our Algorithm 4.1 but with a different model to collect the reward measurements.

Another important difference is that the reduction to the greedy selection is based on the independent arms assumption in [66], while in our case it is based on the exclusive nature of the reward signal. Also, we extend beyond the greedy UCB selection by introducing risk-awareness designed to avoid severe beam misalignment during the learning.

We cast our beam pair refinement as a stochastic continuum-armed bandit (CAB) problem, which has infinitely many arms. CAB assumes the reward function has some smoothness property (e.g., Lipschitz continuous). There are different approaches to solve CAB such as Bayesian optimization (BO) [93], the zooming algorithm [64], and optimistic optimization (OO) [78]. BO does not discretize the arm space but has high complexity. It is more suitable when sampling is expensive or the learning horizon is short. The zooming algorithm and OO rely on smart discretization of the arm space. The zooming algorithm uses an adaptive approach that applies finer discretization in promising region. This is done using an activation rule that is assumed given to the algorithm, but this is a non-trivial problem itself. OO approaches discretize the arm space using a tree and exploit the hierarchy for an efficient search for the best arm. OO approaches designed for stochastic settings include Stochastic Simultaneous OO (StoSOO) and hierarchical OO (HOO) [78]. StoSOO is an explore-first algorithm where the task is to find the best arm given an exploration budget. This does not fit our setting where there is no separate explore and exploit phase. HOO is designed for maximizing the cumulative reward and suits our setting well. Applying HOO in its original

form does not work well. We propose three modifications to suit our beam refinement problem.

## 4.2 Contributions

Our contributions are summarized as follows:

- We propose an online algorithm to learn to select beam pairs with risk-awareness to reduce the probability of severe beam misalignment during the learning. This is done by making the algorithm less likely to select high risk beam pairs. The proposed solution balances the learning burden on early-stage users and the learning speed.
- We provide regret analysis of the proposed algorithms, which provides insights into the cost of the learning due to the introduction of the risk-awareness.
- We formulate the beam pair refinement problem as a CAB problem. Our solution is based on the HOO algorithm [20] with modifications to suit the beam alignment context.
- We integrate the two algorithms into a two-layer online learning solution that learns to select and refine the beam pairs at the same time. The beam pair selection part learns coarse beam directions and the refinement part learns to refine them. This hierarchy is more efficient than learning the refined beam directions directly since now the refinement learning focuses only in promising directions selected by the beam selection part.

Our numerical result shows that the integrated solution learns quickly. Using  $16 \times 16$  arrays at both the transmitter and receiver and a training budget of 30 beam pairs, it achieves an average gain of 1 dB over the exhaustive beam search over  $271^2$  beam pairs in the original codebook without refinement within the first 100 time steps. The gain can reach up to about 1.5 dB over time. Note that unlike prior work [44, 90, 91] that uses simplistic abstract models that match exactly with the underlying statistical assumptions of the problem formulation, we use realistic channels generated by ray-tracing to evaluate our algorithms.

The rest of the chapter is structured as follows. Section 3.3 describes our system model and how we generate the data. Section 4.3 reviews important concepts and notations from Chapter 3, which is the basis for the online learning problem. Section 4.4 describes the proposed two-layer online learning algorithm with the beam pair selection in the first layer and beam pair refinement in the second. The details of the two layers along with analysis are given in Section 4.5 and Section 4.6. Proofs are given in the appendices. Numerical evaluations are given in Section 4.7 followed by the conclusions in Section 4.8.

### 4.3 System Model and Some Background

The system model including the channel and received signal model used here is the exact same as in Chapter 3, and thus we refer the details to that chapter. We recall that our beam codebook consists of beams generated using progressive phase-shift with their main beam directions separated by their 3



dB beamwidths. We note that our learning solutions does not depend on this choice and can work with other codebooks. For convenience, we will provide a brief description of important and relevant notations and concepts from Chapter 3.

We use the power loss probability for quantifying the alignment accuracy. We start with the definition of the power loss. Denote  $\gamma_i = \|\mathbf{h}_i\|^2$  the channel strength with  $\mathbf{h}_i$  the effective channel when the  $i$ -th transmit and receive beam pair is used. Denote  $\mathcal{B}$  the set of all possible beam pairs in the codebook, the power loss when selecting the beam pair  $s$  is defined by

$$\xi = \frac{\max_{i \in \mathcal{B}} \gamma_i}{\gamma_s}. \quad (4.1)$$

The beam pair  $s$  is selected from the selection set  $\mathcal{S}$ , and with accurate beam training  $s = \arg \max_{i \in \mathcal{S}} \gamma_i$ . If the codebook is used without any modification, then  $\mathcal{S} \subset \mathcal{B}$  and  $\xi \geq 1$  always holds. The proposed online learning method also includes a component to refine the beam pairs to adapt the codebook to the environment, in which case  $\xi < 1$  is possible.

The power loss probability is then defined by

$$P_{\text{pl}}(c, \mathcal{S}) = \mathbb{P}[\xi > c], \quad (4.2)$$

for some constant  $c \geq 1$ . We call the case when  $c = 1$  the misalignment probability. A relevant concept to the misalignment probability is the probability of being optimal given by

$$P_{\text{opt}}(\mathcal{S}) = \mathbb{P}[i^* \in \mathcal{S}], \quad (4.3)$$

where  $i^* = \arg \max_{i \in \mathcal{B}} \gamma_i$  denotes the index of the optimal beam pair. We note that

$$P_{\text{opt}}(\mathcal{S}) = \mathbb{P}[\xi = 1] \quad (4.4)$$

$$= 1 - \mathbb{P}[\xi > 1] \quad (4.5)$$

$$= 1 - P_{\text{pl}}(1, \mathcal{S}), \quad (4.6)$$

where (4.5) follows because  $\xi \geq 1$  (without refinement).

We now review two offline statistical beam pair selection methods (AvgPow and MinMisProb) from Chapter 3 that will be used in this chapter. AvgPow is a heuristic that selects the beam pairs by their average channel strengths. Denote  $\bar{\gamma}_i$  the sample average of the channel strength of the  $i$ -th beam pair and  $\arg \max_{i \in \mathcal{B}; M} \{\cdot\}$  the operator that returns the top- $M$  indices, the selection set of size  $|\mathcal{S}_{\text{AP}}| = M$  can be written as

$$\mathcal{S}_{\text{AP}} = \arg \max_{i \in \mathcal{B}; M} \{\bar{\gamma}_i\}. \quad (4.7)$$

MinMisProb is an optimal selection method that minimizes the misalignment probability. Thanks to the modularity of the probability of being optimal (Proposition 3.1 from Chapter 3), MinMisProb is equivalent to the selection by the probability of being optimal. Let  $|\mathcal{S}_{\text{MMP}}| = M$ , then

$$\mathcal{S}_{\text{MMP}} = \arg \max_{i \in \mathcal{B}; M} \left\{ \hat{P}_{\text{opt}}(i) \right\}, \quad (4.8)$$

where  $\hat{P}_{\text{opt}}(i)$  denotes the probability of being optimal of the beam pair  $i$  estimated from the database. In this chapter, we develop an online learning version of MinMisProb while balancing the risk during the learning.

## 4.4 Proposed Two-Layer Online Learning

Our aim in this chapter is to develop an online learning algorithm for fast and efficient beam alignment. We propose a two-layer online solution to achieve this goal. The idea here is to learn coarse beam directions (quantized by the 3 dB beamwidths) that are promising in the first layer and conduct a refinement of those promising directions in the second layer. This kind of hierarchy is efficient because the refinement is only done in promising directions.

An overview of the proposed online learning solution is illustrated in Fig. 4.1. The algorithm runs in an infinite loop, where in each iteration, it recommends a list of beam pairs and updates the learning parameters recorded in the database upon receiving the beam measurements of those pairs. As mentioned earlier, by having the communicating vehicle transmit, there is no extra feedback overhead to collect the beam measurements. We highlight the groups of blocks that correspond to the learning agent and the environment in Fig. 4.1. This shows a typical RL setting where the agent optimizes its action through direct interaction with the environment [94]. In our problem, the action is the subset of beam pairs selected for the training, and the environmental response is the beam measurement results. The algorithm starts by running a detection loop for a request for beam training from the user. If a request is detected, the position (other context can also be used, but we focus on position) is extracted from the training request packet and input to the beam pair selection procedure. Then, the procedure produces a list of beam pairs using the learning parameters corresponding to the location bin

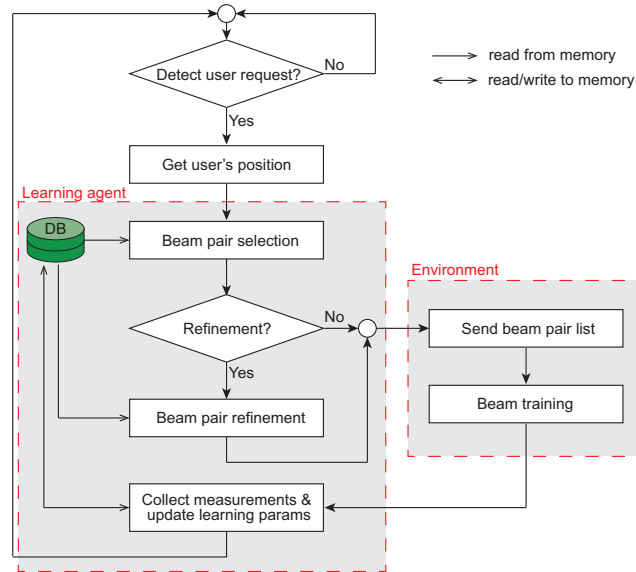


Figure 4.1: A flowchart of the proposed online learning algorithm. The algorithm starts with a training request detection loop. When it detects a request, the algorithm decodes the user's position and input to the beam selection procedure, which then reads the learning parameters corresponding to the position and determines a subset of promising beam pairs. If beam refinement is enabled, the refinement parameters of those selected pairs are selected. The beam subset is then sent to the user and the subset of beam pairs are trained. The beam measurements are used to update the learning parameters and the algorithm returns to the training request detection loop.

stored in the database. If the beam pair refinement is enabled, the refinement parameters of the selected beam pairs are picked by the beam pair refinement procedure. The resulting subset of beam pairs is then sent to the user with an ACK to allow the beam training. After beam training, the measurements of the selected beam pairs are used to update the learning parameters in the database. Then, the algorithm goes back to the detection loop to wait for the next training request.

## 4.5 Online Beam Pair Selection

In this section, we describe the first layer of the two-layer solution. We start with the problem statement. Then, we develop online beam pair selection algorithms, first without and then with risk-awareness. We conclude the section with regret analyses of the proposed algorithms and some discussion.

### 4.5.1 Problem Statement

Our goal here is to develop an online version of the optimal beam pair selection method, MinMisProb. Specifically, the algorithm needs to solve the following optimization problem in an online setting:

$$\begin{aligned} & \underset{\mathcal{S}}{\text{minimize}} && P_{\text{pl}}(1, \mathcal{S}) \\ & \text{subject to} && \mathcal{S} \subset \mathcal{B}, |\mathcal{S}| \leq B_{\text{tr}}, \end{aligned} \tag{4.9}$$

where  $B_{\text{tr}}$  is the desired subset size (the training budget). In an online learning setting,  $P_{\text{pl}}(1, \mathcal{S})$  is not known, and it has to be estimated on the fly. To gain accurate knowledge of the beam pairs, each of them must be sampled multiple times, which means that the learning can be very slow when  $\mathcal{B}$  is a large set, i.e., when large arrays with narrow beams are used. To remedy this problem, we propose to apply a heuristic to screen the beam pairs using a small offline database (of size  $N$ ) to obtain a smaller set  $\hat{\mathcal{B}}$  to apply the learning algorithm on.  $\hat{\mathcal{B}}$  is obtained as the set of the unique beam pairs among the  $NC$  entries of the first  $C$  columns of Table 3.1. In our simulations, the offline database size  $N = 5$  and  $C = 200$  seem to be good enough for this purpose.

### 4.5.2 Greedy UCB Algorithm

We first propose a solution to (4.9) without risk-awareness. A subset  $\mathcal{S}$  can be treated as a super-arm and a single-play MAB algorithm can be used. Such an approach is not efficient because it treats each super-arm as independent and the number of super-arms is large due to the combinatorial nature of the number of all possible subsets.

A more efficient approach to solve (4.9) is to leverage the structure of  $P_{\text{pl}}(1, \mathcal{S})$  to take advantage of the dependence between the subsets. Specifically, we make use of the modularity property of the probability of being optimal (Proposition 3.1 from Chapter 3). We note that by the relationship in (4.6), the problem (4.9) is equivalent to a maximization of  $P_{\text{opt}}(\mathcal{S})$  with the same constraints, i.e.,

$$\begin{aligned} & \underset{\mathcal{S}}{\text{maximize}} && P_{\text{opt}}(\mathcal{S}) \\ & \text{subject to} && \mathcal{S} \subset \hat{\mathcal{B}}, |\mathcal{S}| \leq B_{\text{tr}}, \end{aligned} \tag{4.10}$$

where we also replace  $\mathcal{B}$  by  $\hat{\mathcal{B}}$  as explained earlier. Since  $P_{\text{opt}}(\mathcal{S})$  is modular, it can be decomposed as

$$P_{\text{opt}}(\mathcal{S}) = \sum_{i \in \mathcal{S}} P_{\text{opt}}(i). \tag{4.11}$$

This property is due to the exclusive nature of the events that the  $i$ -th beam pair is optimal (i.e., having the highest channel strength). Recall that the probability of a union of exclusive events is the sum of the probability of each individual event [43]. The main implication of (4.11) is that the reward of  $\mathcal{S}$  can be computed from the individual rewards of each of the beam pairs

---

**Algorithm 4.1** Greedy UCB

---

```
1: // initialize arms' parameters using a small offline database
2:  $X_{\text{tot}}[i] \leftarrow 0$ , for  $\forall i \in \hat{\mathcal{B}}$ 
3:  $X_{\text{tot}}[\arg \max_{i \in \hat{\mathcal{B}}} \bar{\gamma}_i^{\text{init}}] \leftarrow 1$ 
4:  $T_i \leftarrow 1$ , for  $\forall i \in \hat{\mathcal{B}}$ 
5: for  $n = 1, 2, \dots$  do
6:   // Compute UCB values
7:    $\text{UCB}_i \leftarrow \frac{X_{\text{tot}}[i]}{T_i} + \sqrt{\frac{2 \log(n)}{T_i}}$ , for  $\forall i \in \hat{\mathcal{B}}$ 
8:   // Greedy selection using UCB values
9:    $\mathcal{S} \leftarrow \emptyset$ 
10:  for  $k = 1, 2, \dots, B_{\text{tr}}$  do
11:     $\mathcal{S} \leftarrow \mathcal{S} \cup \arg \max_{i \in \hat{\mathcal{B}} \setminus \mathcal{S}} \text{UCB}_i$ 
12:  end for
13:  Train the selected  $B_{\text{tr}}$  beam pairs to get  $\gamma_{i,n}$  for  $\forall i \in \mathcal{S}$ 
14:  // Update the learning parameters
15:   $T_i \leftarrow T_i + 1$ , for  $\forall i \in \mathcal{S}$ 
16:   $X_{\text{tot}}[\arg \max_{k \in \mathcal{S}} \gamma_{k,n}] \leftarrow X_{\text{tot}}[\arg \max_{k \in \mathcal{S}} \gamma_{k,n}] + 1$ 
17: end for
```

---

in  $\mathcal{S}$ . This means the optimal beam pair subset can be obtained by a greedy approach, where one beam pair is selected at a time. Observing this property, we propose to use a greedy UCB algorithm as shown in Algorithm 4.1, that selects the beam pairs by their UCB indices. The UCB index of an arm is a high confidence bound of the expected reward, which consists of the expected reward seen so far and the uncertainty (the confidence margin) [15].

An important part of Algorithm 4.1 is the reward signal. Since the expected reward is the the probability of being optimal, an ideal choice for the

reward signal is

$$x_{i,t} = \begin{cases} 1 & \text{if } i \text{ was best in } \hat{\mathcal{B}} \\ 0 & \text{otherwise} \end{cases}, \quad (4.12)$$

which takes the value 1 if the pair  $i$  is best and 0 otherwise. In an actual setting, it is not known if a pair is the best in  $\hat{\mathcal{B}}$  since only the beam measurements for the beam pairs in the subset  $\mathcal{S} \subset \hat{\mathcal{B}}$  are available. The best guess would be the strongest pair among the beam pairs trained. Considering these limitations, we propose to use an alternative and practical reward signal,

$$x_{i,t} = \begin{cases} 1 & \text{if } i \text{ was best in } \mathcal{S} \\ 0 & \text{all other pairs in } \mathcal{S} \end{cases}, \quad (4.13)$$

which takes the value 1 for the pair with the strongest beam measurements in  $\mathcal{S}$  and 0 for all other pairs in  $\mathcal{S}$ . Denoting  $X_{\text{tot}}[i] = \sum_{t=1}^n x_{i,t}$ , the expected reward of beam pair  $i$  at time  $n$  is estimated by  $\hat{P}_{\text{opt}}(i) = X_{\text{tot}}[i]/T_i$ , where  $T_i$  is the number of times that the pair  $i$  was selected up to time  $n$ .

An intuitive understanding of this alternative reward definition can be drawn from an analogy to a sport tournament. In each round, the winners from each subgroup from the previous round play against each other to decide who will proceed to the next round, which eventually will reach the championship. We, thus, expect that over time only strong beam pairs will receive reward of 1. We believe that under certain assumption on the underlying reward statistics of the beam pairs, it is possible to provide some guarantee that  $\hat{P}_{\text{opt}}(i)$  will converge to the true  $P_{\text{opt}}(i)$  as  $T_i \rightarrow \infty$ . This is outside the focus of this chapter and is left for future work.



As will be seen in Section 4.7.1, Algorithm 4.1 does not perform well. The main reason for this is because it only tries to minimize the cumulative regret and is oblivious to the multiple-play setting in the beam alignment problem. Since multiple beam pairs are trained, the subset  $\mathcal{S}$  can be divided into two parts. One part is for exploitation that uses the knowledge obtained so far to select the beam pairs and the other part is for exploration that aims at improving the accuracy of the learning parameters. By balancing these two parts, it is possible to reduce the risk (high power loss events) at any given round. In other words, in the multiple-play setting, the risk of large losses can be traded off with the speed of learning (time to get accurate statistics of the arms). Another point for improvement for Algorithm 4.1 is that it throws away the magnitude information because the reward signal is binary. Recall that the binary reward signal is needed because we make use of the modularity of  $P_{\text{opt}}(\mathcal{S})$  that is the basis for the greedy selection using the UCB indices. To remedy these weaknesses, we propose a risk-aware version of Algorithm 4.1.

### 4.5.3 Risk-Aware Greedy UCB Algorithm

We first start with the definition of risk. A possible choice for the risk is the power loss, which measures the misalignment severity. Since only the beam pairs in  $\mathcal{S}$  are trained, the channel strength of the optimal beam pair is not necessarily known (especially, during the early stage of the learning) and the power loss cannot be computed directly. Another important point is that this risk needs to be estimated. Therefore, it is crucial to quantify the

uncertainty of the risk estimate for it to be useful for the beam pair selection. For these two reasons, we propose to use a binary risk signal defined in terms of the ratio of the channel strength of the beam pair and the best beam pair in  $\mathcal{S}$ , i.e., the risk signal of the beam pair  $i$  at time  $t$  is given by

$$z_{i,t} = \begin{cases} 1 & \text{if } \frac{\max_{k \in \mathcal{S}} \gamma_{tk}}{\gamma_{i,t}} > \Gamma_{\text{risk}} \\ 0 & \text{otherwise} \end{cases}, \quad (4.14)$$

where  $\Gamma_{\text{risk}}$  is a risk threshold. The choice of  $\Gamma_{\text{risk}}$  will be discussed in Section 4.7.2.

A way to capture the uncertainty is to put a prior distribution on the risk based on the observations seen so far. By the definition (4.14),  $z_{i,t}$  is Bernoulli distributed with some unknown parameter  $\zeta$ . It is well-known that the Beta distribution is the conjugate prior to the Bernoulli distribution [57]. This means that the belief on the risk of the beam pair  $i$  upon seeing the measurements up to time  $n$  can be updated conveniently by updating the parameters of the Beta distribution as

$$\tilde{Z}_n \sim \text{Beta}(1 + Z_{\text{tot}}[i], 1 + T_i - Z_{\text{tot}}[i]), \quad (4.15)$$

where we denote  $Z_{\text{tot}}[i] = \sum_{t=1}^n z_{i,t}$ . Here, we assume that at time 0 without any observation,  $\tilde{Z}_0 \sim \text{Beta}(1, 1)$ , which is the uniform distribution over  $[0, 1]$ . This is a reasonable assumption since no information on the beam pair  $i$  is available at time 0.

We next explain how the risk estimate along with the prior are used in the rejection mechanism to reduce the probability of large power loss events

during the learning. The new algorithm is shown in Algorithm 4.2, which we call the risk-aware greedy UCB algorithm. The new addition to Algorithm 4.1 is the risk-aware feature that rejects a beam pair selected by the greedy UCB with a probability reflecting its risk. The rejection probability is determined using the risk drawn from the prior distribution given in (4.15) and the confidence margin in two steps. First, a random variable  $\tilde{Z}_n$  is drawn from this prior (line 13). Then, it is multiplied by the confidence margin for those beam pairs with  $X_{\text{tot}}[\ell] > 0$ . The obtained  $Z_n$  is the rejection probability. The second step is needed because any beam pair is subject to blockage and their risks are not zero. This means that if  $\tilde{Z}_n$  is used directly as the rejection probability, even good beam pairs will be rejected with non-zero probability even when  $n \rightarrow \infty$ . The second step ensures the algorithm accepts the UCB selection for “good” beam pairs with increasing probability over time.

The proposed rejection mechanism is a random method that rejects the beam pair with a probability  $Z_n$ . First, the algorithm draws a Bernoulli random variable  $\text{Rej}$  with parameter  $Z_n$ . If  $\text{Rej} = 0$ , the algorithm accepts the beam pair, otherwise it rejects the pair. In that case, the replacement beam pair is selected using  $\hat{P}_{\text{opt}}(i)$  when there are pairs with  $X_{\text{tot}}[i] > 0$ , and using the average channel strength  $\bar{\gamma}_i$  when all remaining pairs have  $X_{\text{tot}}[i] = 0$ . Note that unlike Algorithm 4.1, which does not use the amplitudes of the beam measurements  $\gamma_{i,n}$ , here they are used to update the risk parameters and also used for the replacement selection. This new algorithm makes a fuller use of the measurement information as compared to Algorithm 4.1.

---

**Algorithm 4.2** Risk-Aware Greedy UCB
 

---

```

1: // initialize arms' parameters using a small offline database
2:  $X_{\text{tot}}[i] \leftarrow 0$ , for  $\forall i \in \hat{\mathcal{B}}$ 
3:  $X_{\text{tot}}[\arg \max_{i \in \hat{\mathcal{B}}} \bar{\gamma}_i^{\text{init}}] \leftarrow 1$ 
4:  $Z_{\text{tot}}[i] \leftarrow 0$ , for  $\forall i \in \hat{\mathcal{B}}$ 
5:  $T_i \leftarrow 1$ , for  $\forall i \in \hat{\mathcal{B}}$ 
6: for  $n = 1, 2, \dots$  do
7:   // Compute UCB values
8:    $\text{UCB}_i \leftarrow \frac{X_{\text{tot}}[i]}{T_i} + \sqrt{\frac{2 \log(n)}{T_i}}$ , for  $\forall i \in \hat{\mathcal{B}}$ 
9:   // Greedy selection using UCB values
10:   $\mathcal{S} \leftarrow \emptyset$ 
11:  for  $k = 1, 2, \dots, B_{\text{tr}}$  do
12:     $\ell \leftarrow \arg \max_{i \in \hat{\mathcal{B}} \setminus \mathcal{S}} \text{UCB}_i$ 
13:     $\tilde{Z}_n \sim \text{Beta}(1 + Z_{\text{tot}}[\ell], 1 + T_\ell - Z_{\text{tot}}[\ell])$ 
14:     $Z_n \leftarrow \tilde{Z}_n \times \sqrt{\frac{2 \log(n)}{T_\ell}}$  if  $X_{\text{tot}}[\ell] > 0$ , else  $Z_n \leftarrow \tilde{Z}_n$ 
15:     $\text{Rej} \sim \text{Ber}(Z_n)$ 
16:    if  $\text{Rej} = 0$  then
17:       $\mathcal{S} \leftarrow \mathcal{S} \cup \{\ell\}$ 
18:    else
19:      if  $\exists i \in \hat{\mathcal{B}} \setminus \mathcal{S}$  with  $X_{\text{tot}}[i] > 0$  then
20:         $\mathcal{S} \leftarrow \mathcal{S} \cup \arg \max_{i \in \hat{\mathcal{B}} \setminus \mathcal{S}} \hat{P}_{\text{opt}}(i)$ 
21:      else
22:         $\mathcal{S} \leftarrow \mathcal{S} \cup \arg \max_{i \in \hat{\mathcal{B}} \setminus \mathcal{S}} \bar{\gamma}_i$ 
23:      end if
24:    end if
25:  end for
26:  Train the selected  $B_{\text{tr}}$  beam pairs to get  $\gamma_{i,n}$  for  $\forall i \in \mathcal{S}$ 
27:  // Update the learning parameters
28:   $T_i \leftarrow T_i + 1$ , for  $\forall i \in \mathcal{S}$ 
29:   $X_{\text{tot}}[\arg \max_{k \in \mathcal{S}} \gamma_{k,n}] \leftarrow X_{\text{tot}}[\arg \max_{k \in \mathcal{S}} \gamma_{k,n}] + 1$ 
30:   $Z_{\text{tot}}[i] \leftarrow Z_{\text{tot}}[i] + 1$  if  $(\max_{k \in \mathcal{S}} \gamma_{k,n}) / \gamma_{i,n} > \Gamma_{\text{risk}}$ ,  $\forall i \in \mathcal{S}$ 
31: end for

```

---

#### 4.5.4 Regret Analysis

In this subsection, we derive regret bounds of the two algorithms that will provide insights on the effect of the rejection mechanism we introduced in Algorithm 4.2. We make a few simplifications to the problems to allow tractable analysis which we will describe in detail when presenting the results. Proofs are provided in the appendices.

Before stating the results, we first describe the metric used for the evaluation. For this type of online learning problem, a widely used metric is the cumulative regret. It is defined as *the cumulative performance loss as compared to the performance of an oracle that always plays the best arm* [19]. Translating this to the beam alignment problem, *the regret incurred in a time step is non-zero when the algorithm does not select the best subset of beam pairs*  $\mathcal{S}^*$ . Assuming  $|\mathcal{S}^*| = B_{\text{tr}}$ , the optimal selection in (4.8) tells us that  $\mathcal{S}^*$  contains the top- $B_{\text{tr}}$  beam pairs with the highest probability of being optimal  $P_{\text{opt}}(\cdot)$ . Now, we call the beam pairs with the  $B_{\text{tr}}$ -highest  $P_{\text{opt}}(\cdot)$  as optimal and the rest of beam pairs as suboptimal. Then, the cumulative regret increases whenever one or more suboptimal beam pairs are selected in the selection set  $\mathcal{S}$ .

In the following, we present what is called a problem-dependent bound on cumulative regret (we drop ‘cumulative’ from now on for convenience), which quantifies the regret in terms of the optimality gap. The optimality gap is defined as the difference in the probability of being optimal of an optimal

pair  $i^*$  and a suboptimal pair  $\ell$ , i.e.,

$$\Delta_{\ell,i^*} = P_{\text{opt}}(i^*) - P_{\text{opt}}(\ell). \quad (4.16)$$

By definition,  $0 < \Delta_{\ell,i^*} < 1$  if  $P_{\text{opt}}(i^*) > 0$ . Note that  $\Delta_{\ell,i^*}$  measures the difficulty in discriminating the suboptimal pair  $\ell$  from the optimal pair  $i^*$  for the particular problem at hand; thus, the name problem-dependent when the regret bound is expressed using optimality gaps.

For Algorithm 4.1, we assume the reward signal during the learning is the ideal reward and not the alternative one, i.e., we assume the reward signal is given by (4.12) instead of (4.13). We make this assumption because it is intractable to deal directly with the dynamics of the alternative reward estimate of (4.13). The main step in deriving the regret bound is the application of the Chernoff-Hoeffding inequality to bound the probability that the sample average of the reward is within the UCB value. To apply the Chernoff-Hoeffding inequality, it is required that the sample rewards are IID, which cannot be guaranteed when using the alternative reward definition because its distribution depends on the history of the selection done so far. This, however, is a reasonable assumption, since we expect that (4.13) will approach (4.12) for large  $n$ , which is the domain where the regret bound is meaningful. The derived bound is shown in Theorem 4.1. We note that the regret bound is  $\mathcal{O}(\log(n))$ , which is known to be optimal up to the constant coefficient in front of  $\log(n)$  [19]. This confirms that the algorithm is a reasonable solution.

**Theorem 4.1.** *Assuming that the ideal reward signal (4.12) is accessible, the expected regret at time  $n$  of the greedy UCB algorithm is upper bounded by*

$$R_1[n] \leq 8 \log(n) \sum_{\ell \in \mathcal{B} \setminus \mathcal{S}^*} \sum_{i^* \in \mathcal{S}^*} \frac{1}{\Delta_{\ell, i^*}} + \left(1 + \frac{\pi^2}{3}\right) \sum_{\ell \in \mathcal{B} \setminus \mathcal{S}^*} \sum_{i^* \in \mathcal{S}^*} \Delta_{\ell, i^*}. \quad (4.17)$$

For Algorithm 4.2, we make two additional assumptions besides the accessibility of the ideal reward signal. The first assumption is that the rejection probability of any beam pair  $\ell$  is constant, denoted by  $1 - \zeta_\ell$ . This is used because the rejection probability of the algorithm is dynamic (depending on the observations so far) and is not tractable. With a large enough  $n$ , we expect the risk estimate to stabilize, and thus this is not an unreasonable assumption. The second assumption is that when rejected the replacement selection has an optimality gap  $\tilde{\Delta}_{\ell, i^*}$ .

**Theorem 4.2.** *Assuming that the ideal reward signal (4.12) is available, the rejection probability of beam pair  $\ell$  is  $1 - \zeta_\ell$ , and that when rejected the optimality gap of the replacement selection is  $\tilde{\Delta}_{\ell, i^*}$ , then the expected regret at time  $n$  of the risk-aware greedy UCB algorithm is bounded by*

$$\begin{aligned} R_2[n] \leq & \frac{8 \log(n)}{\delta^2} \sum_{\ell \in \mathcal{B} \setminus \mathcal{S}^*} \sum_{i^* \in \mathcal{S}^*} \frac{1}{\Delta_{\ell, i^*}} + \frac{8 \log(n)}{\delta^2} \sum_{\ell \in \mathcal{B} \setminus \mathcal{S}^*} \sum_{i^* \in \mathcal{S}^*} \frac{(1 - \zeta_\ell) \tilde{\Delta}_{\ell, i^*}}{\zeta_\ell \Delta_{\ell, i^*}^2} \\ & + \left(1 + \frac{\pi^2}{2}\right) \sum_{\ell \in \mathcal{B} \setminus \mathcal{S}^*} \sum_{i^* \in \mathcal{S}^*} (\zeta_\ell \Delta_{\ell, i^*} + (1 - \zeta_\ell) \tilde{\Delta}_{\ell, i^*}), \end{aligned} \quad (4.18)$$

where  $\delta = (\sqrt{5} - 1)/2$ .

Theorem 4.2 shows a regret bound for Algorithm 4.2. The algorithm still has  $\mathcal{O}(\log(n))$  regret but with a larger constant. This shows that introducing risk-awareness increases the learning time in the sense that  $R_2[n] > R_1[n]$ .

This is because by rejecting a high-risk beam pair, the algorithm loses the chance to get information on that beam pair. The idea of Algorithm 4.2 is to distribute the learning of these high-risk beam pairs (which has high cost) more evenly among the users by rejecting them with some probability. In other words, Algorithm 4.2 tradeoffs the learning speed to balance the risk of severe misalignment endured by each user at different stages of the learning.

## 4.6 Online Beam Pair Refinement

In this section, we describe our beam pair refinement solution, which is the second layer of the two-layer online learning algorithm. We start with the problem formulation and then describe our modified HOO solution.

### 4.6.1 Problem Statement

We formulate our beam pair refinement as a stochastic CAB problem. The beams are generated by progressive phase-shift and are defined by their main beam directions. The goal is to find the pointing direction of a beam pair to maximize the average channel strength of that beam pair in an online setting. Specifically, denoting  $\phi_i^t, \theta_i^t, \phi_i^r$ , and  $\theta_i^r$  the transmit and receive main beam directions in the azimuth and elevation of the beam pair  $i$  defined in the codebook, and  $\Phi_i^t, \Theta_i^t, \Phi_i^r$ , and  $\Theta_i^r$  the corresponding 3 dB beamwidths, the



problem of refining the beam pair  $i$  can be written as

$$\begin{aligned}
& \underset{\phi^t, \theta^t, \phi^r, \theta^r}{\text{maximize}} && \mathbb{E}[\gamma_i(\phi^t, \theta^t, \phi^r, \theta^r)] \\
& \text{subject to} && \phi^t \in [\phi_i^t - \Phi_i^t/2, \phi_i^t + \Phi_i^t/2], \\
& && \phi^r \in [\phi_i^r - \Phi_i^r/2, \phi_i^r + \Phi_i^r/2], \\
& && \theta^t \in [\theta_i^t - \Theta_i^t/2, \theta_i^t + \Theta_i^t/2], \\
& && \theta^r \in [\theta_i^r - \Theta_i^r/2, \theta_i^r + \Theta_i^r/2].
\end{aligned} \tag{4.19}$$

Any pointing direction  $(\phi^t, \theta^t, \phi^r, \theta^r)$  satisfying the constraints is an arm in this problem. The space to search for the best arm is the hyperrectangle defined by the constraints, which is a continuous space. This means that the directions are fine-tuned within the 3 dB beamwidths of the original beam pair  $i$  defined by the pointing direction  $(\phi_i^t, \theta_i^t, \phi_i^r, \theta_i^r)$ . The coarse search to within the 3 dB beamwidth is supposed to be done by the beam pair selection algorithm.

#### 4.6.2 Modified HOO for Beam Pair Refinement

HOO is a CAB algorithm that runs on a search tree. We start by describing the search tree. Then, we explain the flow of HOO. Finally, we provide the details of the modifications made to the original algorithm to fit the beam refinement task. We describe the algorithm for refining a beam pair  $i$ . Since all the description is in the context of this beam pair  $i$ , we drop explicit references to beam pair  $i$  here to avoid notational clutter.

We now define the search tree  $\mathcal{T}$  which HOO runs on. Each node in the tree is the pair of transmit and receive pointing directions  $(\phi^t, \theta^t, \phi^r, \theta^r)$  satisfying the constraints in (4.19). The root of the tree is the original pointing direction of the beam pair  $i$   $(\phi_i^t, \theta_i^t, \phi_i^r, \theta_i^r)$  as defined in the codebook. Each node in the tree at depth  $\ell < \ell_{\max}$  has 16 children which correspond to all

possible combinations of transmit and receive beam directions perturbed by  $1/2^\ell$  of the beamwidths in the four variables. Denote  $(\phi_{\ell,k}^t, \theta_{\ell,k}^t, \phi_{\ell,k}^r, \theta_{\ell,k}^r)$  the parameters of the  $k$ -th node at depth  $\ell$  in  $\mathcal{T}$ , its set of 16 children nodes can be writting using a Cartesian product as

$$\left\{ \left\{ \begin{bmatrix} \phi_{\ell,k}^t + \Phi_i^t/2^\ell, & \theta_{\ell,k}^t \end{bmatrix}^T \right. \right. \\ \left. \begin{bmatrix} \phi_{\ell,k}^t - \Phi_i^t/2^\ell, & \theta_{\ell,k}^t \end{bmatrix}^T \right. \\ \left. \begin{bmatrix} \phi_{\ell,k}^t, & \theta_{\ell,k}^t + \Theta_i^t/2^\ell \end{bmatrix}^T \right. \\ \left. \begin{bmatrix} \phi_{\ell,k}^t, & \theta_{\ell,k}^t - \Theta_i^t/2^\ell \end{bmatrix}^T \right\} \times \left\{ \begin{bmatrix} \phi_{\ell,k}^r + \Phi_i^r/2^\ell, & \theta_{\ell,k}^r \end{bmatrix}^T \right. \\ \left. \begin{bmatrix} \phi_{\ell,k}^r - \Phi_i^r/2^\ell, & \theta_{\ell,k}^r \end{bmatrix}^T \right. \\ \left. \begin{bmatrix} \phi_{\ell,k}^r, & \theta_{\ell,k}^r + \Theta_i^r/2^\ell \end{bmatrix}^T \right. \\ \left. \begin{bmatrix} \phi_{\ell,k}^r, & \theta_{\ell,k}^r - \Theta_i^r/2^\ell \end{bmatrix}^T \right\}. \quad (4.20)$$

Using this node expansion rule, a node at depth  $\ell + 1$  deviates from its parent node in the pointing direction by beamwidth/ $2^\ell$  and depends only on  $\ell$ . Note that each depth in the tree can be thought of as a grid partitioning the search space defined by the constraints in (4.19). The grid becomes finer deeper in the tree (i.e., as  $\ell$  increases).

We now describe how the modified HOO works. A pseudo-code is shown in Algorithm 4.3. It runs on a finite tree with a maximum depth of  $\ell_{\max}$ . The nodes in the tree are activated on the fly, and only the root node and its children are active at time  $n = 0$ . Thus, the initial tree is  $\mathcal{T} = \{(1, 1)\} \cup \mathcal{C}_{1,1}$ , where  $\mathcal{C}_{\ell,k}$  denotes the set of the indices of the children of the node  $(\ell, k)$ . In each iteration, there are three main parts. First, a node is selected by traversing the active tree starting from the root following the path through nodes that have the largest B-values (line 6-13), which is the best optimistic estimate of the average rewards of the nodes. The second part is the beam measurement for the selected node (line 14). Lastly, after obtaining the measurement, the learning parameters are updated. If the condition is met, a node in the tree is expanded, i.e., activating its 16 children nodes. Note

---

**Algorithm 4.3** Modified HOO for Beam Pair Refinement
 

---

```

1: // Initialization
2:  $\mathcal{T} \leftarrow \{(0, 1)\} \cup \mathcal{C}_{1,1}$ 
3:  $(B_{2,j}, T_{2,j}, \hat{\mu}_{2,j}, \mathbf{Sq}_{2,j}) \leftarrow (\infty, 0, 0, 0)$  for  $\forall j \in \mathcal{C}_{1,1}$ 
4: for  $n = 1, 2, \dots$  do
5:   // Select a node in the tree to sample
6:    $(\ell, k) \leftarrow (1, 1)$  // Start from the root node
7:    $\mathcal{P} \leftarrow \{(\ell, k)\}$ 
8:   for  $\ell = 1, \dots, \min\{\text{depth}(\mathcal{T}), \ell_{\max} - 1\}$  do
9:      $k^* \leftarrow \arg \max_{j \in \mathcal{C}_{\ell,k}} B_{\ell+1,j}$ 
10:     $(\ell, k) \leftarrow (\ell + 1, k^*)$ 
11:     $\mathcal{P} \leftarrow \mathcal{P} \cup \{(\ell, k)\}$ 
12:  end for
13:   $(\ell_s, k_s) \leftarrow (\ell, k)$ 
14:  Obtain the beam measurement for node  $(\ell_s, k_s)$  denoted by  $\gamma$ 
15:  // Update the learning parameters
16:  for  $(\ell, k) \in \mathcal{P}$  do // update sample averages
17:     $T_{\ell,k} \leftarrow T_{\ell,k} + 1$ 
18:     $\hat{\mu}_{\ell,k} \leftarrow (1 - \frac{1}{T_{\ell,k}})\hat{\mu}_{\ell,k} + \gamma/T_{\ell,k}$ 
19:     $\mathbf{Sq}_{\ell,k} \leftarrow \mathbf{Sq}_{\ell,k} + \gamma^2$ 
20:     $\sigma_{\ell,k}^2 \leftarrow (\mathbf{Sq}_{\ell,k} - \hat{\mu}_{\ell,k}^2 T_{\ell,k})/T_{\ell,k}$ 
21:  end for
22:  for all  $(\ell, k) \in \mathcal{T}$  do // update U-values
23:     $U_{\ell,k} \leftarrow \left( \hat{\mu}_{\ell,k} + \sqrt{16\hat{\sigma}_{\ell,k}^2 \frac{\log(n)}{T_{\ell,k}}} \right) \nu(\ell)$ 
24:    // forced exploration
25:     $U_{\ell,k} \leftarrow \infty$  if  $T_{\ell,k} < \lceil \alpha_{\text{norm}} \log(n) \rceil$  or  $T_{\ell,k} < K_{\min}$ 
26:  end for
27:  // expand a node if conditions are met
28:  if  $\ell_s < \ell_{\max} \wedge T_{\ell_s, k_s} > K_{\text{exd}} \wedge (\ell_s, k_s)$  is a leaf then
29:     $\mathcal{T} \leftarrow \mathcal{T} \cup \mathcal{C}_{\ell_s, k_s}$ 
30:     $(B_{\ell_s+1,j}, T_{\ell_s+1,j}, \hat{\mu}_{\ell_s+1,j}, \mathbf{Sq}_{\ell_s+1,j}) \leftarrow (\infty, 0, 0, 0)$  for  $\forall j \in \mathcal{C}_{\ell_s,j}$ 
31:  end if
32:  for  $\ell = \ell_s, \ell_s - 1, \dots, 2$  do // update B-values
33:     $B_{\ell,k} \leftarrow \min\{U_{\ell,k}, \max_{j \in \mathcal{C}_{\ell,k}} B_{\ell+1,j}\}, \forall$  nodes at depth  $\ell$  in  $\mathcal{T}$ 
34:  end for
35: end for

```

---

that to lower the risk of expanding a suboptimal node, it is enforced that a node can be expanded only after it is sampled  $K_{\text{exd}}$  times. The last part of the parameter update is the B-values. They are computed by back calculation from the sampled node back to the root (line 33). The B-value of node  $(\ell, k)$  is the minimum between its own U-value and the maximum B-value of its children nodes, i.e.,

$$B_{\ell,k} \leftarrow \min \left\{ U_{\ell,k}, \max_{j \in \mathcal{C}_{\ell,k}} B_{\ell+1,j} \right\}. \quad (4.21)$$

The U-value is similar to the UCB value but it also accounts for the smoothness property (line 23). The U-value provides an optimistic estimate of its average reward using the parameter of the node, and the maximum B-value among its children nodes provides another optimistic estimate of its reward. By taking the minimum between the two, the obtained B-value provides a refined optimistic estimate of the average reward of the node.

We introduce three main modifications to the original HOO tailored to the beam refinement setting. The first one is the use of a finite tree. The original HOO assumes an infinite tree to represent the arm space. Since small adjustments (e.g., 1/8 of the beamwidth) have a small impact on the gain, a finite tree of maximum depth  $\ell_{\text{max}}$  is used instead to save computation and storage for the learning parameters. The second one is the smoothness bound. The original HOO assumes an additive offset. Due to the multiplicative nature of the antenna gains, a multiplicative factor  $\nu(\ell)$  as shown in line 23 is more suitable. The factor is computed using Lemma 4.1, which will be detailed in the last part of this subsection.

The third modification is the confidence margin. Because the original margin  $\sqrt{2\log(n)/T_{h,i}}$  is too loose in our setting, we propose to use that of the norm-UCB (line 23) [15]. The margin  $\sqrt{2\log(n)/T_{h,i}}$  is derived from the Chernoff-Hoeffding inequality, which is applicable to any distribution with the support in  $[0, 1]$ . While normalizing the channel strengths by a large enough number will approximately guarantee that the support is within  $[0, 1]$ , the average typically takes a value much less than 1 and the margin  $\sqrt{2\log(n)/T_{h,i}}$  is too loose for reasonable learning horizons. The reason that the average is much smaller than 1 is that due to the small scale fading nature of the wireless channel. Fading is the result of the multipath effect and can cause the maximum instantaneous channel strength to be much larger than the average [46]. A good property of this new margin is that the sample variance is also used. Note that to enable regret analysis, the norm-UCB algorithm requires each arm be sampled at least  $\lceil \alpha_{\text{norm}} \log(n) \rceil$  at time  $n$  with  $\alpha_{\text{norm}} = 8$  [15]. This is enforced by setting the U-values of the nodes that need to be explored to infinity (see line 25). Note that we also introduce the condition  $T_{\ell,k} < K_{\min}$ , which is used to ensure that there are at least  $K_{\min}$  samples of the node for computing the sample variance. This is needed when using a small  $\alpha_{\text{norm}}$ .

We next state a lemma defining the smoothness property of the objective function in (4.19). The lemma is used for computing the smoothness coefficient  $\nu(\ell)$ .

**Lemma 4.1.** *Assume a single-path azimuth PAS with the optimal beam direction  $\phi^*$  with isotropic transmit antenna,  $G(\cdot; \phi_0)$  the normalized gain of*

the beam pattern pointing at  $\phi_0$  assumed to be decreasing and concave in  $[\phi_0, \phi_0 + \Phi/2]$  with  $\Phi$  denoting the 3 dB beamwidth (e.g., true for a uniform linear array), for a receive pointing direction  $\phi_0$  such that  $|\phi^\star - \phi_0| \leq \Delta\phi \leq \Phi/2$ ,

$$\bar{\gamma}(\phi_0)/G(\phi_0 + \Delta\phi; \phi_0) \geq \bar{\gamma}(\phi^\star). \quad (4.22)$$

Moreover, for a general PAS with the support within  $[\phi_0 - \Psi, \phi_0 + \Psi]$  with  $\Psi \leq \Phi/2$ ,

$$\bar{\gamma}(\phi_0)/G(\phi_0 + \Delta\phi; \phi_0) \geq \bar{\gamma}(\phi^\star) - \text{Err}, \quad (4.23)$$

where  $\text{Err} \geq 0$  is a residual term that depends on the shape of the PAS and  $\text{Err} \rightarrow 0$  as  $\Delta\phi \rightarrow 0$ .

We now explain how to determine  $\nu(\ell)$  using Lemma 4.1. While we state Lemma 4.1 assuming isotropic transmit antenna to avoid tedious notations, the same argument applies when we also include the transmit beam pattern. In particular,

$$\bar{\gamma}(\phi_0^r, \phi_0^t)/G_r(\phi_0^r + \Delta\phi^r; \phi_0^r)G_t(\phi_0^t + \Delta\phi^t; \phi_0^t) \geq \bar{\gamma}(\phi^{r\star}, \phi^{t\star}) - \text{Err}. \quad (4.24)$$

When steering the elevation only, we get the same relation as (4.24). If we assume square UPAs, the beam pattern in the azimuth and elevation will be the same. Since we only change the azimuth or elevation but not both per (4.20), the smoothness coefficient is given by

$$\nu(\ell) = a/(G_r(\phi_\ell^r + \Phi/2^\ell; \phi_\ell^r)G_t(\phi_\ell^t + \Phi/2^\ell; \phi_\ell^t)) \quad (4.25)$$

$$\simeq a/g^2(\text{beamwidth}/2^\ell), \quad (4.26)$$

where  $a > 1$  is a correction coefficient to account for  $\text{Err}$  if deemed necessary. For convenience, we approximate the gain by  $g(\cdot)$  the beam pattern at broadside as a function of the deviation from the broadside direction. Note that we will need  $a$  only for large  $\Delta\phi$ . Deeper in the tree, the change in the angle is small and thus  $\text{Err}$  will become negligible. Also, for the sake of clear argument, we restrict  $\Psi \leq \Phi/2$ , but with a more elaborate choice of the coefficient of  $\text{Err}$  in the proof, we can allow  $\Psi$  to be larger. This, however, is not a big concern in our setting because  $\Delta\phi$  will be  $\Phi/4$  or less and  $\text{Err}$  is restricted to a small value already.

## 4.7 Numerical Results

We start with the general setting of our numerical evaluations. As described in Section 3.3, our codebook for  $16 \times 16$  UPA has 271 beams and thus there are  $271^2$  beam pairs. Using the heuristic screening to get  $\hat{\mathcal{B}}$  as explained in Section 4.5.1 with the initial database size of  $N = 5$  and  $C = 200$ , the size of the set of beam pairs to be learned  $|\hat{\mathcal{B}}|$  is typically between 400 and 600 depending on the simulation run. Following the ray-tracing described in Section 3.3, we generated 10,000 channel samples using ray-tracing. To eliminate the effect of the ordering of the channel samples on the learning performance, the evaluation metrics are averaged over 100 simulation runs, where in each run we randomly permute these 10,000 channel samples. We apply a moving average with a window size of 50 time steps to better show the trends.

As an evaluation metric, we use the 3 dB power loss probability and the gain defined as the inverse of the power loss in (4.1). The 3 dB power loss probability (i.e.,  $c = 2$  in (4.2)) measures how often the selected beam pair has a loss larger than 3 dB as compared to the best beam pair selected by exhaustive search, and thus capturing the beam alignment accuracy. This metric, however, is not suitable for evaluating the beam pair refinement because it cannot capture the improvement over the exhaustive search in the original codebook. Allowing the refinement, a beam pair better than the best in the original codebook can be selected resulting in power loss taking a value less than one, or equivalently, a positive gain in dB.

The rest of the section is divided into three parts. In Section 4.7.1, we evaluate the beam pair selection alone without the refinement option. In Section 4.7.2, we assess the performance of the beam pair refinement assuming an offline learning for the beam pair selection. Section 4.7.3 provides evaluations of the integrated solution including both the online beam pair selection and refinement components.

#### 4.7.1 Online Beam Pair Selection

In this subsection, we evaluate the performance of the proposed risk-aware greedy UCB algorithm without the beam refinement option. There are two parameters to be decided when running Algorithm 4.2: the training budget  $B_{\text{tr}}$  and the risk threshold  $\Gamma_{\text{risk}}$  in (4.14). We note that our solution does not require that  $B_{\text{tr}}$  be fixed, but for simplicity, we assume that the same



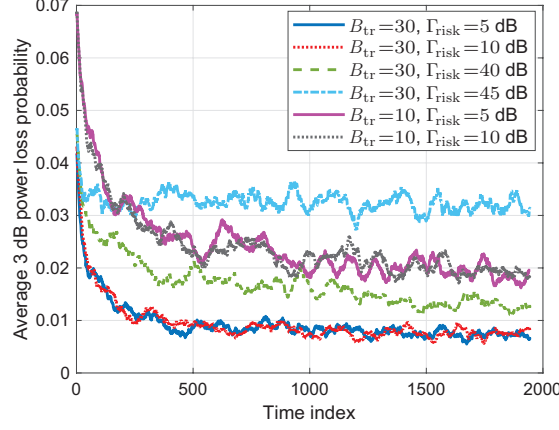


Figure 4.2: Average 3 dB power loss probability using the proposed risk-aware greedy UCB algorithm with different training budgets  $B_{\text{tr}}$  and risk thresholds  $\Gamma_{\text{risk}}$ . For both  $B_{\text{tr}} = 10$  and 30, the plots show similar learning behavior. A smaller training budget  $B_{\text{tr}} = 10$  provides less accuracy beam alignment. The plots using different  $\Gamma_{\text{risk}}$  show that the performance is not sensitive to  $\Gamma_{\text{risk}}$  as long as it is not too large.

$B_{\text{tr}}$  is used during the entire learning horizon. Fig. 4.2 shows the average 3 dB power loss probability versus time for  $B_{\text{tr}} = 10$  and 30 with different  $\Gamma_{\text{risk}}$ . We can confirm from the figure that using a larger training budget  $B_{\text{tr}}$  leads to lower 3 dB power loss probability, i.e., more accurate beam alignment. The learning seems to have two phases: the fast improvement phase in the early time steps and the slower improvement phase after that. For  $B_{\text{tr}} = 30$  and  $\Gamma_{\text{risk}} = 5$  dB, this phase change happens at around time index 500. The slower learning phase starts when the algorithm has identified high-risk beam pairs (with some certainty) and learns those beam pairs at a slow pace due to the rejection mechanism. Regarding the risk threshold, the results show that the algorithm is not sensitive to the choice of  $\Gamma_{\text{risk}}$ . As long as  $\Gamma_{\text{risk}}$  is not too large

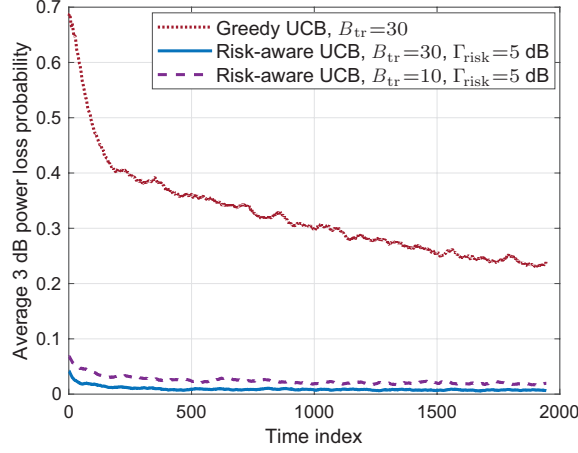


Figure 4.3: Performance comparison of greedy UCB with and without risk-awareness. The performance is an order of magnitude worse without risk-awareness. This is because the risk-aware greedy UCB uses the risk estimates to control the number of high-risk beam pairs selected in the subset  $\mathcal{S}$  reducing the probability of severe misalignment.

(e.g., less than 40 dB), it performs well. The main reason for this behavior is due to the effect of the replacement selection method (line 19-23 in Algorithm 4.2) that selects beam pairs to replace those rejected; even if a good beam pair gets rejected due to risk overestimation (when using a small  $\Gamma_{risk}$ ), it will likely be picked up by the replacement selection.

Fig. 4.3 shows a performance comparison of the greedy UCB algorithm with and without the risk-awareness. The performance without risk-awareness is an order of magnitude worse than that with risk-awareness. This might seem a bit counterintuitive because the regret bound of the risk-aware algorithm is higher. One way to understand this behavior is this. The goal of the UCB selection is to reach a state where we can ensure that a suboptimal arm is

not selected with high probability (call this the optimal state). To reach the optimal state, each arm has to be sampled enough times ( $T_{\ell,i^*}^{(0)}$  in the derivation in Appendix 4.A). Algorithm 4.1 samples the arms to reach this state fast, but it will expose early stage users to select more high-risk arms. Algorithm 4.2 balances the number of high-risk arms at any round by the rejection mechanism which results in a slower speed to reach the optimal state, i.e., a slower learning speed. By not exposing a user to too many high-risk arms, Algorithm 4.2 can ensure that the regret each user has to endure is not too large. In other words, although the cumulative regret is smaller (at a large enough time), users in early stages of Algorithm 4.1 have to sacrifice. Algorithm 4.2 distributes the regret more evenly among the users at different learning stages. We note that because of the large number of arms (400 to 600 as noted earlier), the time to reach the optimal state is large and Algorithm 4.1 is not practical as an online solution as shown in Fig. 4.3.

Lastly, Fig. 4.4 shows a performance comparison to demonstrate the effectiveness of our choice of the reward signal in (4.13). Specifically, we compare the accuracy of the beam selection using the average sample rewards ( $\hat{P}_{\text{opt}}(i)$ ) versus the more intuitive choice of average channel strength  $\bar{\gamma}_i$ . We also compare it with the case where we assume that the ideal reward defined in (4.12) is available to the algorithm during the learning. To evaluate this, we let the risk-aware greedy UCB algorithm run for 2000 time steps. We, then, use  $\hat{P}_{\text{opt}}(i)$  and  $\bar{\gamma}_i$  estimated at time step 2000 to get two sets of beam selections and evaluate the two sets over 500 channel samples. We use  $B_{\text{tr}} = 30$  and

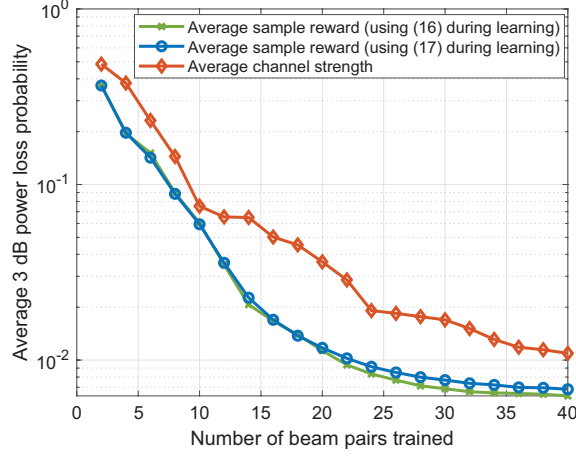


Figure 4.4: A comparison of the accuracy of the selection set produced by the average of the proposed reward signal ( $\hat{P}_{\text{opt}}(i)$ ) and the more intuitive choice of average channel strength. The performance when using  $\hat{P}_{\text{opt}}(i)$  is consistently better for all training budgets. The comparison when using the proposed practical reward signal (4.13) as opposed to the ideal reward signal (4.12) shows negligible performance loss.

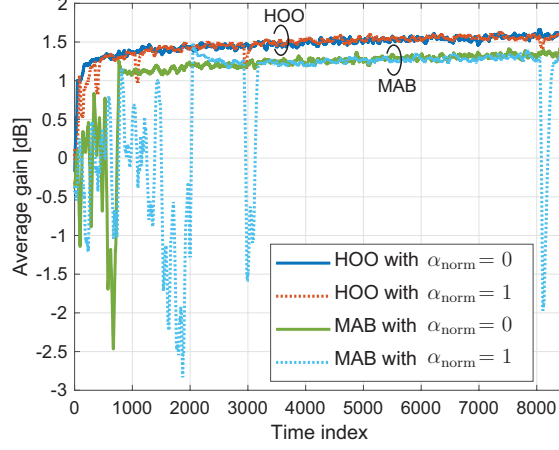
$\Gamma_{\text{risk}} = 5$  dB for the online learning. Fig. 4.4 shows the 3 dB power loss probability against the number of beam pairs trained. We can see that the beam pair selection using  $\hat{P}_{\text{opt}}(i)$  is more accurate than using the average channel strengths. Also, the plots show that the degradation due to the use of the proposed alternative and practical reward signal in (4.13) during the learning results in negligible loss. These results confirm the effectiveness of our choice of the reward signal in (4.13).

### 4.7.2 Online Beam Pair Refinement

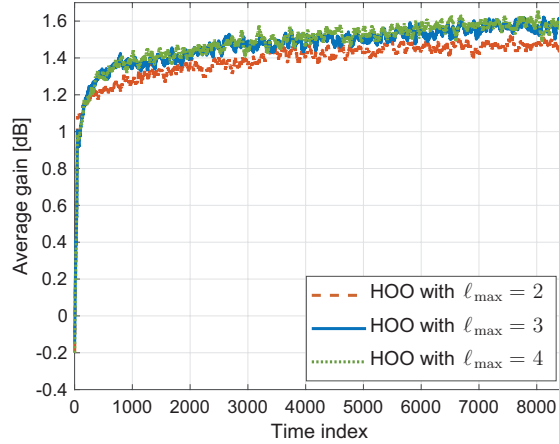
To evaluate the performance of the beam refinement on its own, we perform an offline beam pair selection using the MinMisProb method from Chapter 3 before running the beam refinement. In each simulation run, we use the first 300 channel samples to determine the selection set  $\mathcal{S}$ , and then we run the beam pair refinement on each of the beam pairs in  $\mathcal{S}$  where we set the training budget to  $B_{\text{tr}} = 30$ . For a baseline comparison, we implement an MAB solution using the norm-UCB algorithm from [15]. The MAB solution is run on the leaves of the search tree, and thus the number of arms is  $16^{\ell_{\text{max}}-1}$ . Besides  $B_{\text{tr}}$ , we also need to specify the maximum tree depth  $\ell_{\text{max}}$  and the forced exploration parameter  $\alpha_{\text{norm}}$ . We use  $K_{\text{min}} = 3$  and  $K_{\text{exd}} = 10$ .

We start by comparing the performance of MAB and our modified HOO solution in Fig. 4.5(a). We can see that HOO learns much faster by leveraging the tree structure. We can see the cost of exploration of MAB in the initial stage, where each arm has to be tried  $k_{\text{min}}$  times, where each arm has to be tried  $K_{\text{min}}$  times. Using the search tree, starting from the root, HOO will first explore the nodes at depth 2. At depth 3, it explores only the children nodes of promising nodes at depth 2, and this goes on until reaching  $\ell_{\text{max}}$ . This way, HOO does not have to sample all the leaves uniformly to explore the whole arm space leading to more efficient exploration than MAB.

We next show the effect of  $\alpha_{\text{norm}}$  and  $\ell_{\text{max}}$  on the performance. We noted earlier that  $\alpha_{\text{norm}} = 8$  is required to derive a regret bound in [15]. Forcing exploration this way with  $\alpha_{\text{norm}}$  turns out to result in bad performance for our



(a) HOO vs. MAB with  $\ell_{\max} = 3$ .



(b) HOO with different  $\ell_{\max}$  and  $\alpha_{\text{norm}} = 0$ .

Figure 4.5: A comparison of HOO and MAB with different  $\alpha_{\text{norm}}$  and  $\ell_{\max}$ . Fig. (a) compares the performance when  $\ell_{\max} = 3$ . MAB does not use the hierarchical structure of the search tree as HOO and suffers a larger exploration penalty. The penalty is even more severe as  $\alpha_{\text{norm}}$  increases. The results show that the forced exploration is not needed and  $\alpha_{\text{norm}} = 0$  should be used. Fig. (b) compares the performance of HOO when using different  $\ell_{\max}$ . There is negligible gain for setting  $\ell_{\max}$  beyond 3. We also see that HOO does not have extra degradation due to exploration when we increase  $\ell_{\max}$ .

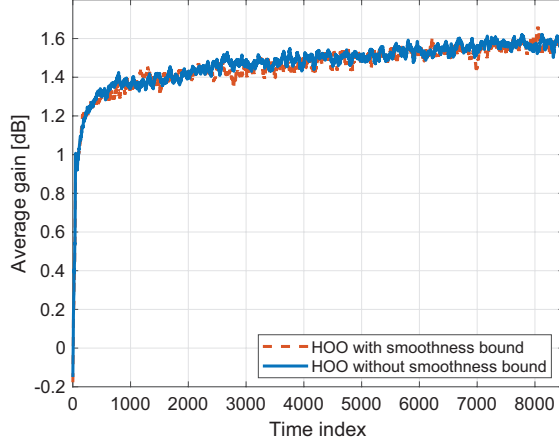


Figure 4.6: A comparison of HOO with and without the smoothness coefficient  $\nu(\ell)$  (for computing the U-values). The smoothness coefficient shows negligible effect. This is likely because the refinement problem searches locally within the 3 dB beamwidth. Since this is in the vicinity of the optimal point, the smoothness bound does not help in eliminating suboptimal nodes resulting in no performance gain.

applications as shown in Fig. 4.5(a). The dips in the gains are due to this forced exploration, and the intervals between dips decrease as  $\alpha_{\text{norm}}$  increases. Note that even with  $\alpha_{\text{norm}} = 0$ , both MAB and HOO still explore because of the confidence margin of the norm-UCB  $\sqrt{16\hat{\sigma}_{\ell,k}^2 \log(n)/T_{\ell,k}}$ . Fig. 4.5(a) shows that for both MAB and HOO,  $\alpha_{\text{norm}} = 0$  provides the best performance. Fig. 4.5(b) compares the performance of HOO for  $\ell_{\text{max}} = 2, 3$  and 4 with  $\alpha_{\text{norm}} = 0$ . We can see that a larger  $\ell_{\text{max}}$  improves the gains, which is expected since it allows a more refined search. Remarkably, thanks to the structure of the search tree, a larger  $\ell_{\text{max}}$  does not require more cost in the exploration. Since the performance improvement is quite small and the number of nodes in the tree increases quickly, we use  $\ell_{\text{max}} = 3$  from now on.

Fig. 4.6 compares the HOO beam refinement with and without the smoothness coefficient  $\nu(\ell)$ . The performance difference is negligible. This is likely because the search region in our problem is already confined to a small local region (within the 3 dB beamwidths of the selected beam pair), so that the constraint derived from the smoothness property does not have much value. This has a welcoming implication. The algorithm can be expected to be robust to small irregularity in the detailed shape of the beam patterns (thus affecting the exact smoothness property), which can be expected with real hardware.

### 4.7.3 Integrated Online Learning Solution

In this subsection, we evaluate the performance when combining the beam pair selection and refinement together. One thing that needs to be specified when combining the two is when to start the refinement for a selected beam pair. We consider the following three variations to start the beam refinement:

1. Refine all: The beam pair refinement is started for any beam pair from the first time it is selected by the online selection algorithm. This is the most straightforward way to combine the two components.
2. Refine after  $X_{\text{tot}}[i] > 0$ : The refinement of the beam pair  $i$  starts from the time step that the beam pair  $i$  receives a reward, i.e., when  $X_{\text{tot}}[i]$  becomes positive. The point for this option is that the algorithm only refines those beam pairs deemed to be most promising.



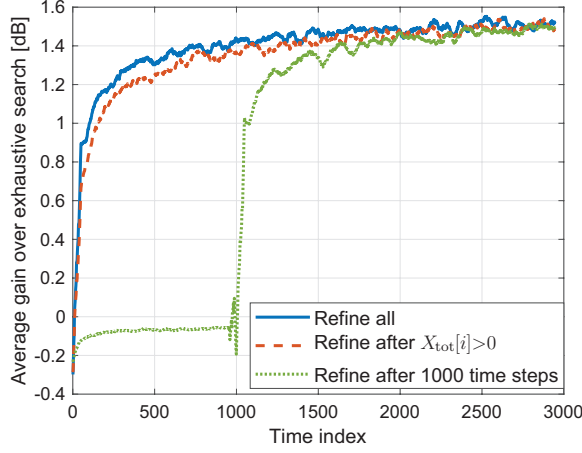


Figure 4.7: A comparison of average gain of the integrated solution with the three options for when to start the beam refinement. The plots show no negative impact of the beam pair refinement on the online learning for beam pair selection. It is best to start the refinement simultaneously with the online beam pair selection.

3. Refine after  $n_0$  time steps: The beam refinement of all selected beam pairs starts after running the online beam pair selection for  $n_0$  time steps. The rationale for this option is to prevent the beam pair refinement algorithm from affecting the learning of the beam pair selection algorithm. This option allows the beam pair selection to run for a while so that it stabilizes to some extent before starting the beam pair refinement.

While it seems more efficient to focus the refinement on promising beam pairs only as in Option 2, refining suboptimal beam pairs as well will maximize their average received signal and could reduce the risk of large power loss. Thus, it is not clear which option provides the best performance.

Fig. 4.7 compares the average gains over the exhaustive search (on the original codebook) of the three options. Here,  $B_{\text{tr}} = 30$ ,  $\Gamma_{\text{risk}} = 5$  dB,  $\ell_{\text{max}} = 3$ ,  $\alpha_{\text{norm}} = 0$ , and no smoothness coefficient is used (i.e.,  $\nu(\ell) = 1$ ). We can see that the first option, which is also the most straightforward one, provides the best performance. Focusing just on promising beam directions as in the second option performs quite well but is slightly worse than the first option. The results show that there is no benefit in waiting for some time before enabling the beam refinement as in the third option.

## 4.8 Conclusion

In this chapter, we proposed position-based online learning algorithms for beam pair selection and refinement. We used the MAB framework to develop a risk-aware greedy UCB algorithm for beam pair selection and a modified HOO for the beam pair refinement. Combining the two solutions together, we can gain up to about 1.5 dB over the received power obtained by exhaustive search over the original beam codebook before refinement. The learning is fast and it achieves an average gain of about 1 dB within the first 100 time steps. While we only use position in this chapter, more side information from sensors on the vehicle or the RSU about the current environment will help further reduce the beam training overhead. As shown in this chapter, even efficient learning algorithms can be impractical without risk-awareness because the focus is on cumulative rather than instantaneous performance. Therefore, we believe risk-awareness is a key to developing practical online

learning solutions to take full advantage of these sensors to enable fast and efficient mmWave communications.

## 4.9 Appendix

### 4.A Proof of Theorem 4.1

The regret is non-zero when one or more suboptimal beam pairs are selected. Thus, the total expected regret can be bounded by the average number of times suboptimal pairs are selected. We note that this derivation follows the steps of the UCB1 derivation from [15, Theorem 1] with the exception of the multiple-play setting. We provide the full details for completeness and readability. Denote  $\ell$  and  $i^*$  the indices of a suboptimal and optimal pair. Denote  $T_{\ell,i^*}[n]$  the number of times  $\ell$  is selected instead of  $i^*$  up to time  $n$ , the expected regret is

$$R_{\ell,i^*}[n] = \mathbb{E}[T_{\ell,i^*}[n]] \Delta_{\ell,i^*}, \quad (4.27)$$

where  $\Delta_{\ell,i^*}$  is the optimality gap defined in (4.16). This is because whether  $\ell$  is selected or not at time  $n$  depends on the rewards up to time  $n - 1$  and the loss depends only on the rewards at  $n$ . Thus, the two are independent by the assumption of independent reward signals across time (typical in an MAB setting).

We will now compute a bound for  $\mathbb{E}[T_{\ell,i^*}[n]]$ . A *necessary* condition for the pair  $\ell$  to be selected instead of the pair  $i^*$  is that  $\text{UCB}_\ell > \text{UCB}_{i^*}$ . After the pair  $\ell$  has been selected  $T_{\ell,i^*}^{(0)}$  times, the number of times  $\ell$  is selected instead

of  $i^\star$  up to time  $n$  can be bounded by

$$T_{\ell, i^\star}[n] \leq T_{\ell, i^\star}^{(0)} + \sum_{t=t_0}^n \mathbf{1} \left\{ \text{UCB}_\ell \geq \text{UCB}_{i^\star}, T_\ell(t-1) \geq T_{\ell, i^\star}^{(0)} \right\} \quad (4.28)$$

$$= T_{\ell, i^\star}^{(0)} + \sum_{t=t_0}^n \mathbf{1} \left\{ \hat{P}_{\text{opt}}(\ell) + c_{t-1, T_\ell(t-1)} \geq \hat{P}_{\text{opt}}(i^\star) + c_{t-1, T_{i^\star}(t-1)}, T_\ell(t-1) \geq T_{\ell, i^\star}^{(0)} \right\} \quad (4.29)$$

$$\leq T_{\ell, i^\star}^{(0)} + \sum_{t=t_0}^n \mathbf{1} \left\{ \max_{T_{\ell, i^\star}^{(0)} < u_\ell < t} \left\{ \hat{P}_{\text{opt}}(\ell) + c_{t-1, u_\ell} \right\} \geq \min_{0 < u < t} \left\{ \hat{P}_{\text{opt}}(i^\star) + c_{t-1, u} \right\} \right\} \quad (4.30)$$

$$\leq T_{\ell, i^\star}^{(0)} + \sum_{t=1}^{n-1} \sum_{u=1}^{t-1} \sum_{u_\ell=T_{\ell, i^\star}^{(0)}}^{t-1} \mathbf{1} \left\{ \hat{P}_{\text{opt}}(\ell) + c_{t, u_\ell} \geq \hat{P}_{\text{opt}}(i^\star) + c_{t, u} \right\}. \quad (4.31)$$

where  $t_0 \geq T_{\ell, i^\star}^{(0)}$  and  $c_{t, u} = \sqrt{2 \log(t)/u}$  is the confidence margin. For  $\{\hat{P}_{\text{opt}}(\ell) + c_{t, u_\ell} \geq \hat{P}_{\text{opt}}(i^\star) + c_{t, u}\}$  to be true, at least one of the followings must hold

$$\hat{P}_{\text{opt}}(i^\star) \leq P_{\text{opt}}(i^\star) - c_{t, u} \quad (4.32)$$

$$\hat{P}_{\text{opt}}(\ell) \geq P_{\text{opt}}(\ell) + c_{t, u_\ell} \quad (4.33)$$

$$P_{\text{opt}}(i^\star) < P_{\text{opt}}(\ell) + 2c_{t, u_\ell}. \quad (4.34)$$

Note that (4.32) means the UCB value underestimates the true reward of pair  $i^\star$ , and (4.33) means the UCB value overestimates the true reward of pair  $\ell$  by larger than the corresponding confidence margins. Setting  $T_{\ell, i^\star}^{(0)} = \lceil 8 \log(n) / \Delta_{\ell, i^\star}^2 \rceil$ , it can be shown that (4.34) is impossible [15, p. 243], and we

can bound  $\mathbb{E}[T_{\ell,i^*}[n]]$  by

$$\mathbb{E}[T_{\ell,i^*}[n]] \leq \left\lceil \frac{8 \log(n)}{\Delta_{\ell,i^*}^2} \right\rceil + \sum_{t=1}^{n-1} \sum_{u=1}^{t-1} \sum_{u_\ell=T_{\ell,i^*}^{(0)}}^{t-1} \mathbb{P}[(4.32) \text{ is true}] + \mathbb{P}[(4.33) \text{ is true}] \quad (4.35)$$

$$\leq \left\lceil \frac{8 \log(n)}{\Delta_{\ell,i^*}^2} \right\rceil + \sum_{t=1}^{\infty} \sum_{u=1}^{t-1} \sum_{u_\ell=T_{\ell,i^*}^{(0)}}^{t-1} (t^{-4} + t^{-4}) \quad (4.36)$$

$$\leq \left\lceil \frac{8 \log(n)}{\Delta_{\ell,i^*}^2} \right\rceil + 2 \sum_{t=1}^{\infty} \sum_{u=1}^t \sum_{u_\ell=1}^t t^{-4} \quad (4.37)$$

$$\leq \left\lceil \frac{8 \log(n)}{\Delta_{\ell,i^*}^2} \right\rceil + 2 \frac{\pi^2}{6}. \quad (4.38)$$

The second line in (4.36) follows because the probability terms can be shown to be bounded by  $t^{-4}$  using the Chernoff-Hoeffding inequality [15].

The total regret bound follows by summing all pairs of optimal and suboptimal beam pairs

$$R_1[n] \leq \sum_{\ell \in \hat{\mathcal{B}} \setminus \mathcal{S}^*} \sum_{i^* \in \mathcal{S}^*} \mathbb{E}[T_{\ell,i^*}[n]] \Delta_{\ell,i^*}. \quad (4.39)$$

Substituting (4.38) in and after some algebra, we obtain (4.17).

## 4.B Proof of Theorem 4.2

The derivation follows similarly to that of Theorem 4.1, but we need to be careful about the rejection mechanism. Even if a pair is selected by the greedy UCB selection, it will not be used for the training if it is rejected. Let  $\tilde{T}_{\ell,i^*}$  and  $T_{\ell,i^*}$  be the number of times the pair  $\ell$  is selected instead of the pair  $i^*$

and the number of times it is accepted for beam training, respectively. We proceed similarly to obtain a bound similar to (4.31) given by

$$\tilde{T}_{\ell,i^*}[n] \leq \tilde{T}_{\ell,i^*}^{(0)} + \sum_{t=1}^{n-1} \sum_{u=1}^{t-1} \sum_{\tilde{u}_\ell = \tilde{T}_{\ell,i^*}^{(0)}}^{t-1} \mathbf{1}\{\hat{P}_{\text{opt}}(\ell) + c_{t,u_\ell} \geq \hat{P}_{\text{opt}}(i^*) + c_{t,u}\}. \quad (4.40)$$

To compute the bound on  $\mathbb{E}[\tilde{T}_{\ell,i^*}[n]]$ , we again use (4.32)-(4.34). The probability bounds on (4.32) and (4.33) are still applicable. Because of the rejection, we cannot guarantee that (4.34) is impossible, but we can bound its probability. Note that  $\tilde{u}_\ell$  is the number of times the pair  $\ell$  is selected, and  $u_\ell$  is the number of times it is accepted for beam training. It can be shown that (4.34) is impossible if  $u_\ell > \frac{8 \log(t)}{\Delta_{\ell,i^*}^2}$  [15, p. 243]. Thus, we can bound the probability that (4.34) holds by  $\mathbb{P}\left[u_\ell \leq \frac{8 \log(t)}{\Delta_{\ell,i^*}^2}\right]$ . With the acceptance probability  $\zeta_\ell$ , we have  $\mathbb{E}[u_\ell] = \mathbb{E}[\tilde{u}_\ell]\zeta_\ell$ . Setting  $\tilde{T}_{\ell,i^*}^{(0)} = \lceil 8 \log(n)/(\zeta_\ell \delta^2 \Delta_{\ell,i^*}^2) \rceil$ , we get the following bound

$$\mathbb{P}\left[u_\ell \leq \frac{8 \log(t)}{\Delta_{\ell,i^*}^2} \middle| \tilde{u}_\ell = t\right] \leq \mathbb{P}\left[u_\ell \leq \frac{8 \log(t)}{\Delta_{\ell,i^*}^2} \middle| \tilde{u}_\ell = \tilde{T}_{\ell,i^*}^{(0)}\right] \quad (4.41)$$

$$\leq \mathbb{P}\left[u_\ell \leq \frac{8 \log(n)}{\Delta_{\ell,i^*}^2} \middle| \tilde{u}_\ell = \tilde{T}_{\ell,i^*}^{(0)}\right] \quad (4.42)$$

$$\leq n^{-4}. \quad (4.43)$$

Here, (4.41) follows because  $t \geq t_0 \geq \tilde{T}_{\ell,i^*}^{(0)}$ , (4.42) holds because  $t < n$ , and (4.43) is the application of the lower tail of the Chernoff bound with  $\delta = (\sqrt{5} - 1)/2$  for the Bernoulli distribution [39, Theorem 4]. Taking the expectation of

(4.40) and substitute the probability bounds for (4.32)-(4.34) to hold, we get

$$\mathbb{E} [\tilde{T}_{\ell, i^*}[n]] \leq \tilde{T}_{\ell, i^*}^{(0)} + \sum_{t=1}^{n-1} \sum_{u=1}^{t-1} \sum_{\tilde{u}_\ell = \tilde{T}_{\ell, i^*}^{(0)}}^{t-1} (2t^{-4} + n^{-4}) \quad (4.44)$$

$$\leq \tilde{T}_{\ell, i^*}^{(0)} + 3 \sum_{t=1}^{\infty} \sum_{u=1}^t \sum_{\tilde{u}_\ell=1}^t t^{-4} \quad (4.45)$$

$$= \tilde{T}_{\ell, i^*}^{(0)} + \pi^2/2. \quad (4.46)$$

To obtain the regret, we note that when the pair  $\ell$  is selected the regret incurred is  $\zeta_\ell \Delta_{\ell, i^*} + (1 - \zeta_\ell) \tilde{\Delta}_{\ell, i^*}$  because when rejected (with probability  $1 - \zeta_\ell$ ), the regret is  $\tilde{\Delta}_{\ell, i^*}$ . The total expected regret is then

$$R_2[n] \leq \sum_{\ell \in \hat{\mathcal{B}} \setminus \mathcal{S}^*} \sum_{i^* \in \mathcal{S}^*} \mathbb{E}[\tilde{T}_{\ell, i^*}[n]] (\zeta_\ell \Delta_{\ell, i^*} + (1 - \zeta_\ell) \tilde{\Delta}_{\ell, i^*}), \quad (4.47)$$

which after rearranging terms will result in (4.18).

#### 4.C Proof of Lemma 4.1

Assuming a normalized PAS, then the single-path PAS can be represented by the delta function  $\delta(\phi - \phi^*)$ . The average received power can be written as

$$\bar{\gamma}(\phi_0) = \int_{\phi_0 - \Delta\phi}^{\phi_0 + \Delta\phi} \delta(\phi - \phi^*) G(\phi; \phi_0) d\phi. \quad (4.48)$$

Since the gain  $G(\phi; \phi_0)$  is decreasing in  $[\phi_0, \phi_0 + \Phi/2]$  and  $|\phi^* - \phi_0| \leq \Delta\phi \leq \Phi/2$  by the assumption of the Lemma,

$$\frac{G(\phi; \phi_0)}{G(\phi_0 + \Delta\phi; \phi_0)} \geq 1 \geq G(\phi; \phi^*), \quad \forall \phi \in [\phi_0 - \Delta\phi, \phi_0] \cup [\phi_0, \phi_0 + \Delta\phi]. \quad (4.49)$$

Multiply both sides by  $\delta(\phi - \phi^*)$  and take the integral to get

$$\int_{\phi_0 - \Delta\phi}^{\phi_0 + \Delta\phi} \delta(\phi - \phi^*) G(\phi; \phi_0) / G(\phi_0 + \Delta\phi; \phi_0) d\phi \geq \int_{\phi_0 - \Delta\phi}^{\phi_0 + \Delta\phi} \delta(\phi - \phi^*) G(\phi; \phi^*) d\phi \quad (4.50)$$

$$\bar{\gamma}(\phi_0) / G(\phi_0 + \Delta\phi; \phi_0) \geq \bar{\gamma}(\phi^*). \quad (4.51)$$

Now, for a more general PAS  $\mathcal{P}(\phi)$  with a bounded support in  $[\phi_0 - \Psi, \phi_0 + \Psi]$ , the average received power can be written as

$$\bar{\gamma}(\phi_0) = \int_{\phi_0 - \Psi}^{\phi_0 + \Psi} \mathcal{P}(\phi) G(\phi; \phi_0) d\phi \quad (4.52)$$

$$\begin{aligned} &= \int_{\phi_0 - \Psi}^{\phi_0 - \Delta\phi} \mathcal{P}(\phi) G(\phi; \phi_0) d\phi + \int_{\phi_0 - \Delta\phi}^{\phi_0 + \Delta\phi} \mathcal{P}(\phi) G(\phi; \phi_0) d\phi \\ &\quad + \int_{\phi_0 + \Delta\phi}^{\phi_0 + \Psi} \mathcal{P}(\phi) G(\phi; \phi_0) d\phi. \end{aligned} \quad (4.53)$$

By the same argument as in the single-path PAS case, we have

$$\mathcal{P}(\phi) G(\phi; \phi_0) \frac{G(\phi_0; \phi_0)}{G(\phi_0 + \Delta\phi; \phi_0)} \geq \mathcal{P}(\phi) G(\phi; \phi^*), \quad \forall \phi \in [\phi_0 - \Delta\phi, \phi_0 + \Delta\phi] \quad (4.54)$$

$$\begin{aligned} \mathcal{P}(\phi) G(\phi; \phi_0) \frac{G(\phi_0 + \Psi - \Delta\phi; \phi_0)}{G(\phi_0 + \Psi; \phi_0)} &\geq \mathcal{P}(\phi) G(\phi; \phi^*), \\ \forall \phi &\in [\phi_0 - \Psi, \phi_0 - \Delta\phi] \cup [\phi_0 + \Delta\phi, \phi_0 + \Psi]. \end{aligned} \quad (4.55)$$

Taking the integral of (4.54) and (4.55), we have

$$\bar{\gamma}(\phi_0) / G(\phi_0 + \Delta\phi) \geq \bar{\gamma}(\phi^*) - \text{Err} \quad (4.56)$$

where

$$\begin{aligned} \text{Err} &= \left( \frac{G(\phi_0 + \Psi - \Delta\phi; \phi_0)}{G(\phi_0 + \Psi; \phi_0)} - \frac{1}{G(\phi_0 + \Delta\phi; \phi_0)} \right) \times \\ &\quad \left( \int_{\phi_0 - \Psi}^{\phi_0 - \Delta\phi} \mathcal{P}(\phi) G(\phi; \phi_0) d\phi + \int_{\phi_0 + \Delta\phi}^{\phi_0 + \Psi} \mathcal{P}(\phi) G(\phi; \phi_0) d\phi \right). \end{aligned} \quad (4.57)$$



Because  $G(\phi; \phi_0)$  is decreasing and concave for  $\phi \in [\phi_0, \phi_0 + \Phi/2]$  by the assumptions of the Lemma (e.g., true for a uniform linear array), the coefficient is positive and decreasing as  $\Delta\phi$  decreases. Since the integrands are positive by definition,  $\mathbf{Err}$  is positive. Further, because the sum of the integrals in (4.57) is upper bounded by  $\bar{\gamma}(\phi_0)$  (which is finite), we have  $\mathbf{Err} \rightarrow 0$  as  $\Delta\phi \rightarrow 0$ . Also, for small  $\Psi$  the integration intervals decrease and when  $\Psi \leq \Delta\phi$  they disappear, i.e.,  $\mathbf{Err} = 0$ .

## Chapter 5

### Conclusion

The high data rate capability of mmWave can enable a whole new range of V2X applications including both safety and non-safety related ones. One difficulty in using mmWave is the need to use directional beams to overcome its propagation characteristics, which is especially challenging in a mobile setting such as V2X. In this dissertation, focusing on the analog beamforming architecture, we addressed this challenge by taking advantage of the side information available from sensors equipped on vehicles. Specifically, using position and past beam measurements, we developed both offline and online learning algorithms that can rank beam directions by their likelihood to support a strong link connection. Performance evaluation of a standard solution defined in IEEE 802.11ad shows that it cannot support large arrays under high mobility, while the lower overhead of our proposed solutions allows the use of such large arrays, which is necessary to support decent link distances (e.g., tens to a few hundred meters). This proves our thesis statement that

*Position-based learning can exploit the propagation characteristics of the environment to reduce mmWave link configuration overhead.*

A summary of the main contributions presented in this dissertation now follows.

## 5.1 Summary

- **Chapter 2:** We presented a theoretical investigation of the impact of the directionality on the temporal variation of vehicular channels. We derived closed-form expressions for channel coherence time incorporating for the first time both the Doppler effect and pointing error due to mobility. Contrary to classical results, due to pointing error, there is a non-zero optimal beamwidth that maximizes the channel coherence time. We also proposed a new concept called the beam coherence time useful for quantifying the beam alignment overhead. It is shown that accounting for both the channel estimation and beam alignment overhead, beam realignment should be done every *beam* coherence time and not every *channel* coherence time.
- **Chapter 3:** We proposed a new framework for fast beam alignment suitable for V2I settings. The core of this framework is the beam pair selection method that selects promising beam directions eliminating unnecessary training overhead. We proposed to leverage position information available from onboard sensors and past beam measurements to learn to select beam pairs. Three beam selection methods are proposed: two based on statistical learning and one based on ML. Numerical evaluations using realistic channels from ray-tracing simulations show that the

proposed methods have an order of magnitude lower training overhead as compared to the IEEE 802.11ad solution.

- **Chapter 4:** We developed an online version of the optimal beam pair selection that minimizes the misalignment probability. We use the UCB idea in the multiple-play MAB setting and combine with a newly introduced risk-aware feature designed to avoid severe misalignment during the learning. We also developed an online beam pair refinement algorithm that can adapt the beam codebook to the environment to further maximize the beamforming gain. The refinement algorithm is based on HOO with modifications to account for the specific characteristics of mmWave beam alignment. An integrated solution combining the two learning components shows a fast learning behavior that quickly achieves positive gains over exhaustive beam search on the unrefined codebook.

## 5.2 Future Research Directions

This section describes several research directions related to the work represented in this dissertation.

### 5.2.1 Experimental Verification of Impact of Beamwidth on Channel Temporal Variation

In Chapter 2, we theoretically investigated the impact of the beamwidth on the variation of the instantaneous channel coefficient (captured by the channel coherence time) and the average received power (captured by the beam

coherence time). While the results make theoretical sense, verification in real measurements is lacking in the literature. The main difficulty in measuring the Doppler effect on the channel coefficient is the need to keep track of the alignment of the beams as the alignment condition will affect the PAS which in turn affects the Doppler spread. While one could use position information to help keep the alignment in the LOS case, it is not clear what is a good strategy for the NLOS case. For the beam coherence part, the difficulty lies in the repeatability of the measurement setting to ensure that the PAS (without the antenna patterns) stays unchanged during the measurements.

### **5.2.2 Learning for Multi-User Beam Alignment**

In this dissertation, we focus on the single-user scenario and show that position information is helpful in reducing the link configuration overhead. We expect position information to be beneficial for the multi-user scenario as well, but there will be different challenges. For example, beam directions that are good for each individual user could result in suboptimal performance when those users are served simultaneously due to interference among themselves. There are also opportunities for improving efficiency. For example, if a group of users shares multiple beam directions in their respective selection set (not necessarily in the same ranking), those beam directions can be trained simultaneously with one single training packet, which will reduce the overhead as compared to training each user individually. There are two research questions here. One is how do we learn to predict the level of interference when serving

a group of users together and the other is how do we optimize the grouping of users to minimize the training overhead?

### 5.2.3 Learning for Beam Tracking

Our proposed beam alignment methods have a low overhead that could allow it to be used for tracking as well. For example, by inserting one or several beam measurements to test a few beam pairs in the database in each transmission round (i.e., using the in-packet training as defined in IEEE 802.11ad), we can keep track of the best beam pair. There is, however, an opportunity for further improvement by leveraging the temporal correlation between the beam pairs across time instances as the vehicle moves. The question here is how to learn this temporal correlation with as little effort as possible? One idea is to initially run the tracking as outlined earlier after the learning for the beam alignment and start collecting temporal data. That is, the tracking learning is conducted after the beam alignment learning. While the beam alignment algorithm provides only a fixed ranking of beam pairs, the tracking algorithm can use the best beam pair sequence seen so far as additional context to alter the ranking to fit the current situation (i.e., exploiting the temporal correlation), and thus it can help further reduce the beam training overhead. Reinforcement learning framework that can take prior belief from the beam alignment learning could be an interesting starting point. The ultimate goal would be to integrate the learning into one single learning algorithm without having a separate beam alignment learning phase.

#### 5.2.4 Online Contextual Bandit for Beam Alignment

As more sensors are equipped on vehicles, more and more context information beyond position will become available. Thus, it is desirable to extend our risk-aware online learning to a framework that can easily incorporate multiple types of side information as we did in the LtR approach in an offline learning setting. Of course, one can define multi-dimensional context bins and define an MAB instance for each bin and apply our risk-aware learning algorithm. This is, however, not necessarily an efficient way to do this for the reason discussed in Section 3.8.1. Contextual bandit is a promising framework for developing such a solution.

#### 5.2.5 Extension to Hybrid Architectures

Equipped with several RF chains, a hybrid architecture can support multiple data streams and/or provide beam diversity leading to superior performance over that of an analog architecture. It, however, also requires more effort in link configuration. Our learning approaches so far can identify promising directions and thus is also useful for a hybrid architecture. For example, the knowledge of promising beam directions can be used as a prior for selecting efficient measurement matrices in a compressive sensing solution [9]. With a more flexible architecture, we can also expect that more efficient learning solutions will be possible. The first question would be how to specify the set of arms. In the analog beamforming case, we use beams generated by progressive phase-shift. More sophisticated set of arms are likely needed to take

full advantage of the multiple RF chains available in a hybrid architecture. It might also depend on the beamforming strategy. For example, we could use a two-step approach where the transceiver first estimates the channel and then calculates the beamformer. In this case, if the channel estimation uses a compressive sensing idea, then one approach would be to learn the probability of selecting a codeword in a dictionary used in a compressive sensing solution. Another possible strategy is to learn to directly choose beamformers. In that case, a different set of arms might be more appropriate. In either case, it is important to design some metric with simplifying property such as the modularity of the probability of being optimal in our case. This way, one can reduce the complexity of the selection of a subset of arms, which in general is an intractable problem.

### **5.2.6 Personalized MmWave Link Configuration**

Being able to collect performance data for a specific context opens up opportunities for customizing or adapting the communication systems to suit that context. In this dissertation, the beam refinement algorithm has its premise in this idea where the codebook is adapted to the propagation environment conditioned by the position. This idea can be applied more generally such as for mmWave link configuration for handheld devices. For example, information from the sensors on the device can provide clues on the orientation of the antenna array. The orientation information can be used to translate the pointing angle to a standard coordinate, and the learning approaches we



presented can still be used. Beyond this one can consider learning the user's habit of using the device. For example, how the user holds the device when running a particular application (e.g., voice call versus playing games) can infer which beam directions are likely blocked by the user's hand or head. Thus, this would allow personalizing the communication systems to the users for optimal performance. As noted in the conclusion in Chapter 4, to do such learning online, exploring too much could be a potential pitfall. Some form of risk-awareness is crucial to avoid large performance losses, which is expected to be severe during the early stage of the learning.

## Bibliography

- [1] 3rd Generation Partnership Project (3GPP). Study on LTE support for vehicle to everything (V2X) services. 3GPP TR 22.885, December 2015.
- [2] 3rd Generation Partnership Project (3GPP). 3GPP study on LTE-based V2X services. TR 36.885, July 2016.
- [3] 3rd Generation Partnership Project (3GPP). Study on enhancement of 3GPP support for 5G V2X services. 3GPP TR 22.886, March 2017.
- [4] A. Maltsev *et al.* MiWEBA D5.1: Channel modeling and characterization. Technical report, MiWEBA project, June 2014.
- [5] W. B. Abbas and M. Zorzi. Context information based initial cell search for millimeter wave 5G cellular networks. In *Proc. Eur. Conf. on Netw. and Commun. (EuCNC)*, pages 111–116, June 2016.
- [6] Milton Abramowitz. *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. Dover Publications, 1974.
- [7] Agilent. Wireless LAN at 60 GHz - IEEE 802.11ad explained. <http://www.cs.odu.edu/~cs752/papers/milli-008.pdf>, 2013. Application Note.

- [8] Ali S. Al-Ghamdi. Analysis of time headways on urban roads: Case study from riyadh. *Transport Eng-J*, 127(4):289–294, 2001.
- [9] Anum Ali and Robert W. Heath Jr. Compressed beam selection in millimeter wave systems with out-of-band partial support information. In *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, 2017.
- [10] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic. Deep learning coordinated beamforming for highly-mobile millimeter wave systems. *IEEE Access*, 6:37328–37348, June 2018.
- [11] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath Jr. MIMO precoding and combining solutions for millimeter-wave systems. *IEEE Commun. Mag.*, 52(12):122–131, December 2014.
- [12] Allied Vision. Prosilica GT 2050. <https://www.alliedvision.com/en/products/cameras/detail/2050.html>. Accessed: Mar. 30, 2016.
- [13] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice-Hall, 1979.
- [14] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro. LTE for vehicular networking: A survey. *IEEE Commun. Mag.*, 51(5):148–157, May 2013.
- [15] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47:235–256, 2002.

- [16] J. C. Aviles and A. Kouki. Position-aided mm-wave beam training under NLOS conditions. *IEEE Access*, 4:8703–8714, November 2016.
- [17] P. Bahl and V. N. Padmanabhan. RADAR: an in-building RF-based user location and tracking system. In *Proc. IEEE Int. Conf. on Comp. and Comm. (INFOCOM)*, pages 775–784, March 2000.
- [18] Constantine A. Balanis. *Antenna Theory: Analysis and Design*. Wiley-Interscience, May 2005.
- [19] S. Bubeck and N. Cesa-Bianchi. *Regret Analysis of Stochastic and Non-stochastic Multi-armed Bandit Problems*. Foundations and Trends in Machine Learning. Now Publishers, 2012.
- [20] Sébastien Bubeck, Rémi Munos, Gilles Stoltz, and Csaba Szepesvári. X-armed bandits. *J. Mach. Learn. Res.*, 12:1655–1695, July 2011.
- [21] Giuseppe Burtini, Jason Loeppky, and Ramon Lawrence. A survey of online experiment design with the stochastic multi-armed bandit. <https://arxiv.org/abs/1510.00757>, November 2015.
- [22] Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM J. Comput.*, 40(6):1740–1766, December 2011.
- [23] Mingzhe Chen, Ursula Challita, Walid Saad, Changchuan Yin, and Merouane Debbah. Machine learning for wireless networks with artificial

intelligence: A tutorial on neural networks. <https://arxiv.org/abs/1710.02913>, October 2017.

- [24] Wei Chen, Yajun Wang, and Yang Yuan. Combinatorial multi-armed bandit: General framework and applications. In *Proc. Int. Conf. on Mach. Learn.*, pages 151–159, June 2013.
- [25] Xiang Cheng, Cheng-Xiang Wang, and D.I. Laurenson. A geometry-based stochastic model for wideband MIMO mobile-to-mobile channels. In *Proc. IEEE Global Commun. Conf.*, pages 1–6, November 2009.
- [26] D. Chizhik, J. Du, R. Feick, M. Rodriguez, G. Castro, and R. Valenzuela. Path loss, beamforming gain and time dynamics measurements at 28 GHz for 90% indoor coverage. Technical report, Bell Laboratories, Nokia, December 2017.
- [27] Dmitry Chizhik. Slowing the time-fluctuating MIMO channel by beamforming. *IEEE Trans. Wireless Commun.*, 3(5):1554–1565, 2004.
- [28] Junil Choi, Vutha Va, Nurial Gonzalez-Prelcic, Robert Daniels, Chandra R. Bhat, and Robert W. Heath Jr. Millimeter-wave vehicular communication to support massive automotive sensing. *IEEE Commun. Mag.*, 54(12):160–167, December 2016.
- [29] S. Collonge, G. Zaharia, and G. E. Zein. Influence of the human activity on wide-band characteristics of the 60 GHz indoor radio channel. *IEEE Trans. Wireless Commun.*, 3(6):2396–2406, November 2004.

- [30] J. J. Condon and S. M. Ransom. Antenna fundamentals. <http://www.cv.nrao.edu/course/astr534/AntennaTheory.html>. Accessed: Jun. 15, 2018.
- [31] F. Devoti, I. Filippini, and A. Capone. Facing the millimeter-wave cell discovery challenge in 5G networks with context-awareness. *IEEE Access*, 4:8019–8034, November 2016.
- [32] Q. Duan, T. Kim, H. Huang, K. Liu, and G. Wang. AoD and AoA tracking with directional sounding beam design for millimeter wave MIMO systems. In *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun.*, pages 2271–2276, August 2015.
- [33] Gregory D. Durgin and Theodore S. Rappaport. Theory of multipath shape factors for small-scale fading wireless channels. *IEEE Trans. Antennas Propag.*, 48(5):682–693, 2000.
- [34] Jack Edmonds. Matroids and the greedy algorithm. *Mathematical Programming*, 1(1):127–136, 1971.
- [35] Nariman Farsad and Andrea Goldsmith. Neural network detection of data sequences in communication systems. <https://arxiv.org/abs/1802.02046>, February 2018.
- [36] FHWA, U.S. Department of Transportation. NGSIM–Next Generation SIMulation. <http://gateway.path.berkeley.edu/ngsimdocs/>.

- [37] Catherine Forbes, Merran Evans, Nicholas Hastings, and Brian Peacock. *Statistical Distributions*. John Wiley & Sons, Inc., November 2010.
- [38] X. Gao, L. Dai, Y. Zhang, T. Xie, X. Dai, and Z. Wang. Fast channel tracking for terahertz beamspace massive MIMO systems. *IEEE Trans. Veh. Technol.*, 66(7):5689–5696, July 2017.
- [39] Michel Goemans. Chernoff bounds, and some applications. <http://math.mit.edu/~goemans/18310S15/chernoff-notes.pdf>, 2015. Lecture Note.
- [40] Andrea Goldsmith. *Wireless Communications*. Cambridge University Press, Cambridge, 2005.
- [41] Nuria Gonzalez-Prelcic, Roi Mendez-Rial, and Robert W. Heath Jr. Radar aided beam alignment in mmwave V2I communications supporting antenna diversity. In *Proc. Inf. Theory Appl. Workshop*, February 2016.
- [42] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, San Diego, 5th edition edition, 1994.
- [43] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes*. Oxford University Press, 3 edition, 2001.
- [44] Morteza Hashemi, Ashutosh Sabharwal, C. Emre Koksall, and Ness B. Shroff. Efficient beam alignment in millimeter wave systems using

- contextual bandits. <https://arxiv.org/abs/1712.00702>, April 2018.
- IEEE INFOCOM*.
- [45] Trevor J. Hastie, Robert John Tibshirani, and Jerome H Friedman. *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*. Springer series in statistics. Springer, New York, 2009.
  - [46] Robert W. Heath Jr. *Introduction to Wireless Digital Communication: A Signal Processing Perspective*. Prentice Hall, 1 edition, March 2017.
  - [47] Robert W. Heath Jr., Nuria Gonzalez-Prelcic, Sundeep Rangan, Wonil Roh, and Akbar Sayeed. An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE J. Sel. Topics Signal Process.*, 10(3):436–453, April 2016.
  - [48] IEEE. IEEE std 802.15.3c-2009. *IEEE Standard*, pages c1–187, October 2009.
  - [49] IEEE. IEEE std 802.11p-2010. *IEEE Standard*, pages 1–35, July 2010.
  - [50] IEEE. IEEE std 802.11ad-2012. *IEEE Standard*, pages 1–628, December 2012.
  - [51] International Telecommunication Union (ITU). Effects of building materials and structures on radiowave propagation above about 100 MHz. ITU-R P.2040-1, July 2015.



- [52] J. Levinson, *et al.* Towards fully autonomous driving: Systems and algorithms. In *Proc. IEEE Intell. Veh. Symp.*, pages 163–168, June 2011.
- [53] J. Wang *et al.* Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems. *IEEE J. Sel. Areas Commun.*, 27(8):1390–1399, 2009.
- [54] William C. Jakes and Donald C. Cox. *Microwave Mobile Communications*. Wiley-IEEE Press, September 1994.
- [55] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo. Machine learning paradigms for next-generation wireless networks. *IEEE Wirel. Commun.*, 24(2):98–105, April 2017.
- [56] Daniel Jiang, Qi Chen, and Luca Delgrossi. Optimal data rate selection for vehicle safety communications. In *Proc. 5th ACM Int. Workshop on Veh. Inter-Netw.*, pages 30–38, 2008.
- [57] Michael Jordan. The exponential family: Conjugate priors. <https://people.eecs.berkeley.edu/~jordan/courses/260-spring10/other-readings/chapter9.pdf>, 2009. Lecture Note.
- [58] K. Hosoya *et al.* Multiple sector ID capture (MIDC): A novel beamforming technique for 60-GHz band multi-Gbps WLAN/PAN systems. *IEEE Trans. Antennas Propag.*, 63(1):81–96, January 2015.

- [59] Tobias Kadur, Hsiao-lan Chiang, and Gerhard Fettweis. Effective beam alignment algorithm for low cost millimeter wave communication. In *Proc. IEEE Veh. Technol. Conf.*, pages 1–5, 2016.
- [60] J. B. Kenney. Dedicated short-range communications (DSRC) standards in the United States. *Proc. IEEE*, 99(7):1162–1182, July 2011.
- [61] Joongheon Kim and A. F. Molisch. Enabling Gigabit services for IEEE 802.11ad-capable high-speed train networks. In *Proc. IEEE Radio and Wireless Symp. (RWS)*, pages 145–147, January 2013.
- [62] Seong-Woo Kim, Baoxing Qin, Zhuang Jie Chong, Xiaotong Shen, Wei Liu, M.H. Ang, E. Frazzoli, and D. Rus. Multivehicle cooperative driving using cooperative perception: Design and experimental validation. *IEEE Trans. Intell. Transp. Syst.*, 16(2):663–680, April 2015.
- [63] Taejoon Kim and David J. Love. Virtual AoA and AoD estimation for sparse millimeter wave MIMO channels. In *Proc. IEEE Workshop on Signal Process. Adv. in Wireless Commun.*, pages 146–150, 2015.
- [64] Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. Multi-armed bandits in metric spaces. In *Proc. Annu. ACM Symp. on Theory of Comput.*, STOC ’08, pages 681–690, New York, NY, USA, 2008. ACM.
- [65] Evgeny Kupershtein, Mati Wax, and Israel Cohen. Single-site emitter localization via multipath fingerprinting. *IEEE Trans. Signal Process.*, 61(1):10–21, 2013.

- [66] Branislav Kveton, Csaba Szepesvári, Zheng Wen, and Azin Ashkan. Cascading bandits: Learning to rank in the cascade model. In *Proc. Int. Conf. on Mach. Learn.*, pages 767–776, 2015.
- [67] B. D. S. Lakmali and D. Dias. Database correlation for GSM location in outdoor & indoor environments. In *Proc. 4th Int. Conf. on Inform. and Autom. for Sustain.*, pages 42–47, December 2008.
- [68] Bin Li, Zheng Zhou, Haijun Zhang, and Arumugam Nallanathan. Efficient beamforming training for 60-GHz millimeter-wave communications: A novel numerical optimization framework. *IEEE Trans. Veh. Technol.*, 63(2):703–717, 2014.
- [69] Bin Li, Zheng Zhou, Weixia Zou, Xuebin Sun, and Guanglong Du. On the efficient beam-forming training for 60 GHz wireless personal area networks. *IEEE Trans. Wireless Commun.*, 12(2):504–515, 2013.
- [70] E. S. Li and K. Sarabandi. Low grazing incidence millimeter-wave scattering models and measurements for various road surfaces. *IEEE Trans. Antennas Propag.*, 47(5):851–861, May 1999.
- [71] C. Liu, M. Li, S. V. Hanly, I. B. Collings, and P. Whiting. Millimeter wave beam alignment: Large deviations analysis and design insights. *IEEE J. Sel. Areas Commun.*, 35(7):1619–1631, July 2017.
- [72] Tie-Yan Liu. Learning to rank for information retrieval. *Found. Trends Inf. Retr.*, 3(3):225–331, March 2009.

- [73] Zhinus Marzi, Dinesh Ramasamy, and Upamanyu Madhow. Compressive channel estimation and tracking for large arrays in mm-wave picocells. *IEEE J. Sel. Topics Signal Process.*, 10(3):514–527, 2016.
- [74] Kota Mase, Junji Inoue, and Junji Kizu. Performance evaluation of a roadside-to-vehicle communication system using narrow antenna beam switching based on traffic flow model. In *Proc. IEEE Global Commun. Conf.*, pages 1–5, 2008.
- [75] Muriel Médard. The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel. *IEEE Trans. Inf. Theory*, 46(3):933–946, 2000.
- [76] R. Mendez-Rial, C. Rusu, A. Alkhateeb, N. Gonzalez-Prelcic, and R. W. Heath Jr. Channel estimation and hybrid combining for mmwave: Phase shifters or switches? In *Proc. Inf. Theory Appl. Workshop*, pages 90–97, February 2015.
- [77] N. Moraitis and Philip Constantinou. Indoor channel measurements and characterization at 60 GHz for wireless local area network applications. *IEEE Trans. Antennas Propag.*, 52(12):3180–3189, December 2004.
- [78] Rmi Munos. From bandits to Monte-Carlo tree search: The optimistic principle applied to optimization and planning. *Found. Trends Mach. Learn.*, 7(1):1–129, 2014.

- [79] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14(1):265–294, 1978.
- [80] Thomas Nitsche, Adriana B. Flores, Edward W. Knightly, and Joerg Widmer. Steering with eyes closed: Mm-wave beam steering without in-band measurement. In *Proc. IEEE Int. Conf. on Comp. and Comm. (INFOCOM)*, volume 26, pages 2416–2424, 2015.
- [81] Ole Norklit and Rodney G Vaughan. Angular partitioning to yield equal Doppler contributions. *IEEE Trans. Veh. Technol.*, 48(5):1437–1442, 1999.
- [82] Sophocles J. Orfanidis. *Electromagnetic Waves and Antenna*. Rutgers University, 2014. online book <http://www.ece.rutgers.edu/~orfanidi/ewa/>.
- [83] Simon Pike. 5G access. <https://www.techuk.org/insights/meeting-notes/item/6870-meeting-notes-from-uk-spf-cluster-2-event-on-mmwave>, December 2015.
- [84] H.S. Rad and S. Gazor. The impact of non-isotropic scattering and directional antennas on MIMO multicarrier mobile communication channels. *IEEE Trans. Commun.*, 56(4):642–652, April 2008.
- [85] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir. Broadband millimeter-wave propagation measurements

- and models using adaptive-beam antennas for outdoor urban cellular communications. *IEEE Trans. Antennas Propag.*, 61(4):1850–1859, April 2013.
- [86] Theodore S. Rappaport, Robert W. Heath Jr., Robert C. Daniels, and James N. Murdock. *Millimeter Wave Wireless Communications*. Pearson, September 2014.
- [87] Remcom. Wireless InSite. <http://www.remcom.com/wireless-insite>.
- [88] S. Hur *et al.* Millimeter wave beamforming for wireless backhaul and access in small cell networks. *IEEE Trans. Commun.*, 61(10):4391–4403, October 2013.
- [89] Mathew K Samimi and Theodore S Rappaport. Ultra-wideband statistical channel model for non line of sight millimeter-wave urban channels. In *Proc. IEEE Global Commun. Conf.*, pages 3483–3489, 2014.
- [90] Maria Scalabrin, Nicolo Michelusi, and Michele Rossi. Beam training and data transmission optimization in millimeter-wave vehicular networks. <https://arxiv.org/abs/1804.10181>, April 2018.
- [91] J. Seo, Y. Sung, G. Lee, and D. Kim. Training beam sequence design for millimeter-wave MIMO systems: A POMDP framework. *IEEE Trans. Signal Process.*, 64(5):1228–1242, March 2016.

- [92] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proc. IEEE*, 104(1):148–175, January 2016.
- [93] Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proc. Int. Conf. on Mach. Learn.*, pages 1015–1022, 2010.
- [94] Richard S. Sutton and Andrew G. Barto. *Introduction to Reinforcement Learning*. MIT Press, Cambridge, MA, USA, 1st edition, 1998.
- [95] Paul D. Teal, T.D. Abhayapala, and R.A. Kennedy. Spatial correlation for general distributions of scatterers. *IEEE Signal Process. Lett.*, 9(10):305–308, October 2002.
- [96] The 5G Infrastructure Public Private Partnership (5GPPP). 5G automotive vision. <https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Automotive-Vertical-Sectors.pdf>, October 2015.
- [97] The European Commission. Commission implementing decision of 11 December 2013 amending decision 2006/771/EC on harmonisation of the radio spectrum for use by short-range devices and repealing decision 2005/928/EC. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32013D0752>, December 2013. EC Decision 2013/752/EU.

- [98] TOYOTA. Toyota to display new map generation system at CES 2016.  
<http://newsroom.toyota.co.jp/en/detail/10765074/>. Accessed: Jul. 19, 2018.
- [99] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [100] United States Department of Transportation. Intelligent transportation systems - connected vehicle. [http://www.its.dot.gov/pilots/cv\\_pilot\\_apps.htm](http://www.its.dot.gov/pilots/cv_pilot_apps.htm). Accessed: Mar. 19, 2016.
- [101] V. Nurmela *et al.* Deliverable d1.4 METIS channel models. Technical report, METIS Project, February 2015.
- [102] V. Va, J. Choi, and Robert W. Heath Jr. The impact of beamwidth on temporal channel variation in vehicular channels and its implications. *IEEE Trans. Veh. Technol.*, 66(6):5014–5029, May 2017.
- [103] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath. Inverse multipath fingerprinting for millimeter wave V2I beam alignment. *IEEE Trans. Veh. Technol.*, 67(5):4042–4058, May 2018.
- [104] V. Va, T. Shimizu, G. Bansal, and R. W. Heath. Position-aided millimeter wave V2I beam alignment: A learning-to-rank approach. In *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun.*, pages 1–5, October 2017.



- [105] Vutha Va and Robert W. Heath Jr. Basic relationship between channel coherence time and beamwidth in vehicular channels. In *Proc. IEEE Veh. Technol. Conf.*, pages 1–5, September 2015.
- [106] Vutha Va, Takayuki Shimizu, Gaurav Bansal, and Robert W. Heath Jr. Beam design for beam switching based millimeter wave vehicle-to-infrastructure communications. In *Proc. IEEE Int. Conf. on Commun.*, pages 1–6, May 2016.
- [107] Vutha Va, Takayuki Shimizu, Gaurav Bansal, and Robert W. Heath Jr. Millimeter wave vehicular communications: A survey. *Found. Trends Netw.*, 10(1), 2016.
- [108] Vutha Va, Takayuki Shimizu, Gaurav Bansal, and Robert W. Heath Jr. Online learning for position-aided millimeter wave beam training, September 2018. Submitted to *IEEE Access*.
- [109] Vutha Va, Xinchun Zhang, and Robert W. Heath Jr. Beam switching for millimeter wave communication to support high speed trains. In *Proc. IEEE Veh. Technol. Conf.*, pages 1–5, September 2015.
- [110] Velodyne. Velodyne LiDAR HDL-64E. <http://velodynelidar.com/hdl-64e.html>. Accessed: Jun. 15, 2018.
- [111] W. Roh *et al.* Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results. *IEEE Commun. Mag.*, 52(2):106–113, February 2014.

- [112] Junyi Wang, Zhou Lan, Chang woo Pyo, T. Baykas, Chin-Sean Sum, M.A Rahman, Jing Gao, R. Funada, F. Kojima, H. Harada, and S. Kato. Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems. *IEEE J. Sel. Areas Commun.*, 27(8):1390–1399, October 2009.
- [113] Yuyang Wang, Murali Narasimha, and Robert W. Heath Jr. Mmwave beam prediction with situational awareness: A machine learning approach. <https://arxiv.org/abs/1805.08912>, March 2018. to appear in *IEEE SPAWC*.
- [114] E. Ward and J. Folkesson. Vehicle localization with low cost radar sensors. In *Proc. IEEE Intell. Veh. Symp.*, pages 864–870, June 2016.
- [115] Eric W. Weisstein. Convolution. Accessed on 18 Jun. 2016.
- [116] J. Wenger. Automotive radar - Status and perspectives. In *Proc. Comp. Semicond. Integr.*, pages 21–24, 2005.
- [117] Y Shi *et al.* GAPfm: optimal top-n recommendations for graded relevance domains. In *Proc. ACM Conf. on Inform. and Knowledge Manage.*, pages 2261–2266, 2013.
- [118] H. Ye, L. Liang, G. Y. Li, J. Kim, L. Lu, and M. Wu. Machine learning for vehicular networks: Recent advances and application examples. *IEEE Veh. Technol. Mag.*, 13(2):94–101, June 2018.

- [119] Su Khiong Yong and Chia-Chin Chong. An Overview of Multigigabit Wireless through Millimeter Wave Technology: Potentials and Technical Challenges. *EURASIP J. Wireless Commun. and Netw.*, 2007:1–10, 2007.
- [120] A. G. Zajic and G. L. Stuber. Space-time correlated mobile-to-mobile channels: Modelling and simulation. *IEEE Trans. Veh. Technol.*, 57(2):715–726, March 2008.
- [121] Hang Zhao, R. Mayzus, Shu Sun, M. Samimi, J.K. Schulz, Y. Azar, K. Wang, G.N. Wong, F. Gutierrez, and T.S. Rappaport. 28 GHz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in New York city. In *Proc. IEEE Int. Conf. on Commun.*, pages 5163–5167, June 2013.

# Vita

Vutha Va received his BS and MS degree in Electrical and Electronic Engineering from Tokyo Institute of Technology in 2011 and 2013, respectively. He is currently a graduate student at the Wireless Networking and Communications Group at the University of Texas at Austin. He was a research intern at Toyota InfoTechnology Center, USA, Inc., Mountain View, CA, in summer 2015 and 2016. His research interest includes vehicular communications, millimeter wave communications, machine learning for communication systems, and the fifth generation (5G) cellular networks.

Permanent address: 1715 Guadalupe St.  
Austin, Texas 78701

This dissertation was typeset with  $\text{\LaTeX}^\dagger$  by the author.

---

<sup>†</sup> $\text{\LaTeX}$  is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's  $\text{\TeX}$  Program.