Millimeter Wave Beam Alignment Using Deterministic Compressive Sensing

By

Erfan Khordad

A thesis submitted to Macquarie University

for the degree of Master of Research

Department of Engineering

June 2019

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Erfan Khordad

Acknowledgements

I would like to express my sincere thanks to my supervisor, Prof. Stephen V. Hanly and my co-supervisor Prof. Iain B. Collings for their patience, motivation and continuous support of my research. I am also grateful that I have conducted my research under their supervision and expert guidance leading to writing this dissertation.

Abstract

Designing the beamforming vectors for channel estimation in mmWave systems is challenging because of the narrow beams required and the small number of useful directions. The state of the art employs random or structured random beamforming to leverage compressive sensing techniques to solve this problem using a small number of measurements. In this dissertation, inspired by existing deterministic sensing matrices from the theory of compressive sensing, two novel deterministic compressive sensing approaches are proposed for channel estimation in mmWave systems. In the proposed approaches, the Kronecker product or row-by-row Kronecker product of existing deterministic sensing matrices are used to design the structure of pilot beam patterns for the beam alignment process. These approaches not only result in significant overhead reduction, but also present improvement in terms of performance for some scenarios.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The next generation of wireless cellular networks (5G) aim to greatly increase the number of connected users and at the same time boost per user-data rates by orders of magnitude. It also aims to deliver new low latency services for machine to machine communications and improved spectral and energy efficiency for future networks. Previous generations of wireless systems have tried to cope with such demands; however, 5G's precursors were not sufficient to realize new concepts and ideas such as the Internet of Things and smart cities. In this regard, there have been extensive studies considering challenges ahead for 5G, and there are still many issues to be dealt with.

The main focus for developing 5G has been on five major technologies including device-centric architectures, smarter devices, machine-to-machine communications, massive multiple input multiple output (MIMO) and millimeter wave [1].

Device-centric architectures bypass connections to base stations (BSs) and allow direct data transfers between devices without intermediate nodes; therefore, they will change the traditional architecture. Device-centric architectures makes devices more active to have participation in exchanging data in the network rather than just being receive nodes. Device-to-Device (D2D) connections and multi-hop communications make up device-centric architectures. In D2D connections, devices act as a bridge to deliver coverage to other devices without following the traditional architecture where a link is established through a BS, uplink and downlink connections. This can reduce latency, avoid possibly strong pathloss in communicating with BS, and prevent waste of power in uplink and downlink directions. Devices can also act as relays in a multi-hop manner over long distances between BSs and in non-line of sight conditions, resulting in less pathloss and hence less energy consumption as compared to single-hop communications. Therefore, device-centric communications can have a significant impact upon future networks [2].

Smarter devices have a crucial role in realizing 5G concepts [1]. Nearby users can wirelessly exchange digital content through their devices. Managing such connection by the infrastructure side

might cause a waste of signalling resources and power in uplink and downlink. This shows the importance of upgrading devices to be able to participate in D2D communications. In addition, devices need to be more compatible with data processing and storage in cloud-based procedures so as to use the high volume of memory and processing resources shared in the cloud.

Machine-to-machine (M2M) communications is defined as exchanging information or communicating among machines operating as a network with low human involvement. There are some major challenges for M2M communications including delivering network coverage to a large number of low-rate devices, maintaining a minimum data rate in almost all conditions and transferring data with very low latency [1]. M2M communications is established for a wide variety of purposes such as measurements in industry or cooperating between sensors. The ideas of D2D and M2M might seem analogous at first glance. D2D is more associated with communications and the fact that the data should be directly exchanged among devices without BSs or intermediate nodes involvement. M2M is more associated with the application. In M2M, machines collaborate with each other in order to work as a team and to pursue a common objective. They might use access points or multi-hop communications across the wireless network to be connected to each other.

Massive MIMO is the idea of increasing the number of antennas by more than an order of magnitude in wireless systems exploiting multiple antennas at transceivers [3]. This technology has enormous potential to pave the road for 5G requirements owing to high spectral efficiency, nearly flat channel responses because of decreasing small-scale effects as a result of large number of channel measurements, and simple design of transceivers due to near orthogonality of the channels between each BS and users. In other words, massive MIMO is one of the key components in fulfilling 5G requirements. There are some major challenges for massive MIMO including pilot contamination, overhead reduction, antenna correlations and coupled measurements, and millimeter wave massive MIMO [3].

Millimeter Wave (mmWave) is another key component, providing the required spectrum for 5G [1][3]. Conventional wireless systems have operated at microwave frequencies from several hundred MHz to a few GHz with wavelengths in the range of a few centimeters up to about a meter. This spectrum is now quite congested and there is a need for more bandwidth. One of the promising solutions is using mmWave frequencies for communications systems to satisfy the demand for more bandwidth. Challenges regarding mmWave systems such as propagation-related issues, efficient beamforming designs and channel characterization should be addressed to take full advantage of mmWave frequencies. This thesis focuses on challenges and signal processing problems concerning

mmWave systems.

1.1 mmWave Systems

One of the potential solutions to provide a large amount of bandwidth is utilizing mmWave frequencies with the wavelength between 1 millimeter to 10 millimeters, ranging from 30 to 300 GHz. Some considerations in the past claimed that employing mmWave might be impractical owing to adverse propagation characteristics of mmWave signals. However, in recent years, there has been much progress in mmWave circuit design resulting in more cost- and power-effective communications devices. In fact, high frequency electronics allows new signal processing techniques (such as discussed in this thesis) to help overcome the challenges of the mmWave channels [3].

Propagation characteristics related to the nature of mmWaves are now the major challenge and pose difficulties including pathloss and attenuation owing to rain, foliage and atmospheric absorption. Pathloss increases with the frequency squared; thus, increasing the carrier frequency by an order of magnitude for example from 3 to 30 GHz, results in 20 dB power loss without regard to the distance between transmitter and receiver [3]. Oxygen absorption or heavy rain can cause attenuation on the order of 10 to 20 dB per Km. Small cells with small radii ranging from 50 to 200 meters are envisaged for cellular structures [4]. For such small cells, this attenuation would be only a few dB, which is less significant. In fact, this absorbtion and pathloss might be favorable owing to the fact that it reduces the interference from other BSs. On the other hand, mmWave signals are very sensitive to blockages since they show more specular propagations and low diffraction in comparison to microwave frequencies. As such, a link established by mmWave signals may be completely disconnected due to obstacles [5].

Hardware limitations including high cost and power consumption of analog/digital signal devices, expensive Radio Frequency (RF) chains, and difficulty of implementing a full RF chain make use of digital baseband processing infeasible for every antenna element. An RF chain for each antenna and digital baseband processing are used in traditional MIMO systems, but it is not power- and cost-efficient to use such an approach in mmWave sytems owing to the large number of antenna elements required at mmWave frequencies [6][7].

The key element addressing the propagation issues and providing high gains and improved signal-to-noise ratios (SNRs) in mmWave systems is employing directional beamforming generated by the aid of large antenna arrays. In fact, since mmWave systems employ small wavelength, a large number of antenna elements can be compacted in a small size in order to boost array gain and help overcome

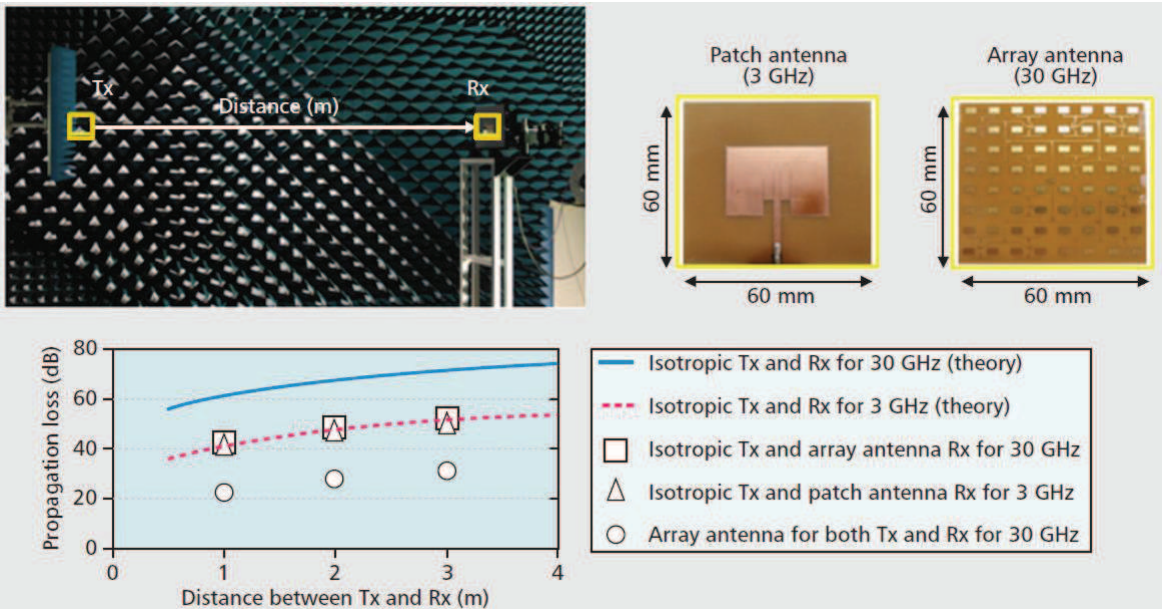


FIGURE 1.1: Measurements of propagation loss conducted in [9].

pathloss. In other words, mmWave signals do not bring about any increased pathloss, providing that physical antenna size is kept constant and directional transmissions are deployed [4][8][9]. To verify this, one can refer to measurements conducted in [9]. A patch antenna and an array antenna of the same physical aperture size which, respectively, operate in 3 and 30 GHz were used in these measurements as shown in Fig. 1.1. As seen, the patch antenna (3GHz) or the array antenna (30GHz) used at the receiver side both result in the same amount of propagation loss. Also, employing array antennas at the transmitter and receiver decreases the propagation loss and increases the receive power by 20 dB in the scenario considered in [9].

Analog beamforming is a sub-optimal solution to the hardware limitations in mmWave systems. A single RF chain and a group of phase shifters to adjust the phase of the signal transmitted by each antenna are employed in this beamforming technique. Sub-optimal performance of the analog beamforming is due to the fact that amplitudes of the weights applied to the signal by the antennas are constant and the resolution of phase adjustment is low. Also, the analog beamforming could not support transmission of multiple data streams owing to using a single RF chain [6][7].

Hybrid analog/digital structures have been employed to make multi-stream transmission possible. Two different parts constitute the hybrid structures. A large number of phase shifters improve the antenna array gain and a small number of RF chains are used in order to cancel multiuser interference; these are employed in analog and digital parts respectively [10].

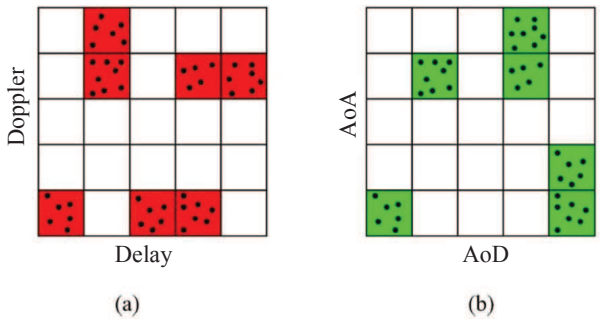


FIGURE 1.2: Resolution bins associated with (a) Doppler-Delay domain (b) Angle domain [14].

1.2 Channel Estimation in mmWave Systems

The physical model for the multipath wireless channel depends nonlinearly on a large number of parameters including angle of departure (AoD) from the transmitter, angle of arrival (AoA) to the receiver, Doppler shift and delay. It is a daunting task to estimate and analyze this nonlinear channel model. In addition, propagation paths in mmWave channels are formed by a small number of clusters, indicating the sparse structure of mmWave channels [11][12]. To approximate the channel model by a tractable and linear model and also to exploit the sparsity of mmWave channels, the virtual channel representation (the beamspace representation) is employed [13][14]. The virtual model is generated by sampling the space the four parameters generate and by virtue of four-dimensional (4D) Fourier series. The sampling provides 4D resolution bins associated with different discretized parameters as illustrated in Fig. 1.2 which divides the four dimensions to two 2D spaces: the Doppler-delay space and the angle space. In the virtual channel model, each 4D bin has a corresponding value known as a virtual channel coefficient. All paths lying within a bin are determined by merely one virtual channel coefficient. The dots in Fig. 1.2 show the paths within the bins. In order to gain more insight, Fig. 1.3 can be considered, showing the angle space for two bins in the Doppler-delay space.

Numerous channel estimation studies assumed an environment with many uniformly-distributed line of sight and diffuse paths, i.e., rich multipath environment. This means that nearly all bins in the virtual channel representation contain paths. By way of illustration, this indicates that nearly all bins in Fig. 1.2 or Fig. 1.3 should contain some dots. There are also many scenarios for which it has been found that propagation paths tend to form clusters around scatterers[14–17].

For channels including clusters of paths, when the bandwidth, symbol duration or the number of antennas employed in the system are large, the angle-delay-Doppler sampling rate is large as well, so the 4D bins become smaller than the space between multipath clusters [14]. As a consequence, the

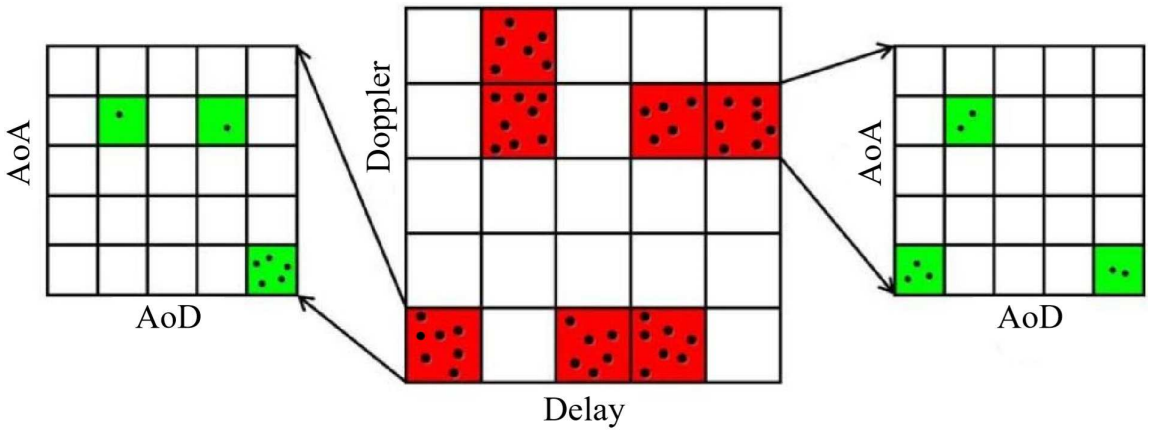


FIGURE 1.3: Angle space for two resolution bins in the Doppler-delay space [14].

number of bins containing the paths belonging to the clusters noticeably decreases and the number of empty bins significantly increases. Fig. 1.2 and Fig. 1.3 illustrate the case that the bandwidth or the number of antennas are very large. In other words, out of a large number of virtual channel coefficients just a few of them have values above an acceptable threshold, showing that the beamspace representation of a clustered multipath channel is sparse. Since only a few virtual channel coefficients have significant values, estimating and analyzing the channel narrows down to finding those few virtual channel coefficients. Otherwise, owing to very large dimensions of MIMO channel matrices, a plethora of channel gains would be estimated, and it would be a daunting task to do so [18]. In particular, a mmWave channel is a multipath channel with a small number of clusters, and its virtual model, owing to large number of antennas, includes only a few significant virtual channel coefficients. Therefore, the virtual model we have described captures the sparse nature of mmWave channels.

The nonzero virtual channel coefficients correspond to the bins in the 4D space (AoA-AoD-Doppler-delay), and the pair of the green and red bins linked by arrows in Fig. 1.3 correspond to these 4D bins. The sparsity of the channel is illustrated by both the sparsity in the Doppler-delay projection (sparse red bins) and the sparsity in the AoA-AoD projection (sparse green bins) [14].

The sparsity of the model representing the mmWave channel arises because there are only a few significant paths between the BS and user equipments (UEs) in cellular communications operating in mmWave frequencies. Thus, BSs and UEs need to be aware of the significant paths so as to align their beams in the direction of those paths. The channel estimation for mmWave systems is finding the best beam directions resulting in a desirable SNR for the connection, which is also known as beam alignment (BA). There have been efforts dealing with the channel estimation for mmWave systems in the literature. In the sequel, we will elaborate on related studies.

1.2.1 Existing Methods for Channel Estimation in mmWave Systems

A mmWave cellular system with multi-antenna BSs and multi-antenna mobile stations (MSs) is considered in [6]. Also, hybrid beamforming is incorporated, so the BSs and MSs have multiple RF chains. A narrowband block-fading model represents the channel between BS and MS in this network. This channel is, then, expressed in terms of propagation paths, antenna array response vectors at BS and MS, the complex gains of the paths and the azimuth angles of departure and arrival. Note that along with considering uniform linear arrays (ULAs), only the azimuth is considered so two-dimensional beamforming is applied in the horizontal plane.

It is assumed that BS sends equal symbols in successive time slots and the MS also makes measurements in successive time slots in [6]. Due to the sparsity of the mmWave channel, a sparse formulation for channel estimation is introduced, after quantizing AoA/AoDs. Compressive sensing (CS)-based methods can be employed for such sparse estimations where the number of nonzero elements of an unknown vector is known to be small relative to the length of the vector (see chapter 2). An adaptive CS technique is used in [6] due to the superior performance in low-SNR scenarios which occur in mmWave systems.

A codebook based on a hybrid analog/digital approach is presented in [6] for constructing the beamforming vectors applied by BS and MS. The designed codebook is called a hierarchical codebook and contains several levels or stages. Each of the levels has beamforming vectors which generate certain beamwidths covering specific ranges of angles. These angles are further divided into certain subranges of angles corresponding to other beamforming vectors in the next level. In each level, compared to the previous level, a search over a smaller angular space is carried out. The beamforming vectors of each level are used in updating the channel estimation algorithm. In other words, transceivers first use wider beams and then, based on the feedback exchanged between them, they refine the beams to finally find the best beam pair with the desired resolution. It must be pointed out that the codebook proposed in [6] with its very low complexity has superior performance compared to analog-only beamforming codebooks due to digital as well as analog processing. It also can be used for ULAs and non-ULAs.

Using the proposed hierarchical codebook, channel estimation is presented for two different cases: First, for the channel with merely one path, for which an upper bound of the probability of error for the channel estimation is proposed, and then the method is extended for the case where there are multiple paths.

Let K be the number by which each angle range is divided in the next level of the codebook design in [6]. Providing the desired angle resolution for channel estimation is $\frac{2\pi}{N}$, estimation time for

each path is proportional to $K^2 \log_K N$ [18]. This shows that [6] achieved substantial improvement in comparison to exhaustive search-based methods for BA problems. However, in order to have a robust channel estimation for fast channel variations, the speed of this channel estimation should be increased. In addition, the time delay caused by performing many measurements in high SNR scenarios can be decreased as done in [18]. Note that in the exhaustive search which is a simple approach for BA, transceivers use all possible combinations of generated beams to probe the whole angular space in order to find the best beam pair. This method is particularly favorable since narrow beams can be used at the transceivers to obtain high SNRs [19].

A fast channel estimation based on overlapped beam patterns is proposed in [18]. When beams are overlapped, they have intersections in certain angular ranges in the angular space or they cover a certain angular range simultaneously. Using such beams leads to estimating channels in less time compared to non-overlapped beam designs, as shown in [18].

Two channel estimation methods are introduced in [18], considering the sparse nature of mmWave wireless systems and using overlapped beams in the angular domain. The basic assumptions including the channel model and the structure of arrays in [18] are the same as that of [6].

To develop the first channel estimation method proposed in [18], first the vector of measurements at the receiver are taken, assuming that the same pilot signal is sent for M time slots by the BS and the UE makes measurements in a sequence of M time slots. Consequently, using the maximum likelihood technique, AoA/AoDs are extracted based on the measurements. This method works in a multi-level manner similar to that of [6]. The fading coefficients corresponding to AoA/AoDs of the best path are estimated from all the measurements of all hierarchical levels so as to get more accurate values whereas in [6], these coefficients are based entirely on the measurement of the final level.

Furthermore, a scheme to adaptively estimate the channel, by taking into account the maximum value of probability of estimation error, is discussed in [18]. A constraint is checked, based on this algorithm, and if it is not fulfilled, additional measurements will be requested. Satisfying the constraint is examined after completing measurements in each level dividing the angular space in sub-ranges. For this technique, a lower bound of the minimum signal-to-energy ratio corresponding to a definite number of channel measurements is also studied, and it is shown that, for a given probability of estimation error, utilizing this scheme results in a 6 dB gain for the energy-to-noise ratio in comparison to the algorithm in [6]. It is worth noting that apart from [18], the temporal correlations between two channel realizations can be used to improve the BA efficiency, which is called beam tracking or channel tracking in the literature [20][21].

From the analytical point of view, a direct comparison between the hierarchical channel estimation and the exhaustive search is not provided in [18], although some bounds related to the performance of the proposed approach are derived. Analytical assessment is considered for the two approaches in [22] to gain more insight, by employing analog beamforming and assuming the same number of pilot symbols.

Probability of misalignment is used as a performance measure in [22]. A misalignment happens whenever transceivers cannot estimate the beam pair contributing to the highest channel gain.

Upper and lower bounds of probability of misalignment for hierarchical search are derived in [22] under the assumption of a single path. It is shown that the upper and lower bounds closely follow each other when the training sequence is large. Moreover, the misalignment probability for hierarchical and exhaustive search is studied asymptotically, and the superior performance of the exhaustive search as compared to hierarchical search in case of large training sequences is shown. It is also interesting that this result holds for low SNR non-asymptotic cases.

In the next step, it is demonstrated that in low SNR scenarios the exhaustive search has lower misalignment probability than the hierarchical search for multi-path channels and imperfect beamforming patterns. Nonetheless, both techniques show the same performance for high SNR cases.

In [22], the study is limited to a single user scenario where the BS communicates with only one user. There is potential to extend misalignment characterization of hierarchical design and exhaustive search to the case where there exist multiple users in the mmWave system. Also, transceivers in mmWave systems are capable of generating multiple beams since they use directional beamforming. In [22], however, the study is focused on single beam transmission. The misalignment probability of hierarchical design or exhaustive search could be investigated for the scenario where one BS and one user both have the capability of simultaneously generating multiple beams try to find the best beam pair.

Beamspace multi-user MIMO (MU-MIMO) [23] or high-dimensional MU-MIMO [24] can be employed to make use of the full potential of mmWave systems. In this scenario, a BS generates multiple orthogonal beams and transmits them to multiple users simultaneously, and each user can generate one or more beams. The misalignment probability of different BA designs could also be explored for this case. Moreover, the performance of hierarchical designs and exhaustive search in channel estimation of beamspace MU-MIMO is not clear.

Note that in the above-mentioned studies, i.e. [6] and [18], a set of phase shifters are employed for the hybrid structures. It would be interesting to compare the effects of phase shifters and antenna

switches in hybrid beamforming structures for channel estimation of mmWave systems. Generally, using switches at UEs leads to lower power consumption and lower complexity. However, switch-based approaches show inferior array gain due to the relatively small number of active antennas equal to the number of RF chains [25].

A channel estimation for mmWave systems is investigated in [26] using switches in lieu of phase-shifters in the transceivers' structures. Also, the sensitivity of a CS-based method exploited in [25] to the array phase mismatch is studied in [26].

In practice, there are often some uncertainties about the array response including gain and phase mismatch. Performance of CS-based methods is dependent upon the basis by which the sparse representation is obtained. Using a priori known basis can cause performance degradation when the mismatches are considered in the system model [27]. In fact, finding the basis of the sparse channel for CS-based methods can be a very difficult task when there are array uncertainties; therefore, CS-based methods might not be effective in such cases.

The channel estimation technique proposed in [26] is not sensitive to mismatches. The presented approach also has lower complexity in comparison to the CS-based approach presented in [25]. Note that the proposed approach in [26] can be used for other array structures such as planar arrays since it is a basis-free design.

In addition, the strong incoherence property for mmWave channels is discussed in [26]. If a matrix has the strong incoherence property, then with high probability it can be reconstructed with a known number of uniformly sampled entries and the reconstructed matrix is error-free [28]. In [26], it is shown that strong coherence property holds for mmWave channels without phase mismatches.

Returning to the exhaustive and hierarchical searches, we now consider multi-user scenarios. The main disadvantage of the exhaustive search is that it brings about a large training overhead because of the large number of measurements. The hierarchical search can be employed to reduce the total number of measurements by performing the search in a closed-loop fashion. However, in multiuser scenarios, the hierarchical search might not be an efficient approach since it needs to be carried out for every single user, and the training overhead depends linearly upon the number of users. Employing the CS tool is a new approach to BA in order to exploit the sparse nature of the mmWave channels. Non-adaptive CS approaches are specifically useful for multiuser scenarios because each UE can estimate its own channel separately, which means that growing the number of users leads to no extra training overhead.

Analyzing the channel estimation problem for a multi-user mmWave system is crucial owing to

the fact that 5G aims to significantly increase the number of connected users to the network. In the sequel, some approaches which focus more on multiuser scenarios will be discussed.

Generally, BA methods whose training overhead scales sublinearly with the number of users are of interest. Considering this fact, all the users are simultaneously trained in [29]. A multi-antenna BS equipped with some RF chains communicates with a generic multi-antenna mobile station with one RF chain in [29].

The received matrix at the MS after the training procedure in [29] is formulated in terms of training signals. Then, by quantizing the AoA/AoDs, using the idea of virtual channel representation [14][13], the sparse formulation is expressed, where the sparse vector contains path gains associated with the quantized beam directions. The elements of the sensing matrix are generated based on angles randomly selected from quantized angle sets, and CS is leveraged to recover the sparse vector. The orthogonal matching pursuit algorithm discussed in subsection 2.2.2 is employed in [29] due to its low complexity.

Non-uniformly distributed angle sets are considered in [30], and it is proved that the rows of the array response matrices resulting from these non-uniform quantized angles are orthogonal. This orthogonality holds when the number of quantized angles is higher than or equal to the number of antennas and helps in designing pilot beam patterns and reducing the mutual coherence of the equivalent sensing matrix in the CS formulation. The mutual coherence of a sensing matrix is given by the maximum value of all the inner products between any two distinct normalized columns, as explained in section 2.1.

Pilot beam patterns are designed based on minimizing the total coherence of the equivalent sensing matrix in [30]. Note that the total coherence, as defined in [30], is the sum of the squared inner products of any two distinct columns of a sensing matrix. This design includes optimizing the baseband and RF processors jointly, and since it is hard to do so, just the baseband processors are designed and unitary matrices are employed as the RF processors. Moreover, the random permutation of the columns of the matrices used as the RF processors is suggested to improve the performance of the proposed pilot beam patterns.

Some other CS-based BA methods which are suitable and efficient for multiuser scenarios have been proposed in [31–34]. These methods rely on the assumption that the instantaneous channel stays constant during the BA process . In practice, this assumption might not be generally correct since mmWave systems have large Doppler spread resulting in rapidly changing channels, albeit in dense mmWave networks line of sight paths which are dominant tend to have only very slowly varying path

gains [29][35].

Second order statistics of the channel is used in [36] in order to propose a robust method against the significant variation of the channel. A multiuser scenario where ULAs are deployed at the BS and users is considered. It is assumed that AoA/AoDs do not significantly vary over the BA period since the angle coherence time is very much longer than the channel coherence time. The angle coherence time is the time during which AoA/AoDs of scatterers vary significantly. In order to exploit the sparsity of the channel, the frequency domain channel matrix is quantized with respect to AoA/AoDs after presenting the channel and the details of the OFDM signalling employed.

In [36], the BS and UEs generate the same beam patterns for a comb of subcarriers with each RF chain. The beam patterns are generated based on the pseudo-random BF vectors at the BS and random BF vectors at the UEs. The BF vectors at the BS and the UEs are the linear combination of the columns of the discrete Fourier transform (DFT) basis after quantizing the angles. Also, the number of the narrow beams which are constant during the measurements depends on the SNR before BF.

The BS sends beacon slots each of which is one OFDM symbol and probes the channel using beamforming vectors selected from the pseudo-random beamforming codebook known by all the users, during the channel probing process. It is assumed that beacon OFDM symbols are orthogonal in frequency domain. Each user applies beamforming vectors from its own codebook in the sensing step. The users, then, estimate the strongest multipath component and send the information of the strongest multipath component back to the BS. Next, the BS sends an acknowledgement. If an error occurs during this process and the BS cannot send the correct acknowledgement back to the user, the user starts the BA process again after a specified time period has elapsed.

The channel estimation in [36] is carried out using the fact that the strong paths are associated with components of the channel matrix which have large second moment. Therefore, the optimization problem for the channel estimation is presented in terms of a sparse unknown vector based on the second moments of the channel matrix components. This convex optimization problem is a non-negative least squares (NNLS) problem since the unknown vector is non-negative. The unknown sparse vector in [36], owing to its non-negativity, can be recovered using numerical methods for NNLS such as the gradient projection without onerous computational complexity.

In conventional BA schemes often there are several rounds in which signals are exchanged between the BS and users to refine the beams for finding the best beam pair. This is problematic when the channel changes. The approach proposed in [36] circumvents this problem using second order statistics and the fact that the angles themselves only change slowly. Also, all the users are simultaneously

trained using this approach.

After making the measurements in CS-based BA methods, the structure of the resulting sensing matrix has a key role in reliably reconstructing the sparse vector. It has been shown that if the components of the sensing matrix are generated based on some distributions such as i.i.d random Gaussian or Bernoulli, the sparse vector can be recovered with high probability when the number of measurements is more than a lower bound [37][38]. In [29], the sensing matrix used for the sparse recovery is obtained based on random BF which is unstructured. The UEs must have access to the BF codebook deployed by the BS in order to estimate the channel using the CS methods. Therefore, the BF codebook of the BS must be transmitted to the UEs or stored in them, bringing about signalling or storage overhead. Also, the number of nonzero elements in the sensing matrix might be large when employing random BF codebooks, and this increases the computational burden at the UEs. In [36], the sensing matrix used for the sparse recovery is obtained based on pseudo-random BF codebooks at the BS and random codebooks at the UEs. So, the UEs can obtain the pseudo-random BS codebooks in a structured way. Even though this reduces the signalling overhead, the sensing matrix still has a random structure and it is not guaranteed that a specific realization of the random sensing matrix works.

A structured random CS method is incorporated in [39] in order to reduce the signalling or storage overhead resulting from unstructured random BF codebooks at the BS. In [39], a column of the DFT matrix is selected as a beamformer to generate a beam in the first stage and in the second stage the beam is spread over the entire angular range using a unimodular sequence. After employing the notion of virtual channel representation, the sparse formulation based on circulant matrices is given. The circulant matrices are constructed using the DFT basis and the unimodular sequence. Employing the circulant matrices lead to a circulant convolution between the virtual representation of the channel matrix and the circulant matrices, spreading the information of the virtual representation of the channel uniformly in the angle domain. This process which is based on a CS technique has comparable recovery performance to that of completely random measurements [40].

A beam discovery technique using the idea of linear block codes is proposed in [41] to reduce the number of measurements needed for channel estimation. The work in [41] assumes perfect sparsity where AoA/AoDs are always along the quantized angles; therefore, the number of nonzero values in the virtual model of the channel is exactly the same as the number of clusters in the mmWave channel. In channel coding, error syndromes are calculated using the parity check matrix. Then, error patterns via a look-up table are determined, and finally the codewords are obtained. Likewise, finding the

positions of clusters in the virtual channel model is similar to finding the error patterns of a codeword as stated in [41]. So, beam patterns for making the measurements are designed based on the parity check matrix of linear block codes.

There exist two major challenges in designing suitable training beam patterns for mmWave channels. Firstly, since the number of antennas tend to be large in mmWave channels, examining all possible combinations of beam pairs between the BS and UEs would be a daunting task, so reducing the number of measurements is of interest. Deploying CS methods is a promising solution because of the sparse structure of the mmWave channels.

Second, making the measurements in the training step should not impose difficulties from signal processing point of view. Generating BF vectors completely randomly works well with CS methods in terms of sparse recovery. However, some issues arises by employing random BF codebooks:

1. Owing to the unstructured nature of the random BF employed at the BS, all the random BF vectors must be signalled to or stored in the UEs. The signalling or storage overhead owing to this issue becomes more significant in cellular systems where different BSs use different random codebooks [39].

2. Random BF codebooks works well with high probability; however, it is not *guaranteed* that a specific realization of the sensing matrix resulting from random BF codebooks always works [42].

3. The number of nonzero components of the sensing matrix resulting from the random BF codebooks might be large, and this makes the measurement process slow since the number of calculations increases [43]. This constrains the latency achievable in the mmWave systems, which is obviously unfavorable in realizing 5G's objectives.

To address the above-mentioned issues, one potential solution can be using deterministic sensing matrices for CS-based BA methods. In deterministic CS structures, the number of nonzero elements of the sensing matrix can be reduced, which contributes to faster measurement process and lower amount of computational burden at the UEs' receiver. Also, the storage or signalling overhead in the mmWave systems can be considerably reduced due to the fact that the UEs can generate the deterministic sensing matrix by just knowing a few parameters.

In this dissertation, inspired by a deterministic CS method, we address the BA problem for mmWave channels. The main accomplishments of this dissertation are summarized as follows:

- We propose two deterministic sensing matrices. We construct the first deterministic sensing matrix in section 4.3 by performing a Kronecker product between two DeVore's sensing matrices.

The process of generating the DeVore's sensing matrix is explained in section 4.2. Also, we construct the second deterministic sensing matrix in section 4.4 by doing a row-by-row Kronecker product between rows of two matrices which include DeVore's sensing matrices and circular shifts of DeVore's sensing matrices.

- We design the pilot beam patterns (BF vectors) for the BA process based on the proposed deterministic sensing matrices.
- Our proposed approach results in significant reduction of number of measurements compared to the exhaustive search.
- We show that our proposed method contributes to significant overhead reduction compared to the random or structured random BF methods.
- We show that in some scenarios our proposed approach presents an improvement in terms of performance compared to other methods.

1.3 Thesis Organization

In chapter 2, we provide an overview to compressed sensing. The system model is introduced in chapter 3. The process of constructing proposed deterministic sensing matrices and designing the pilot beam patterns are presented in chapter 4. In chapter 5, simulation results are presented. Finally, chapter 6 concludes the thesis.

Chapter **2**

An Overview of Compressive Sensing

A signal is sparse when its vector representation in a basis has only a few nonzero elements. A signal may not be sparse in general but it might exhibits sparsity by applying some transformations. To be more accurate, the signal representation is sparse only in a certain basis. The topic of compressed sensing (CS) is about reconstructing sparse signals using a small number of measurements. Objectives of CS include efficient reconstruction of sparse signals by minimizing the number of measurements, perfect recovery in the noiseless case, and robustness results for scenarios with noise [44].

2.1 General Mathematical Model for CS

In order to represent the CS reconstruction formulation, we start with the following equation,

$$\mathbf{y} = \mathbf{\Phi x}, \tag{2.1}$$

where $\mathbf{y}_{m\times 1}$, $\mathbf{\Phi}_{m\times n}$ and $\mathbf{x}_{n\times 1}$ are the measurement vector, measurement matrix and input signal respectively. The number of measurements are much fewer than the input signal dimension, i.e, $m \ll n$; therefore, (2.1) is heavily undetermined.

Assume $\mathbf{\Psi}$ as the basis in which $\mathbf{x}$ can be represented as a sparse vector, using the linear combination of $\mathbf{\Psi}_i$ which are columns of $\mathbf{\Psi}$ the vector $\mathbf{x}$ can be written as,

$$\mathbf{x} = \sum_{i=1}^{n} s_i \mathbf{\Psi}_i = \mathbf{\Psi s}, \tag{2.2}$$

where $\mathbf{s}_{n\times 1}$ is the equivalent sparse representation of $\mathbf{x}$, i.e., $\mathbf{s}$ has only a few nonzero elements. Here, we assume $\mathbf{s}$ is an L-sparse vector having at most $L \ll n$ nonzero values (or its level of sparsity is L).

Referring to (2.1) and (2.2), the sparse problem can be written as follows,

$$\mathbf{y} = \mathbf{\Theta s}, \tag{2.3}$$

where the reconstruction matrix or sensing matrix $\mathbf{\Theta}_{m\times n}$ is defined as $\mathbf{\Theta} = \mathbf{\Phi} \times \mathbf{\Psi}$.

Properties of $\mathbf{\Theta}$ determine the possibility of perfect recovery. The restricted isometry property (RIP) is a sufficient condition for stable reconstruction. Stable reconstruction indicates that small changes in the measurements should cause small changes in the reconstruction [45][46].

The RIP is defined as follows: $\mathbf{\Theta}$ satisfies the RIP of order L if for all L-sparse vectors $\mathbf{u}$ and a constant $0 < \delta_L < 1$ which is not too close to one [47], the following condition holds:

$$1 - \delta_L \leq \frac{\|\mathbf{\Theta u}\|_2}{\|\mathbf{u}\|_2} \leq 1 + \delta_L. \tag{2.4}$$

Given $\mathbf{\Theta}$, L and δ_L, it is an arduous task to verify RIP [48]. An alternative condition which is easier to verify is the mutual incoherence property. To measure the mutual coherence of $\mathbf{\Theta}$ the following expression is used,

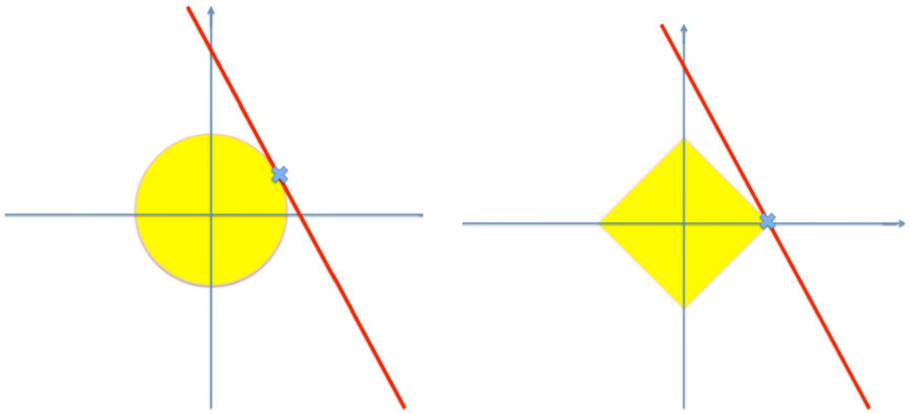
$$\mu(\mathbf{\Theta}) = \max_{i\neq j} \frac{|\langle \boldsymbol{\theta}_i, \boldsymbol{\theta}_j \rangle|}{\|\boldsymbol{\theta}_i\|_2 \|\boldsymbol{\theta}_j\|_2}, \tag{2.5}$$

where $\boldsymbol{\theta}_i$ are the columns of $\mathbf{\Theta}$ and $\langle . \rangle$ indicates the inner product between these columns.

The value of μ is bounded between the Welch bound $\sqrt{\frac{n-m}{m(n-1)}}$ and one, i.e., $\mu \in \left[\sqrt{\frac{n-m}{m(n-1)}}, 1\right]$, and a small value of μ is desirable [38][49]. Indeed, the mutual incoherence property (which indicates small values of μ) is a stronger condition than RIP and from the mutual incoherence property can be implied that RIP is satisfied but the converse does not hold [50]. In addition, the mutual coherence parameter and the RIP constant are related according to the following preposition.

Proposition 1 *[51]* $\mathbf{\Theta}$ *with unit-norm columns and the coherence parameter* μ_N *satisfies RIP of order* L *with constant* $\delta_L = (L-1)\mu_N$

The structure of the sensing matrix $\mathbf{\Theta}$ has a key role in successful recovery. It has been shown that random sensing matrices with elements constructed from Gaussian or Bernoulli distributions satisfy RIP of order L for $m \geq O(L \log(n/L))$ with high probability [52]. However, employing random sensing matrices owing to generating and storing their components might pose practical limitations from signal processing point of view. The alternative approach is using structured measurement matrices such as Toeplitz and circulant matrices or deterministic sensing matrices [38].



(a) l_2-norm (b) l_1-norm

FIGURE 2.1: The intersection of the line $\mathbf{\Theta s} = \mathbf{y}$ with the l_2 and l_1 balls [44].

Extracting the sparse vector $\mathbf{s}$ can be stated as an l_0-optimization problem:

$$\min_{\mathbf{s}} \|\mathbf{s}\|_0 \quad \text{subject to} \quad \mathbf{\Theta s} = \mathbf{y}, \tag{2.6}$$

where $\|\mathbf{s}\|_0$ indicates the non-convex l_0-norm counting the number of nonzero entries of $\mathbf{s}$. Problem (2.6) has been shown to be NP-hard in general since it searches over all possible combinations, which is a computationally onerous task [53]. One of the alternative methods is employing linear programming to find a similar solution in near polynomial time. In this method known as basis pursuit (BP), l_1-norm which has the most similarity to l_0-norm [44][54] is employed to approximate (2.6) as a convex optimization problem as follows:

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{subject to} \quad \mathbf{\Theta s} = \mathbf{y}, \tag{2.7}$$

where $\|\mathbf{s}\|_1$ is the l_1-norm of $\mathbf{s}$ and it adds up absolute values of all components of $\mathbf{s}$. It is worth noting that the estimate of $\mathbf{x}$ is obtained after the reconstruction of the estimate of $\mathbf{s}$, by the inverse transform used to represent $\mathbf{x}$ as a sparse vector $\mathbf{s}$.

In order to understand why l_1-norm can find the sparse solution better than l_2-norm, we can consider Fig. 2.1. The feasible set is the line $\mathbf{\Theta s} = \mathbf{y}$ in this example illustrating l_1-norm and l_2-norm in two dimensions. Using l_2-norm in lieu of l_1-norm in (2.7), the solution is a point on the line $\mathbf{\Theta s} = \mathbf{y}$ with the minimum Euclidean distance from the origin. As seen in Fig. 2.1(a), this point locates away from the coordinate axes; therefore, it does not give a sparse solution. However, for the l_1-norm case shown in Fig. 2.1(b), the intersection of the l_1-norm ball and the line $\mathbf{\Theta s} = \mathbf{y}$ lies on the coordinate axis, i.e., l_1-norm leads to a sparse solution [44].

It is important to mention that the fundamental results of [53] and [55] show that satisfying RIP makes l_0-norm and l_1-norm problems equivalent with high probability. Based on this fact, Gaussian and Bernoulli random matrices which satisfy RIP lead to the same sparse solution for l_0-norm and l_1-norm problems with high probability.

Different CS reconstruction approaches have been proposed thus far. Convex, greedy, thresholding, combinatorial, non-convex and Bayesian are the classifications for these approaches. Measurements in practical scenarios come with perturbation resulting from noise. This makes it necessary for practical CS approaches to be robust against noisy measurements. Convex and greedy approaches among the aforementioned ones are robust to noise [38]. To be exact, in sparse reconstruction when measurements are contaminated with noise, the basis pursuit, which is one of the convex approaches, and greedy approaches can provide good results if the sensing matrix satisfies RIP of order $2L$ with an RIP constant δ_{2L} which is not too close to one [42]. In the next section, some of the convex or greedy approaches will be discussed.

2.2 CS Approaches for Noisy Recovery

When measurements are perturbed with noise, the sparse problem in (2.3) can be rewritten as follows:

$$\mathbf{y} = \mathbf{\Theta s} + \mathbf{n}, \tag{2.8}$$

where $\mathbf{n}$ is the noise.

Basis pursuit which is discussed in the previous section, basis pursuit denoising (BPDN), Dantzig selector and least absolute shrinkage selection operator (LASSO) are convex approaches which all have been used to solve the problem of noisy recovery. Also, matching pursuit (MP), orthogonal matching pursuit (OMP) and their variants such as compressive sampling matching pursuits (CoSaMP) or regularized OMP are greedy algorithms which have been used for CS noisy recovery. In the following, these approaches are concisely reviewed. Note that we will only review MP and OMP for greedy algorithms and we refer the reader to [38] and [56] for more details about the variants of MP and OMP.

2.2.1 Convex Approaches

By relaxing the equality constraint in (2.7) in order to consider noise effect, the BPDN formulation can be presented as follows

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{subject to} \quad \frac{1}{2}\|\mathbf{y} - \mathbf{\Theta s}\|_2^2 \leq \epsilon, \tag{2.9}$$

where $\|.\|_2$ indicates the Euclidean norm or l_2-norm and the constraint limits the squared error between $\mathbf{y}$ and $\mathbf{\Theta s}$ to be less than or equal to ϵ. BPDN can also be cast in a Lagrangian form as follows

$$\min_{\mathbf{s}} \quad \lambda\|\mathbf{s}\|_1 + \frac{1}{2}\|\mathbf{y} - \mathbf{\Theta s}\|_2^2, \tag{2.10}$$

where the value of λ controls the trade-off between reconstruction accuracy and sparsity of the solution. Algorithms such as fixed-point continuation and primal dual interior-point method are employed to solve (2.10). By swapping the constraint with the objective function in (2.9), the LASSO formulation is obtained and is written as follows

$$\min_{\mathbf{s}} \ \frac{1}{2}\|\mathbf{y} - \mathbf{\Theta s}\|_2^2 \quad \text{subject to} \quad \|\mathbf{s}\|_1 \leq \epsilon. \tag{2.11}$$

Dantzig selector formulation is also similar to (2.9). The only difference is that the squared error between $\mathbf{y}$ and $\mathbf{\Theta s}$ is measured by l_∞-norm as follows

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{subject to} \quad \frac{1}{2}\|\mathbf{y} - \mathbf{\Theta s}\|_\infty^2 \leq \epsilon, \tag{2.12}$$

where $\|\mathbf{a}\|_\infty$ selects the element with the maximum absolute value from $\mathbf{a}$.

Solvers for convex problems can be employed to solve the above-mentioned problems. We refer the reader to [38] for more details regarding the different solvers.

2.2.2 Greedy Approaches

Greedy approaches deal with CS problems in an iterative fashion. Iterative steps are carried out based on the correlation of the measurements and columns of the sensing matrix. The columns resulting in high correlation, called atoms, are picked to update the solution. High speed, simple operation and low computational complexity are the advantages of greedy approaches; however, they need to have information about the sparsity of the result a priori.

In MP, $\mathbf{r} \in \mathbb{R}^m$ is defined as a residual vector. The column having maximum correlation with $\mathbf{r}$ is found at each iteration. Mathematically speaking, in each iteration the index of the atom, i.e., λ_{it}, is determined as follows

$$\lambda_{it} = arg \max_k |\langle \mathbf{r}_{it}, \boldsymbol{\theta}_k \rangle|, \tag{2.13}$$

where $\boldsymbol{\theta}_k$ is the k-th column of $\boldsymbol{\Theta}$, it indicates the iteration counter and $\langle . \rangle$ denotes the inner product. Note that $\mathbf{r}_1$ which is the residual in the first iteration is assumed equal to the measurement vector $\mathbf{y}$. In iteration it, $s_{\lambda_{it}}$ which is the λ_{it}-th index of the solution is updated with the scalar projection of $\mathbf{r}_{it}$ along $\boldsymbol{\theta}_{\lambda_{it}}$ as follows

$$s_{\lambda_{it}} = \frac{\langle \mathbf{r}_{it}, \boldsymbol{\theta}_{\lambda_{it}} \rangle}{\|\boldsymbol{\theta}_{\lambda_{it}}\|_2}, \tag{2.14}$$

and the new residual is obtained by subtracting the vector projection from the previous residual as follows

$$\mathbf{r}_{it+1} = \mathbf{r}_{it} - \frac{\langle \mathbf{r}_{it}, \boldsymbol{\theta}_{\lambda_{it}} \rangle \boldsymbol{\theta}_{\lambda_{it}}}{\|\boldsymbol{\theta}_{\lambda_{it}}\|_2^2}. \tag{2.15}$$

The iterations are repeated until the norm of $\mathbf{r}_{it}$ meets the desired threshold or the number of iterations becomes equal to the sparsity level [38][49]. Note that the complexity of MP is $O(mnT)$, assuming T the number of iterations.

OMP is a variant of MP, with the complexity $O(mnL)$ where L is the number of nonzero components of $\mathbf{s}$. In this algorithm, the indexes of atoms are stored in a set $\boldsymbol{\Omega}$, and the atoms form the matrix $\boldsymbol{\Theta}_{\boldsymbol{\Omega}}$. Based on $\boldsymbol{\Theta}_{\boldsymbol{\Omega}}$ and employing the least square method, all coefficients of the solution are updated simultaneously in each iteration. The index of the atom is found as follows

$$\lambda_{it} = arg \max_{k \notin \boldsymbol{\Omega}_{it-1}} |\langle \mathbf{r}_{it}, \boldsymbol{\theta}_k \rangle|, \tag{2.16}$$

then the set of indexes is updated as $\boldsymbol{\Omega}_{it} = \boldsymbol{\Omega}_{it-1} \cup \{\lambda_{it}\}$, and finally, the solution and the residual are obtained as given in (2.17) and (2.18) respectively.

$$\hat{\mathbf{s}}_{\boldsymbol{\Omega}_{it}} = arg \min_{\mathbf{s}} \|\boldsymbol{\Theta}_{\boldsymbol{\Omega}_{it}} \mathbf{s} - \mathbf{y}\|_2^2, \tag{2.17}$$

$$\mathbf{r}_{it+1} = \mathbf{y} - \boldsymbol{\Theta}_{\boldsymbol{\Omega}_{it}} \hat{\mathbf{s}}_{\boldsymbol{\Omega}_{it}}. \tag{2.18}$$

Chapter 3

System Model

We assume a mmWave wireless system comprising a BS with N_T antennas and a generic UE with N_R antennas. The BS and the UE both have one RF chain and they are equipped with uniform linear arrays (ULAs). The space between antenna elements in the arrays is $d = \frac{\lambda}{2}$, where λ is the wavelength and it is calculated by $\lambda = \frac{c_0}{f_0}$, where c_0 and f_0 are the speed of light and the carrier frequency respectively.

In addition, in this thesis, we assume that phase shifting as well as the amplitude control can be performed in the analog domain. This is a practically feasible assumption for mmWave systems as it has been shown in the literature [57][58]. In the next two sections, we describe the channel model and the process of channel estimation.

3.1 Channel Model

Assuming $\theta_l \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ and $\phi_l \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ respectively the AoA and AoD of the lth propagation path between the BS and UE, the array response vectors are given by

$$\mathbf{a}(\theta_l) = \frac{1}{\sqrt{N_T}}[1, e^{j\pi \sin(\theta_l)}, ..., e^{j(N_T-1)\pi \sin(\theta_l)}]^T, \tag{3.1}$$

and

$$\mathbf{b}(\phi_l) = \frac{1}{\sqrt{N_R}}[1, e^{j\pi \sin(\phi_l)}, ..., e^{j(N_R-1)\pi \sin(\phi_l)}]^T. \tag{3.2}$$

We also assume that AoAs and AoDs of the propagation paths have uniform distribution within the angular range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Since a small number of clusters contributes to the propagation paths in the mmWave channels [11][13], we use the clustered physical channel model as follows:

$$\mathbf{H} = \sqrt{\frac{N_T N_R}{L}} \sum_{l=1}^{L} \alpha_l \mathbf{b}(\phi_l)\, \mathbf{a}^H(\theta_l), \tag{3.3}$$

where we assume that the physical channel includes L clusters of scatterers each of which creates a propagation path and $L \ll \max\{N_T, N_R\}$ [36][59]. Also, $\alpha_l \sim \mathcal{CN}(0, \sigma_{\alpha_l}^2)$ is the complex channel gain of the lth propagation path. In addition, we assume that all path gains, i.e., α_l, are constant during the beam alignment (BA) procedure. This is relevant to dense mmWave networks [35].

In (3.3), the AoAs and AoDs have continuous values. To have a tractable channel model, we approximate the channel model in (3.3) with a discrete representation using the idea of a virtual channel model (or beamspace representation) introduced in section 1.2 [13][14]. To do so, we use the following expressions to quantize the AoAs and AoDs [36]:

$$\sin(\theta_{c_1}^q) = \frac{2(c_1 - 1)}{N_T} - 1;\ c_1 = 1, 2, ..., N_T, \tag{3.4}$$

$$\sin(\phi_{c_2}^q) = \frac{2(c_2 - 1)}{N_R} - 1;\ c_2 = 1, 2, ..., N_R, \tag{3.5}$$

where $\theta_{c_1}^q$ and $\phi_{c_2}^q$ indicate the quantized angles. Since ULAs are employed at the BS and UE, the array response vectors corresponding to all $\theta_{c_1}^q$ and all $\phi_{c_2}^q$ form orthonormal bases or unitary DFT matrices [13][29]. In other words, the array response vectors $\mathbf{a}(\theta_{c_1}^q)$ and $\mathbf{b}(\phi_{c_2}^q)$ form respectively the DFT matrices $\mathbf{F}_{N_T}$ and $\mathbf{F}_{N_R}$ as follows:

$$\mathbf{F}_{N_T} = [\mathbf{a}(\theta_1^q), \mathbf{a}(\theta_2^q), ..., \mathbf{a}(\theta_{N_T}^q)], \tag{3.6}$$

and

$$\mathbf{F}_{N_R} = [\mathbf{b}(\phi_1^q), \mathbf{b}(\phi_2^q), ..., \mathbf{b}(\phi_{N_R}^q)], \tag{3.7}$$

which can be obtained by

$$\mathbf{F}_{N_{T(i_1,i_2)}} = (N_T)^{-\frac{1}{2}} e^{2\pi j (i_1 - 1)(\frac{i_2 - 1}{N_T} - \frac{1}{2})}, \tag{3.8}$$

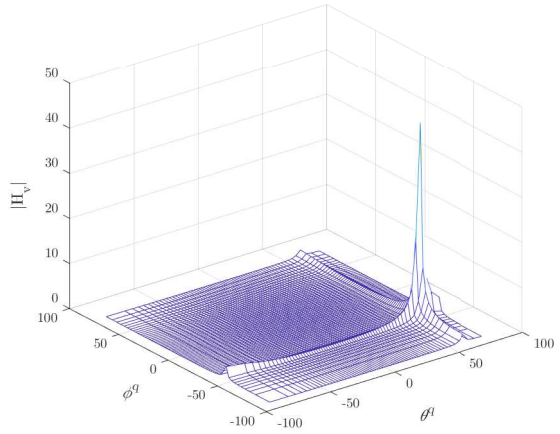


FIGURE 3.1: The virtual channel representation.

and

$$\mathbf{F}_{N_{R(i_3,i_4)}} = (N_R)^{-\frac{1}{2}} e^{2\pi j(i_3-1)(\frac{i_4-1}{N_R}-\frac{1}{2})}, \tag{3.9}$$

where $i_1, i_2 = 1, 2, ..., N_T$ and $i_3, i_4 = 1, 2, ..., N_R$.

Now, using the DFT matrices, we can represent the channel model by

$$\mathbf{H} = \mathbf{F}_{N_R} \mathbf{H}_v \mathbf{F}_{N_T}^H, \tag{3.10}$$

where $\mathbf{H}_v$ is the virtual channel representation which is a sparse matrix with L components having significant nonzero values corresponding to the AoAs and AoDs of the propagation paths. Note that due to the quantization error, there are other components with relatively insignificant values in $\mathbf{H}_v$. If the AoAs and AoDs are exactly aligned with the quantized angles in (3.4) and (3.5), $\mathbf{H}_v$ has exactly L components with nonzero values.

Fig. 3.1 illustrates the magnitude of $\mathbf{H}_v$ for a channel realization. In the example in Fig. 3.1, it is assumed that the BS and UE both have 64 antennas, and there is only one strong propagation path between them with $\theta_l = 60°$, $\phi_l = -45°$ and $\alpha_l = 1$. As seen, $\mathbf{H}_v$ has only one component with a nonzero significant value, and this component locates on the grid point whose corresponding quantized angles are the closest angles to the true angles. Also, components with very small values can be seen, resulting from the quantization error.

3.2 Channel estimation

In the training process, the BS transmits pilot signals x_t using unit-norm transmit BF vectors $\mathbf{w}_t \in \mathbb{C}^{N_T \times 1}$, where t denotes the t-th measurement. Then, the UE applies its unit-norm receive BF vectors $\mathbf{g}_t \in \mathbb{C}^{N_R \times 1}$ to make the t-th measurement which is given by

$$y_t = \mathbf{g}_t^H \mathbf{H} \mathbf{w}_t x_t + \mathbf{g}_t^H \mathbf{n}_t, \tag{3.11}$$

where $\mathbf{n}_t \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$ is the noise vector. Without loss of generality, we assume that $x_t = \sqrt{P}$, where P is the average received power of the pilot signals.

By applying the vectorization identity $vec(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) vec(\mathbf{B})$ to both sides of (3.11) where $\otimes$ indicates the Kronecker product and defining $\mathbf{h} = vec(\mathbf{H})$, we can write

$$y_t = \sqrt{P}(\mathbf{w}_t^T \otimes \mathbf{g}_t^H)\mathbf{h} + \mathbf{g}_t^H \mathbf{n}_t. \tag{3.12}$$

Referring to (3.10), we can obtain $\mathbf{h}$ in terms of the virtual channel representation as follows:

$$\mathbf{h} = vec(\mathbf{H}) = (\mathbf{F}_{N_T}^* \otimes \mathbf{F}_{N_R})\mathbf{h}_v, \tag{3.13}$$

where $\mathbf{h}_v = vec(\mathbf{H}_v)$.

In this thesis, we use the linear combinations of the columns of the DFT matrices to design the beam patterns as described in [36]. The vectors $\mathbf{w}_t^{(1)} \in \{0,1\}^{N_T \times 1}$ and $\mathbf{g}_t^{(1)} \in \{0,1\}^{N_R \times 1}$ select the columns of the DFT matrices $\mathbf{F}_{N_T}$ and $\mathbf{F}_{N_R}$ respectively. Therefore, the transmit and receive BF vectors can respectively be expressed as

$$\mathbf{w}_t = \mathbf{F}_{N_T} \frac{\mathbf{w}_t^{(1)}}{\sqrt{z_1}}, \tag{3.14}$$

and

$$\mathbf{g}_t = \mathbf{F}_{N_R} \frac{\mathbf{g}_t^{(1)}}{\sqrt{z_2}}, \tag{3.15}$$

where z_1 and z_2 indicate the number of ones in $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$. Note that each component of the vectors $\mathbf{w}_t^{(1)}$ or $\mathbf{g}_t^{(1)}$ is related to one quantized angle. If a component of these vectors equals one, it indicates that the antenna elements generate a narrow beam aligned with the corresponding quantized angle to that component. In fact, the ones in the vectors $\mathbf{w}_t^{(1)}$ or $\mathbf{g}_t^{(1)}$ can be thought of as switching on the corresponding narrow beams and the zeros are for switching off the corresponding narrow beams.

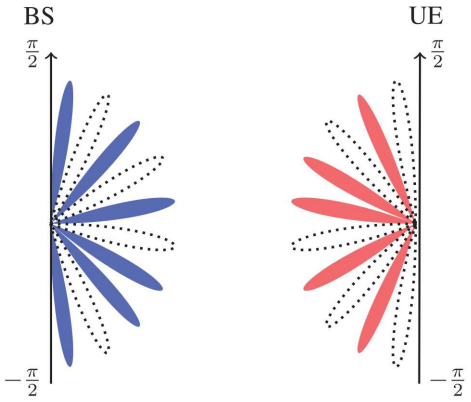


FIGURE 3.2: Pilot beam patterns [36].

Multiple ones in the vectors $\mathbf{w}_t^{(1)}$ or $\mathbf{g}_t^{(1)}$ result in multiple narrow beams. To illustrate this, Fig. 3.2 which is presented in [36] can be considered. In this example, the BS and UE both have 10 antennas and are measuring the channel using the vector $\mathbf{w}_t^{(1)} = [1,0,1,0,1,0,1,1,0,1]^T$ (at the BS) and the vector $\mathbf{g}_t^{(1)} = [0,1,0,1,1,0,1,0,1,0]^T$ (at the UE). As seen, the antenna elements at the BS or UE are capable of generating 10 narrow beams along with quantized angles. Out of all the 10 directions only those narrow beams corresponding to the components of $\mathbf{w}_t^{(1)}$ or $\mathbf{g}_t^{(1)}$ with value one probe the channel and make the measurement.

Using (3.12) to (3.15), we can write

$$y_t = \sqrt{\frac{P}{z_1 z_2}}(\mathbf{w}_t^{(1)^T}\mathbf{F}_{N_T}^T \otimes \mathbf{g}_t^{(1)^H}\mathbf{F}_{N_R}^H)(\mathbf{F}_{N_T}^* \otimes \mathbf{F}_{N_R})\mathbf{h}_v + \frac{1}{\sqrt{z_2}}\mathbf{g}_t^{(1)^H}\mathbf{F}_{N_R}^H\mathbf{n}_t. \tag{3.16}$$

Also, using the mixed-product property of the Kronecker product, i.e., the identity $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD}$ and assuming $P = 1$, (3.16) can be rewritten as

$$y_t = \frac{1}{\sqrt{z_1 z_2}}(\mathbf{w}_t^{(1)^T} \otimes \mathbf{g}_t^{(1)^T})\mathbf{h}_v + \tilde{n}_t, \tag{3.17}$$

where $\tilde{n}_t \sim \mathcal{CN}(0, \sigma_n^2)$, and as $\mathbf{g}_t^{(1)}$ is a real-valued vector (it contains only ones and zeros), $\mathbf{g}_t^{(1)^H}$ has been substituted with $\mathbf{g}_t^{(1)^T}$.

The measurement vector $\mathbf{y}$ can be formed at the UE by stacking all the measurements and it can be written as follows:

$$\mathbf{y} = \mathbf{S}\mathbf{h}_v + \tilde{\mathbf{n}}, \tag{3.18}$$

where the sensing matrix $\mathbf{S}$ is given by

$$\mathbf{S} = \frac{1}{\sqrt{z_1 z_2}} \begin{pmatrix} \mathbf{w}_1^{(1)^T} \otimes \mathbf{g}_1^{(1)^T} \\ \mathbf{w}_2^{(1)^T} \otimes \mathbf{g}_2^{(1)^T} \\ \vdots \\ \mathbf{w}_k^{(1)^T} \otimes \mathbf{g}_k^{(1)^T} \end{pmatrix}. \tag{3.19}$$

Note that k denotes the number of measurements and $\tilde{\mathbf{n}}$ in (3.18) is the noise vector.

Since $\mathbf{h}_v$ has a sparse structure, a proper sensing matrix $\mathbf{S}$ (as discussed in chapter 2) should be employed for a good reconstruction of $\mathbf{h}_v$. As seen, the structure of $\mathbf{S}$ depends on the vectors $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$ which make the transmit and receive BF vectors. In fact, the vectors $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$ show how the measurements are made in the angular domain. In the next chapter, we design these vectors based on deterministic sensing matrices.

Chapter 4

Pilot Beam Pattern Design

In this chapter, we propose two deterministic sensing matrices, and based on the proposed deterministic sensing matrices, we design the structure of pilot beam patterns for the beam alignment (BA) process. We construct the first deterministic sensing matrix by doing a Kronecker product between two DeVore's sensing matrices. To construct the second deterministic sensing matrix, rather than performing Kronecker product between two DeVore's sensing matrices, we use a novel row-by-row Kronecker product between the rows of two matrices which include DeVore's sensing matrices and circular shifts of DeVore's sensing matrices. In the sequel, the process of constructing the proposed deterministic sensing matrices and designing the pilot beam patterns are explained.

Since we need to have a deterministic sensing matrix for the sparse formulation, the beamforming (BF) codebooks at the BS and UE should have a deterministic structure. As we showed in the previous chapter, each measurement is made based on the Kronecker product of the two vectors $\mathbf{w}_t^{(1)^T}$ and $\mathbf{g}_t^{(1)^T}$. Therefore, if we can obtain the sensing matrix by performing a Kronecker product between two matrices $\mathbf{U}$ and $\mathbf{V}$, the rows of $\mathbf{U}$ and $\mathbf{V}$ (after applying a transpose operation) can be used respectively as $\mathbf{w}_t^{(1)}$ at the BS and $\mathbf{g}_t^{(1)}$ at the UE to generate the pilot beam patterns for the channel estimation. The two following prepositions show that indeed this is possible when $\mathbf{U}$ and $\mathbf{V}$ themselves are sensing matrices.

Proposition 2 *[60] Let* $\mathbf{U}$ *and* $\mathbf{V}$ *satisfy RIP of order L with RIP constants* $\delta_{L,\mathbf{U}}$ *and* $\delta_{L,\mathbf{V}}$ *respectively, also define* $\mathbf{S} = \mathbf{U} \otimes \mathbf{V}$. $\mathbf{S}$ *satisfies RIP of order L with the RIP constant* $\delta_{L,\mathbf{S}} \leq \delta_{L,\mathbf{U}}\delta_{L,\mathbf{V}} + \delta_{L,\mathbf{U}} + \delta_{L,\mathbf{V}}$.

Proposition 3 *[61] If* $\mu(\mathbf{U})$ *and* $\mu(\mathbf{V})$ *are the mutual coherence of* $\mathbf{U}$ *and* $\mathbf{V}$ *respectively, and* $\mu(\mathbf{S})$ *is the mutual coherence of* $\mathbf{S}$ *where* $\mathbf{S} = \mathbf{U} \otimes \mathbf{V}$*, we have* $\mu(\mathbf{S}) = \max\{\mu(\mathbf{U}), \mu(\mathbf{V})\}$.

In this thesis, we use the approach proposed by DeVore in [62] to construct $\mathbf{U}$ and $\mathbf{V}$. DeVore designed binary deterministic sensing matrices with dimensions $p^2 \times p^{r+1}$ with mutual coherence $\frac{r}{p}$, where p is a prime power and $1 \leq r < p$. Also, the DeVore's matrix with normalized columns satisfies

RIP of order $L < \frac{p}{r} + 1$ with RIP constant $\delta_L = (L-1)\frac{r}{p}$

If we use the DeVore's approach, $\mathbf{U}$ and $\mathbf{V}$ respectively with dimensions $p_1^2 \times p_1^{r_1+1}$ and $p_2^2 \times p_2^{r_2+1}$ can be constructed. So, using preposition 3, the sensing matrix $\mathbf{S}$ can be constructed by the Kronecker product of $\mathbf{U}$ and $\mathbf{V}$, which has $\mu(\mathbf{S}) = \max\{\frac{r_1}{p_1}, \frac{r_2}{p_2}\}$. Note that the idea of making sensing matrices by doing Kronecker product between two other sensing matrices is originally introduced in [61] for the compressive sensing context. We employ this technique to design the pilot beam patterns (BF vectors) for the channel estimation process.

It is worth pointing out that we use the row-normalized version of $\mathbf{U}$ and $\mathbf{V}$, owing to the fact that we use unit-norm BF vectors in the channel estimation process. In the DeVore's matrix, the number of ones in each row is a constant value. Assuming that there are c_U ones in each row of $\mathbf{U}$, by multiplying the matrix by normalization factor $\frac{1}{\sqrt{c_U}}$, all the rows of $\mathbf{U}$ become normalized. In fact, referring to (3.14), we assume that $z_1 = c_U$ and we define $\tilde{\mathbf{w}}_t = \frac{\mathbf{w}_t^{(1)}}{\sqrt{z_1}}$. By the same token, and referring to (3.15), we have $z_2 = c_V$ where c_V is the number of ones in each row of $\mathbf{V}$. Also, we define $\tilde{\mathbf{g}}_t = \frac{\mathbf{g}_t^{(1)}}{\sqrt{z_2}}$. Note that the normalization mentioned does not change the mutual coherence of $\mathbf{U}$ and $\mathbf{V}$.

As we will discuss later in this chapter, obtaining the sensing matrix $\mathbf{S}$ by just using the Kronecker product between other two sensing matrices limits the number of pilot beam patterns which can be used by the BS and UE to probe the channel. This in fact degrades the performance. As an alternative approach, we construct another sensing matrix using $\mathbf{U}$ and $\mathbf{V}$, but rather than performing Kronecker product between $\mathbf{U}$ and $\mathbf{V}$, we use the Kronecker product between the rows of two matrices which include $\mathbf{U}$, circular shifts of $\mathbf{U}$ and $\mathbf{V}$. Throughout the rest of this thesis, we will call the first and second approach respectively the matrix-by-matrix Kronecker product (MbMKP) approach and the row-by-row Kronecker product (RbRKP) approach.

The construction of the binary sensing matrices by DeVore's approach is done using finite fields, so in the following section we will concisely review some properties of the finite fields, which are required for constructing the DeVore's sensing matrix.

4.1 Finite Fields

To define a field, first Abelian groups should be defined [63]. A set G with a binary operation + form an Abelian group, which is denoted by $\{G, +, 0\}$, if the following properties are satisfied:

1. The operation + is commutative. This indicates that for any $a, b \in G$, we have $a + b = b + a$.

2. The operation + is associative. This indicates that for any $a, b, c \in G$, we have $(a + b) + c = a + (b + c)$.

3. The operation + has an identity element, i.e., if we denote the identity element by 0, for any $a \in G$ we have $a + 0 = 0 + a = a$.

4. For any element $a \in G$, there exist an (additive) inverse of a, i.e., $-a \in G$ where $a + (-a) = (-a) + a = 0$

A finite set F with two binary operations + (addition) and $\cdot$ (multiplication) is a finite field (or a Galois field) denoted by $\{F, +, \cdot\}$ if the following properties hold:

1. $\{F, +, 0\}$ is an Abelian group.

2. $\{F - \{0\}, \cdot, 1\}$ is an Abelian group, which indicates that the nonzero elements of the field with multiplication form an Abelian group whose identity element is denoted by 1. For any nonzero $a \in F$, the multiplicative inverse is denoted by a^{-1}.

3. The multiplication is distributive over the addition, i.e., $a \cdot (b + c) = (b + c) \cdot a = a \cdot b + a \cdot c$.

Note that when the order of the field, i.e., the number of the elements of the field, is finite, it is denoted by $GF(\,.\,)$. The binary field $GF(2)$ with modulo-2 addition and multiplication is an example of a finite field.

Let q be the order of a finite field then $GF(q)$ exists if and only if $q = p^m$, where p is a prime number and m is a positive integer [63]. When $q = p$, the finite field with modulo-p addition and multiplication can be denoted by $GF(p) = \{0, 1, 2, ..., p - 1\}$, which is also called a prime field. For example, $GF(3) = \{0, 1, 2\}$ is a finite field with modulo-3 addition and multiplication. In the case that $q = p^m$, the finite field is known as an extension of $GF(p)$, and p is called the characteristic of $GF(p^m)$. To construct extension fields, polynomials over $GF(p)$ are used.

A polynomial with degree m over $GF(p)$ is defined as

$$g(X) = g_0 + g_1 X + g_2 X^2 + \cdots + g_m X^m \tag{4.1}$$

where $0 \leq i \leq m$, $g_m \neq 0$ and $g_i \in GF(p)$. For addition and multiplication of polynomials over finite fields, the addition and multiplication of coefficients are done modulo-p.

A polynomial is called monic if $g_m = 1$. Also, a polynomial is irreducible if it is not the product of two lower degree polynomials over the same finite field. A polynomial which is monic and irreducible

TABLE 4.1: Addition and Multiplication Over $GF(4)$ [63].

+	0	1	X	$X+1$
0	0	1	X	$X+1$
1	1	0	$X+1$	X
X	X	$X+1$	0	1
$X+1$	$X+1$	X	1	0

·	0	1	X	$X+1$
0	0	0	0	0
1	0	1	X	$X+1$
X	0	X	$X+1$	1
$X+1$	0	$X+1$	1	X

is a prime polynomial. To generate $GF(p^m)$, a prime polynomial, say $t(X)$, over $GF(p)$ is used. If we consider all the polynomials over $GF(p)$ with degree less than m, the set of all such polynomials with addition modulo-p and multiplication modulo-$t(X)$ constitute the finite field $GF(p^m)$. Note that the number of these polynomials is p^m. Also, to do the multiplication modulo-$t(X)$, first, two polynomials are multiplied, then the remainder of dividing the resulting polynomial by $t(X)$ is obtained [63]. It is worth mentioning that all the nonzero elements of $GF(p^m)$ can be obtained by powers of X.

For instance, we can consider the example given in [63]. $t(X) = X^2 + X + 1$ is a prime polynomial over $GF(2)$. This polynomial has $m = 2$, so we can use it to construct $GF(2^2) = GF(4)$. All the polynomials which have degree less than $m = 2$ over $GF(2)$ are 0, 1, X and $X + 1$. The set of all these polynomials with addition modulo-2 and multiplication modulo-$t(X)$ form the Galois filed $GF(4)$. Addition and multiplication of these polynomials are given in Table 4.1. Note that for this example we have $X = X^1$, $X + 1 = X^2$ and $1 = X^3$, which shows that all nonzero elements of $GF(4)$ can be written using powers of X.

In the following section, we will explain the process of generating DeVore's sensing matrix using the finite fields [62].

4.2 DeVore's Sensing Matrix

The process of constructing the DeVore's matrix using prime order finite fields is explained in this section, the same method can be applied for constructing the DeVore's matrix using extension finite fields. Note that we use the similar notation adopted in [62].

Let F denote the finite field $GF(p)$ and consider the set of ordered pairs $\mathcal{S}_F = F \times F$ which has p^2 elements. Also, for any given integer r where $0 < r < p$, define the set $\mathcal{P}_r$ to include all polynomials (over F) of degree less than or equal r. Note that the number of such polynomials is p^{r+1}. A generic polynomial Q in set $\mathcal{P}_r$ can be written as $Q(X) = a_0 + a_1X + \cdots + a_rX^r$, where all the coefficients $a_0, a_1, ..., a_r$ are the elements of F. A generic polynomial in $\mathcal{P}_r$ can be considered as a mapping Q from F to F with the graph $m(Q)$ which is the set of ordered pairs $(X, Q(X))$. Note that such a graph

TABLE 4.2: All the polynomials in $\mathcal{P}_r$.

Lexicographically ordered coefficients	Polynomials	The ordered pairs $(X, Q(X))$
0 0 0	0	$(0,0),(1,0),(2,0)$
0 0 1	1	$(0,1),(1,1),(2,1)$
0 0 2	2	$(0,2),(1,2),(2,2)$
0 1 0	X	$(0,0),(1,1),(2,2)$
0 1 1	$X+1$	$(0,1),(1,2),(2,0)$
0 1 2	$X+2$	$(0,2),(1,0),(2,1)$
0 2 0	$2X$	$(0,0),(1,2),(2,1)$
0 2 1	$2X+1$	$(0,1),(1,0),(2,2)$
0 2 2	$2X+2$	$(0,2),(1,1),(2,0)$
1 0 0	X^2	$(0,0),(1,1),(2,1)$
1 0 1	X^2+1	$(0,1),(1,2),(2,2)$
1 0 2	X^2+2	$(0,2),(1,0),(2,0)$
1 1 0	X^2+X	$(0,0),(1,2),(2,0)$
1 1 1	X^2+X+1	$(0,1),(1,0),(2,1)$
1 1 2	X^2+X+2	$(0,2),(1,1),(2,2)$
1 2 0	X^2+2X	$(0,0),(1,0),(2,2)$
1 2 1	X^2+2X+1	$(0,1),(1,1),(2,0)$
1 2 2	X^2+2X+2	$(0,2),(1,2),(2,1)$
2 0 0	$2X^2$	$(0,0),(1,2),(2,2)$
2 0 1	$2X^2+1$	$(0,1),(1,0),(2,0)$
2 0 2	$2X^2+2$	$(0,2),(1,1),(2,1)$
2 1 0	$2X^2+X$	$(0,0),(1,0),(2,1)$
2 1 1	$2X^2+X+1$	$(0,1),(1,1),(2,2)$
2 1 2	$2X^2+X+2$	$(0,2),(1,2),(2,0)$
2 2 0	$2X^2+2X$	$(0,0),(1,1),(2,0)$
2 2 1	$2X^2+2X+1$	$(0,1),(1,2),(2,1)$
2 2 2	$2X^2+2X+2$	$(0,2),(1,0),(2,2)$

is a subset of $\mathcal{S}_F$.

After ordering the elements of $\mathcal{S}_F$ lexicographically as $(0,0),(0,1),...,(p-1,p-1)$, we define a vector v_Q for each polynomial $Q(X)$ in $\mathcal{P}_r$ such that it is indexed on $\mathcal{S}_F$ and it takes one at any ordered pair from the graph $m(Q)$ and takes zero otherwise. Out of the first p components of the vector v_Q, one of them takes the value one and the rest take the value zero. Similarly, out of the second p components, one of them takes the value one and the rest take the value zero, and so forth. Thus, the vector v_Q has exactly p ones.

Finally, we order all the vectors v_Q where $Q \in \mathcal{P}_r$ lexicographically with respect to the coefficients of the polynomials. The columns of the DeVore's sensing matrix are these ordered vectors. Because each vector v_Q has p^2 components and the number of all the polynomials in $\mathcal{P}_r$ is p^{r+1} which is also the number of all vectors v_Q, the dimensions of the DeVore's sensing matrix is $p^2 \times p^{r+1}$ [62]. To make

the process of generating the DeVore's sensing matrix more clear, we provide the following example:

We will follow the process mentioned above to generate the DeVore's sensing matrix with $p = 3$ and $r = 2$. With these values the resulting sensing matrix is 9×27. For this example, we have $GF(3) = \{0, 1, 2\}$ and consequently $\mathcal{S}_F = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,0), (2,1), (2,2)\}$ which is ordered lexicographically. All the polynomials of degree less than or equal to $r = 2$ constitute $\mathcal{P}_r$. All these polynomials with lexicographically ordered coefficients are shown in Table 4.2. Also, all the ordered pairs $(X, Q(X))$ of the graphs $m(Q)$ corresponding to the polynomials are presented in Table 4.2.

Now, to construct the vectors v_Q, we consider the third column of Table 4.2 and the set $\mathcal{S}_F$. We will describe how to construct v_Q for the polynomial $2X^2 + 2X + 1$ in Table 4.2. Constructing v_Q for other polynomials is done in the same way. As mentioned, v_Q is indexed on $\mathcal{S}_F$, so for this example v_Q is indexed on $\{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,0), (2,1), (2,2)\}$. The ordered pairs $(X, Q(X))$ of the graph $m(Q)$ corresponding to the polynomial $2X^2 + 2X + 1$ are $(0,1)$, $(1,2)$ and $(2,1)$ which are indexed 2, 6 and 8 with respect to $\mathcal{S}_F$, so the second, sixth and eighth component of the vector v_Q take the value one and the other components are zero, i.e., $v_Q = [0,1,0,0,0,1,0,1,0]^T$.

Finally, to obtain the DeVore's sensing matrix, we horizontally stack all the vectors v_Q ordered based on the first column of Table 4.2 which shows the lexicographical order of the coefficients of the polynomials.

4.3 MbMKP Approach

So far, we have described how to construct DeVore's sensing matrix. As mentioned in the beginning of this chapter, for the MbMKP approach we use the Kronecker product of two DeVore's sensing matrices $\mathbf{U}$ and $\mathbf{V}$ to construct the sensing matrix $\mathbf{S}$ for our sparse formulation (3.19), so

$$\mathbf{S} = \mathbf{U}_{p_1^2 \times p_1^{r_1+1}} \otimes \mathbf{V}_{p_2^2 \times p_2^{r_2+1}}. \tag{4.2}$$

Because the number of the columns of $\mathbf{U}$ and $\mathbf{V}$ are $p_1^{r_1+1}$ and $p_2^{r_2+1}$ respectively, the number of the columns of the sensing matrix $\mathbf{S}$ is $p_1^{r_1+1} \times p_2^{r_2+1}$ which is also the number of components of the sparse vector $\mathbf{h}_v$ in (3.18). The number of components of $\mathbf{h}_v$ is the product of the number of antennas at the BS (N_T) and the number of antennas at the UE (N_R), so we assume that N_T and N_R are equal to the number of columns of $\mathbf{U}$ and $\mathbf{V}$ respectively, i.e., $N_T = p_1^{r_1+1}$ and $N_R = p_2^{r_2+1}$. In other words, in

this thesis, the number of antennas at the BS and UE are assumed to be prime powers. Note that for constructing DeVore's sensing matrices $\mathbf{U}$ and $\mathbf{V}$, p_1 and p_2 can be prime powers or prime numbers. For example, to construct $\mathbf{U}$, we can set $p_1 = 2^2$ which is a prime power, and if we set $r = 2$ then $N_T = p_1^{r_1+1} = 64$.

The matrices $\mathbf{U}$ and $\mathbf{V}$ can be written as follows:

$$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_{p_1^2}^T \end{pmatrix} , \quad \mathbf{V} = \begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_{p_2^2}^T \end{pmatrix}. \tag{4.3}$$

We use each row of $\mathbf{U}$ as the vectors $\tilde{\mathbf{w}}_t$ and each row of $\mathbf{V}$ as the vectors $\tilde{\mathbf{g}}_t$, i.e., we have $\tilde{\mathbf{w}}_j = \mathbf{u}_j$ for $j = 1, 2, ..., p_1^2$ and $\tilde{\mathbf{g}}_j = \mathbf{v}_b$ for $b = 1, 2, ..., p_2^2$. Therefore, referring to (4.2) and (4.3), the sensing matrix $\mathbf{S}$ will be given as

$$\mathbf{S} = \begin{pmatrix} \mathbf{u}_1^T \otimes \mathbf{V} \\ \mathbf{u}_2^T \otimes \mathbf{V} \\ \vdots \\ \mathbf{u}_{p_1^2}^T \otimes \mathbf{V} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{w}}_1^T \otimes \tilde{\mathbf{g}}_1^T \\ \tilde{\mathbf{w}}_1^T \otimes \tilde{\mathbf{g}}_2^T \\ \vdots \\ \tilde{\mathbf{w}}_1^T \otimes \tilde{\mathbf{g}}_{p_2^2}^T \\ \tilde{\mathbf{w}}_2^T \otimes \tilde{\mathbf{g}}_1^T \\ \tilde{\mathbf{w}}_2^T \otimes \tilde{\mathbf{g}}_2^T \\ \vdots \\ \tilde{\mathbf{w}}_2^T \otimes \tilde{\mathbf{g}}_{p_2^2}^T \\ \vdots \\ \tilde{\mathbf{w}}_{p_1^2}^T \otimes \tilde{\mathbf{g}}_1^T \\ \tilde{\mathbf{w}}_{p_1^2}^T \otimes \tilde{\mathbf{g}}_2^T \\ \vdots \\ \tilde{\mathbf{w}}_{p_1^2}^T \otimes \tilde{\mathbf{g}}_{p_2^2}^T \end{pmatrix}. \tag{4.4}$$

As seen, each vector $\tilde{\mathbf{w}}_j$, where $j = 1, 2, ..., p_1^2$, is repeated p_2^2 times for all the different vectors $\tilde{\mathbf{g}}_b$ where $b = 1, 2, ..., p_2^2$. This means that the BS repeats the same transmit BF vector (or the same beam pattern) for p_2^2 times while the UE probes the channel using its all possible beam patterns. Then, the BS uses its second beam pattern and repeats it for p_2^2 times while the UE again probes the channel using its all possible beam patterns. This process continues until all possible combinations of the BS's

beam patterns and UE's beam patterns are used to probe the channel.

For instance, if we use $p_1 = 4$, $p_2 = 3$, $r_1 = r_2 = 2$ which indicate $N_T = p_1^{r_1+1} = 64$ and $N_R = p_2^{r_2+1} = 27$, there exist $p_1^2 = 16$ and $p_2^2 = 9$ different beam patterns for the BS and UE respectively. The BS repeats the same beam pattern for 9 times while the UE measures the channel by its different 9 beam patterns. This process is repeated until the BS uses its 16 different beam patterns.

Note that as mentioned in chapter 3 and illustrated in Fig. 3.2, each pilot beam pattern at the BS or UE is formed by several narrow beams. Each nonzero value in the vectors $\tilde{\mathbf{w}}_t$ or $\tilde{\mathbf{g}}_t$ represents a narrow beam. In other words, in the DeVore's matrices $\mathbf{U}$ and $\mathbf{V}$, each row which is considered as a BF vector in our proposed scheme has several nonzero values, indicating that each pilot beam pattern generated by the BS or UE to measure the channel consists of multiple narrow beams.

Returning to our example above, for any $\tilde{\mathbf{w}}_t$, out of 64 components, 16 components have a nonzero value, so, there are 16 narrow beams in each pilot beam pattern of the BS. Also, for any $\tilde{\mathbf{g}}_t$, out of 27 components, 9 components are nonzero, which means that there are 9 narrow beams in each pilot beam pattern of the UE. Note that in fact the DeVore's approach determines the number of nonzero values in each row of $\mathbf{U}$ and $\mathbf{V}$ or equivalently the number of nonzero values in $\tilde{\mathbf{w}}_t$ and $\tilde{\mathbf{g}}_t$.

If we consider the sensing matrix $\mathbf{S}$ in (4.4), we can see that the number of measurements are $p_1^2 p_2^2 = 144$. In our example, the BS uses $p_1^2 = 16$ different pilot beam patterns to probe the channel in all 144 measurements. On the other hand, we intuitively know that if we increase the number of different pilot beam patterns, the possibility of finding the propagation paths between the BS and UE increases. Indeed, if we increase the number of pilot beam patterns of the BS and UE up to 144, the full potential in making independent and different measurements is exploited. So, just using 16 different pilot beam patterns for 144 measurements is not an efficient approach. Inspired by this fact, we will propose RbRKP approach in the next section.

4.4 RbRKP Approach

As discussed, in the MbMKP approach, we used the Kronecker product of two DeVore's sensing matrices $\mathbf{U}$ and $\mathbf{V}$ (with normalized rows) to make the sensing matrix $\mathbf{S}$. In this section, we use the Kronecker product between each row of $\mathbf{U}$ (and its circular shifts) and each row of $\mathbf{V}$. In so doing, we can increase the number of different pilot beam patterns at the BS. Note that for this approach, the number of rows of the two DeVore's sensing matrices we employ must be equal, so in obtaining $\mathbf{U}$ and $\mathbf{V}$, we use $p_1 = p_2$, but r_1 and r_2 can be different; thus, the number of columns of $\mathbf{U}$ and $\mathbf{V}$ can be

different.

If we denote by $\otimes_r$ the row-by-row Kronecker product of two matrices, referring to (4.3), the matrix $\hat{\mathbf{S}}$ which is obtained using row-by-row Kronecker product of $\mathbf{U}$ and $\mathbf{V}$ is given as

$$\hat{\mathbf{S}}_{p_1^2 \times p_1^{r_1+r_2+2}} = \mathbf{U} \otimes_r \mathbf{V} = \begin{pmatrix} \mathbf{u}_1^T \otimes \mathbf{v}_1^T \\ \mathbf{u}_2^T \otimes \mathbf{v}_2^T \\ \vdots \\ \mathbf{u}_{p_1^2}^T \otimes \mathbf{u}_{p_1^2}^T \end{pmatrix}. \tag{4.5}$$

We have observed numerically that the mutual coherence parameter for $\hat{\mathbf{S}}$ is not less than 1, indicating that $\hat{\mathbf{S}}$ cannot be used as a sensing matrix for sparse reconstruction. We propose the following process to obtain a sensing matrix, i.e., a matrix with the mutual coherence less than one, using $\hat{\mathbf{S}}$.

First, we need to define some matrices. If we circularly shift all the columns of the DeVore's sensing matrix $\mathbf{U}$ by i positions to the right, we get the matrix $\mathbf{U}_{CSi}$. Also, we define $\hat{\mathbf{S}}_i$ which is the row-by-row Kronecker product between $\mathbf{U}_{CSi}$ and $\mathbf{V}$. Mathematically speaking, we have

$$\hat{\mathbf{S}}_i = \mathbf{U}_{CSi} \otimes_r \mathbf{V}. \tag{4.6}$$

Finally, by stacking $\hat{\mathbf{S}}$ and all the matrices $\hat{\mathbf{S}}_i$ for $i = 1, 2, ..., n$ vertically, we obtain the matrix $\hat{\mathbf{B}}_n$ as follows:

$$\hat{\mathbf{B}}_n = \begin{pmatrix} \hat{\mathbf{S}} \\ \hat{\mathbf{S}}_1 \\ \vdots \\ \hat{\mathbf{S}}_n \end{pmatrix}. \tag{4.7}$$

We have provided Fig. 4.1 to numerically verify that the mutual coherence of the matrix $\hat{\mathbf{B}}_n$ becomes less than one when n increases. In this figure, we have plotted the mutual coherence of the matrix $\hat{\mathbf{B}}_n$ for three cases where $N_T = N_R = 27$, $N_T = N_R = 64$ and $N_T = N_R = 125$. For a certain n, say $n = n_s$, the matrix $\hat{\mathbf{B}}_n$ becomes a sensing matrix, i.e., its mutual coherence parameter becomes less than one. The value of n_s is provided in the figure for the three cases.

The process of constructing the sensing matrix $\hat{\mathbf{B}}_{n_s}$ using the RbRKP approach is presented in Algorithm 1.

Note that Algorithm 1 gives the first sensing matrix. After getting the first sensing matrix by Algorithm 1, if we continue the process described, we will get other sensing matrices with more

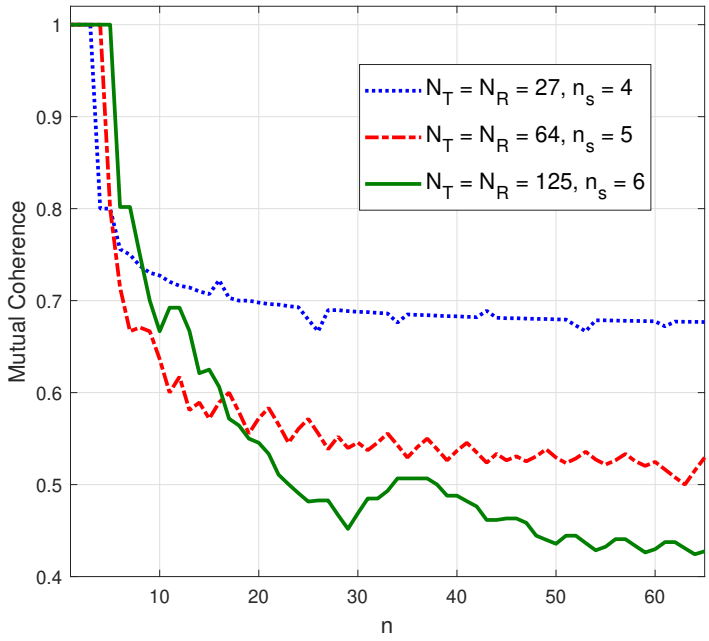


FIGURE 4.1: Mutual coherence of the matrix $\hat{\mathbf{B}}_n$ vs n

number of rows, which is required when we need to make more measurements.

We define the matrix $\hat{\mathbf{U}}$ which is constructed by stacking vertically $\mathbf{U}$ and all $\mathbf{U}_{CSi}$ for $i = 1, 2, ..., n_s$, and also, we define the matrix $\hat{\mathbf{V}}$ which is constructed by stacking vertically the matrix $\mathbf{V}$ for n_s times, i.e.,

$$\hat{\mathbf{U}} = \begin{pmatrix} \mathbf{U} \\ \mathbf{U}_{CS1} \\ \vdots \\ \mathbf{U}_{CSn_s} \end{pmatrix}, \quad \hat{\mathbf{V}} = \begin{pmatrix} \mathbf{V} \\ \mathbf{V} \\ \vdots \\ \mathbf{V} \end{pmatrix}, \tag{4.8}$$

so, it is obvious that $\hat{\mathbf{B}}_{n_s} = \hat{\mathbf{U}} \otimes_r \hat{\mathbf{V}}$.

Based on the RbRKP approach, the BS and UE use the rows of the matrices $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ respectively to generate the pilot beam patterns. In other words, the rows of $\hat{\mathbf{U}}$ are used as the vectors $\tilde{\mathbf{w}}_t$ at the BS and the rows of $\hat{\mathbf{V}}$ are used as the vectors $\tilde{\mathbf{g}}_t$ at the UE.

4.5 Beam Alignment Process

For the MbMKP approach, knowing p_1 and r_1, the BS can construct the DeVore's sensing matrix $\mathbf{U}$. Also, the UE can construct the DeVore's sensing matrix $\mathbf{V}$ with knowing p_2 and r_2. Since the channel estimation is done by the UE, the UE must know the values of p_1 and r_1 to construct $\mathbf{U}$ and

Algorithm 1 : Obtaining a sensing matrix by RbRKP approach.

Input p_1, r_1, r_2 and $n = 0$
Output The sensing matrix $\hat{\mathbf{B}}_{n_s}$ and the index n_s

1: Generate the matrices $\mathbf{U}$ and $\mathbf{V}$ based on the DeVore's approach.
2: Obtain $\hat{\mathbf{S}}$ by row-by-row Kronecker product between $\mathbf{U}$ and $\mathbf{V}$.
3: **Repeat**
4: Increment n.
5: Compute $\mathbf{U}_{CSn}$.
6: Compute $\hat{\mathbf{S}}_n$.
7: Stack vertically $\hat{\mathbf{S}}$ and $\hat{\mathbf{S}}_l$ for $l = 1, 2, ..., n$ to obtain $\hat{\mathbf{B}}_n$.
8: Compute the mutual coherence for $\hat{\mathbf{B}}_n$, i.e., $\mu(\hat{\mathbf{B}}_n)$.
9: **Until** $\mu(\hat{\mathbf{B}}_n) < 1$.
10: $\hat{\mathbf{B}}_{n_s} \leftarrow \hat{\mathbf{B}}_n$ and $n_s \leftarrow n$.

consequently $\mathbf{S}$. So it is assumed that the BS signals the parameters p_1 and r_1 to the UE before the process of BA starts, or it is assumed that these two parameters are stored in the UE's memory. The UE, knowing p_1 and r_1, makes the DeVore's sensing matrix $\mathbf{U}$ and then it can construct the sensing matrix $\mathbf{S}$ by the Kronecker product of $\mathbf{U}$ and $\mathbf{V}$ as discussed earlier. Measurements are made by the BS and UE based on the pilot beam patterns which are generated using the rows of the matrices $\mathbf{U}$ and $\mathbf{V}$. Finally the UE estimates the channel using the sensing matrix $\mathbf{S}$ and the values of measurements.

For the RbRKP approach, the BS can construct the matrix $\hat{\mathbf{U}}$ by p_1, r_1 and n_s. The UE can also construct the matrix $\hat{\mathbf{V}}$ by knowing p_2, r_2 and n_s. Note that as explained earlier for the RbRKP approach $p_2 = p_1$. We assume that the values of p_1, r_1 and n_s are sent to the UE by the BS before the BA procedure or these values are stored in the UE's memory. Therefore, the UE can construct $\hat{\mathbf{U}}$ and consequently $\hat{\mathbf{B}}_{n_s}$. Note that based on Algorithm 1, for different values of p_1, r_1 and r_2 the value of n_s is determined, so one assumption made here is that the BS knows the value of r_2 used by the UE or the UE sends the value of r_2 to the BS before BA process. Equivalently, if the BS knows N_R, the value of r_2 is obtained by the BS using $N_R = p_1^{r_2+1}$.

It should be pointed out that the number of measurements for the MbMKP is always equal to $p_1^2 p_2^2$ which is constrained by the values of p_1 and p_2. For the RbRKP approach, the number of measurements are $p_1^2(n_s + 1)$. As mentioned earlier, n_s determines the first sensing matrix $\hat{\mathbf{B}}_{n_s}$ constructed based on Algorithm 1. For any $n > n_s$, Algorithm 1 results in sensing matrices with $p_1^2(n + 1)$ number of rows. When more measurements are needed sensing matrices constructed by $n > n_s$ can be used.

TABLE 4.3: Overhead of Different Methods

Method	**Overhead**
MbMKP	$O(\log p_1)$
RbRKP	$O(\log n_s)$
[39]	$O(M(r_1+1)\log p_1)$
Bernoulli Random BF	$O(Mp_1^{r_1+1})$

4.6 Overhead Calculation

The overhead of the MbMKP approach depends on the values of p_1 and r_1 which can be signaled to or stored in the UE by $(\log_2 p_1 + \log_2 r_1)$ bits. We know that $r_1 < p_1$; therefore, the overhead of the MbMKP approach is $O(\log p_1)$.

The overhead of the RbRKP approach is calculated based the values of p_1, r_1, r_2 and n_s. The values of p_1, r_1 and n_s can be signaled to or stored in the UE by $(\log_2 p_1 + \log_2 r_1 + \log_2 n_s)$ bits. In addition, the value of r_2 can be sent to the BS by the UE (or it can be stored in the BS) with $\log_2 r_2$ bits. Therefore, the overall overhead is $(\log_2 p_1 + \log_2 r_1 + \log_2 r_2 + \log_2 n_s)$ bits. We know that $r_1, r_2 < p_1$, and without loss of generality we assume $p_1 < n_s$, so the overhead of the RbRKP approach is $O(\log n_s)$.

Table 4.3 compares the overhead of our approaches with that of in [39] and also the approach which uses the Bernoulli random BF codebook. The overhead of the proposed approach in [39] is $O(M \log N_T)$, where M is the number of measurements. Note that in our approaches $N_T = p_1^{r_1+1}$, and also the number of measurements in our approaches has no effect on the overhead. However, as seen in Table 4.3, the overhead of the proposed approach in [39] and the approach using the Bernoulli random BF codebook scale linearly with the number of measurements.

Chapter **5**

Simulation Results

In this chapter, we present the simulation results and compare the performance of our proposed approaches with other methods. As mentioned in chapter 4, the performance of our first proposed approach (MbMKP) is limited, and the number of measurements is always constant for this approach. So, in this chapter, we mostly present the results of our second proposed approach (RbRKP).

The simulation setup and performance metrics we have used in our simulations are described as follows:

- The BS and UE are equipped with $N_T = N_R = \{27, 64, 125\}$ antennas, i.e., the pairs of $\{p_1 = p_2 = 3, r_1 = r_2 = 2\}$, $\{p_1 = p_2 = 4, r_1 = r_2 = 2\}$ and $\{p_1 = p_2 = 5, r_1 = r_2 = 2\}$ are used. The number of antennas at the BS and UE are equal throughout all simulations. Also, the numbers in the legend of the figures show the number of antennas.

- The AoAs and AoDs of propagation paths have uniform distribution within the angular range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

- The average power of the channel gain of lth propagation path is one, i.e., $\sigma_{\alpha_l}^2 = 1$.

- To reconstruct the sparse vector the orthogonal matching pursuit (OMP) algorithm is used.

- The ratio of the average received power of the pilot signals to the noise power is considered as the SNR in our simulations, i.e., SNR = $\frac{P}{\sigma_n^2}$.

- The normalized mean square error is used as one of the performance metrics. NMSE is defined as follows:

$$\text{NMSE} = 10\log_{10}\left(E\left[\frac{\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2}{\|\mathbf{H}\|_F^2}\right]\right), \tag{5.1}$$

  where $\hat{\mathbf{H}}$ is the estimate of the channel and $\|\,.\,\|_F$ denotes the Frobenius norm.

- In mmWave systems, typically the SNR in the beam alignment process (SNR before beamforming) is very low [36][64]. Thus, it is reasonable that first the directions of the propagation paths between the BS and UE are found, and then the path gains are estimated when the beams are aligned in those directions by the BS and UE. So, in addition to NMSE which might not be an effective performance metric in very low SNR regimes, as another performance metric, we use the probability of correctly estimating (PCE) at least one of the directions of the propagation paths, which is equivalent to the probability of correctly estimating at least one of the indexes of the nonzero elements in the sparse vector $\mathbf{h}_v$.

- We use the SNR after BF ($\mathrm{SNR_{AB}}$) as another performance metric. If we denote by ε the index of a nonzero element in $\mathbf{h}_v$ after the beam alignment process, the indexes of the value one in the vectors $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$ are calculated by $\varepsilon_{\mathbf{w}} = \lfloor(\varepsilon-1)/N_R\rfloor + 1$ and $\varepsilon_{\mathbf{g}} = \big((\varepsilon-1) \mod N_R\big) + 1$ respectively [39]. Then, the $\mathrm{SNR_{AB}}$ is calculated as follows:

$$\mathrm{SNR_{AB}} = \frac{|\mathbf{g}_t^{(1)^H}\mathbf{H}_v\mathbf{w}_t^{(1)}|^2}{\sigma_n^2} \tag{5.2}$$

We compare the performance of our approaches with the following methods:

1. The method adopted in [36]. Based on this method, the number of ones in $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$ are constant but their positions are randomly permuted, which makes random BF vectors. We call this method random permutation and we denote it by the abbreviation RdPerm. Note that in our approach the number of ones and their positions in $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$ are constant because we have designed the BF vectors based on our proposed deterministic sensing matrices.

2. The method which is based on a structured random sensing matrix proposed in [39]. We denote this method by the abbreviation StRd. Note that, in the StRd approach , BF vectors are designed by complex linear combinations of the columns of the DFT matrix using the coefficients which are unit magnitude complex values. By contrast, following the RdPerm approach, in designing the BF vectors, we have used the linear combinations of the columns of the DFT matrix using the coefficients having the value zero or one (see (3.14) and (3.15)).

3. The method using Bernoulli random values (zeros and ones with $p = 0.5$) for the elements of $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$. Hereafter, we denote this method by the abbreviation RdBR.

In order to compare the performance of other methods with our approach (RbRKP) for low SNR

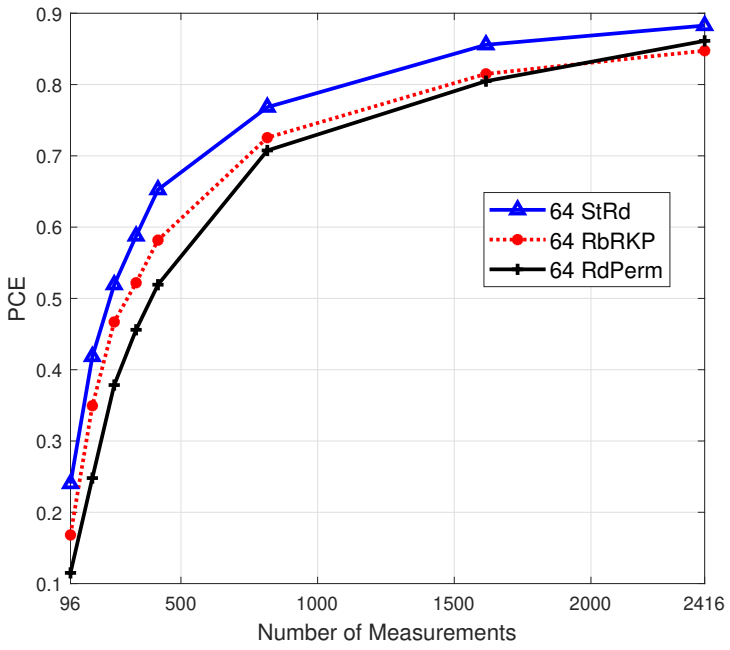


FIGURE 5.1: PCE vs number of measurements for StRd, RdPerm and RbRKP when SNR = -10 dB and $L = 1$.

scenarios, we have provided Fig. 5.1 and Fig. 5.2. The results of theses two figures are obtained for SNR = -10 dB. As mentioned in the beginning of this chapter, PCE is a suitable performance metric for low SNR scenarios. Therefore, we have used PCE as the performance metric for these figures.

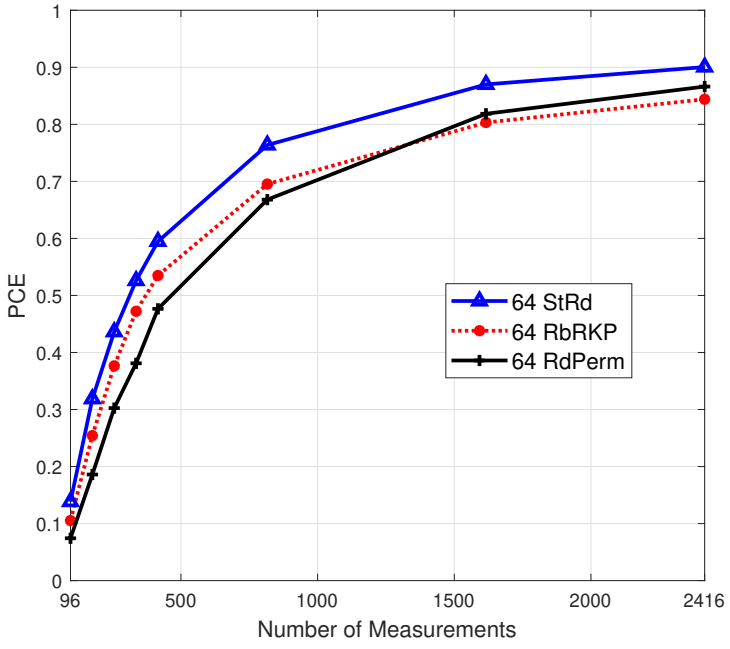


FIGURE 5.2: PCE vs number of measurements for StRd, RdPerm and RbRKP when SNR = -10 dB and $L = 2$.

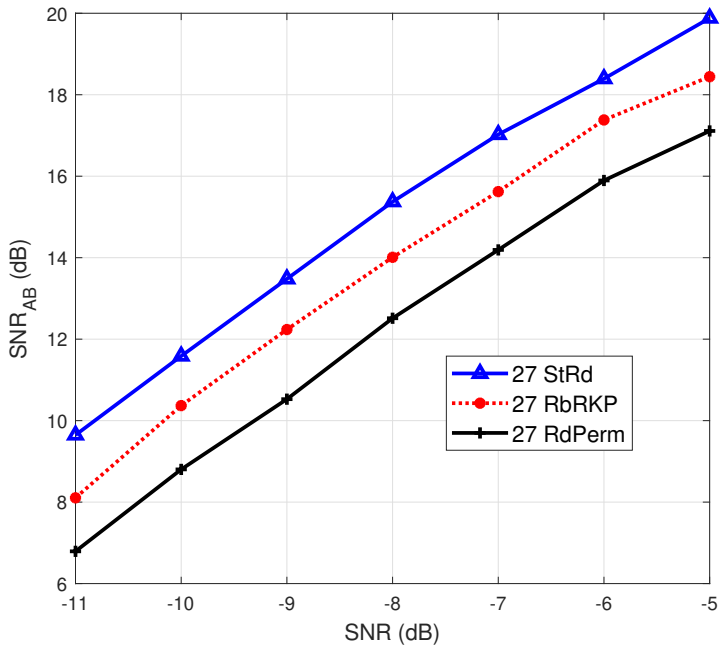


FIGURE 5.3: SNR_{AB} vs SNR when the number of measurements is 45 and $L = 1$.

Fig. 5.1 and Fig. 5.2 which show PCE against the number of measurements are for the cases when the channel has one and two clusters respectively. As illustrated in Fig. 5.1 and Fig. 5.2, our approach (RbRKP) has better performance than RdPerm but worse than StRd. Note that the purpose of compressed sensing is reconstructing the sparse vector with a very small number of measurements compared to the length of the sparse vector. For the results in Fig. 5.1 and Fig. 5.2 the length of the sparse vector $\mathbf{h}_v$ is 4096. Having said that, as seen in Fig. 5.1 and Fig. 5.2, the performance of our approach (RbRKP) is better than RdPerm for small number of measurements. Our approach (RbRKP) performs worse compared to RdPerm for the large number of measurements, but as mentioned, the large number of measurements is not useful in compressed sensing.

We have provided Fig. 5.3 and Fig. 5.4 to compare the SNR_{AB} of other methods with our approach (RbRKP) for low SNR scenarios. The results in Fig. 5.3 and Fig. 5.4 are obtained for the channel with one and two clusters respectively. As seen, the values of the SNR_{AB} for our approach (RbRKP) are more than those of RdPerm and less than those of StRd. For Fig. 5.4, since the channel has two clusters, for each channel realization, the average value of two SNR_{AB}s corresponding to the two propagation paths is obtained. In other words, first the vectors $\mathbf{w}_t^{(1)}$ and $\mathbf{g}_t^{(1)}$ corresponding to the direction of each propagation path are obtained, and then using (5.2), the SNR_{AB} for each propagation path is calculated. Finally, the average value of the two SNR_{AB}s is computed.

As mentioned in the beginning of this chapter, NMSE might not be a good performance metric

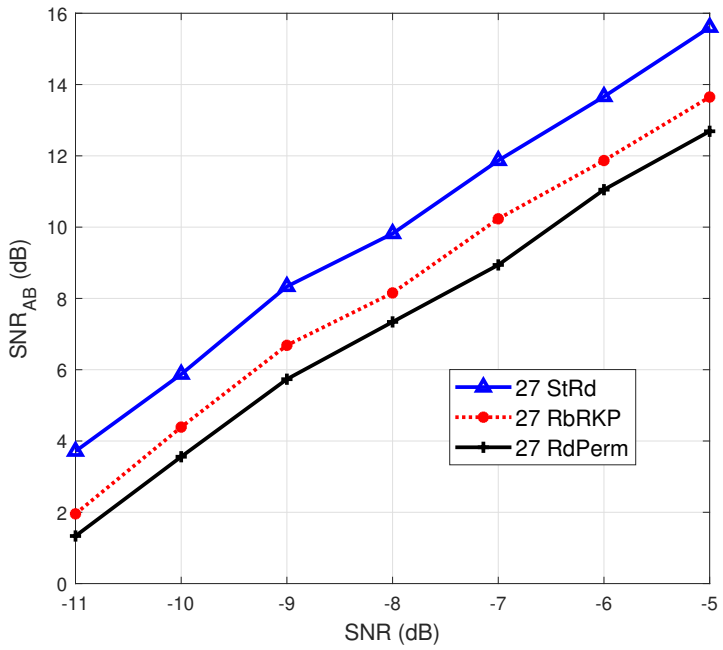


FIGURE 5.4: SNR_{AB} vs SNR when the number of measurements is 45 and $L = 2$.

because typically the SNR in the beam alignment process (SNR before BF) is very low for mmWave systems. However, since NMSE has been widely used in the compressed sensing context, we have provided some results for which we have used NMSE as the performance metric.

Fig. 5.5 and Fig. 5.6 compare the performance of our MbMKP approach with RdPerm and RdBR.

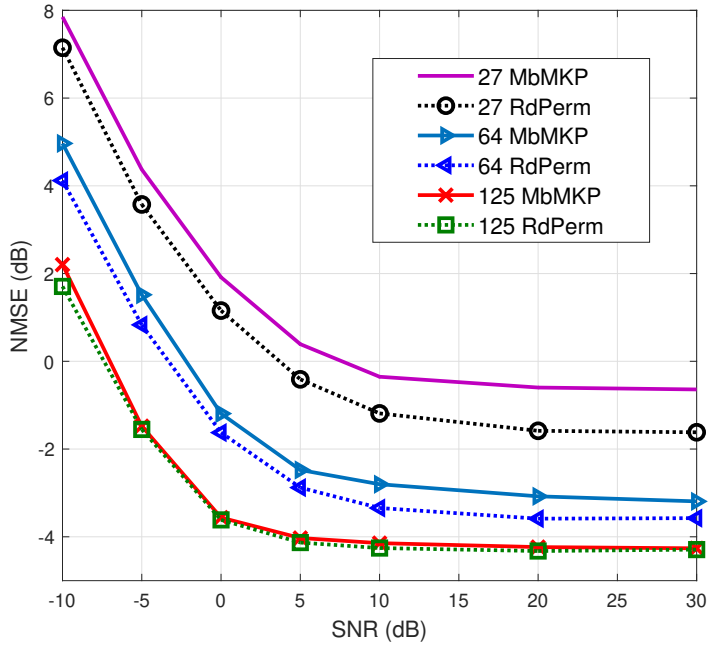


FIGURE 5.5: NMSE vs SNR for MbMKP and RdPerm when $L = 6$.

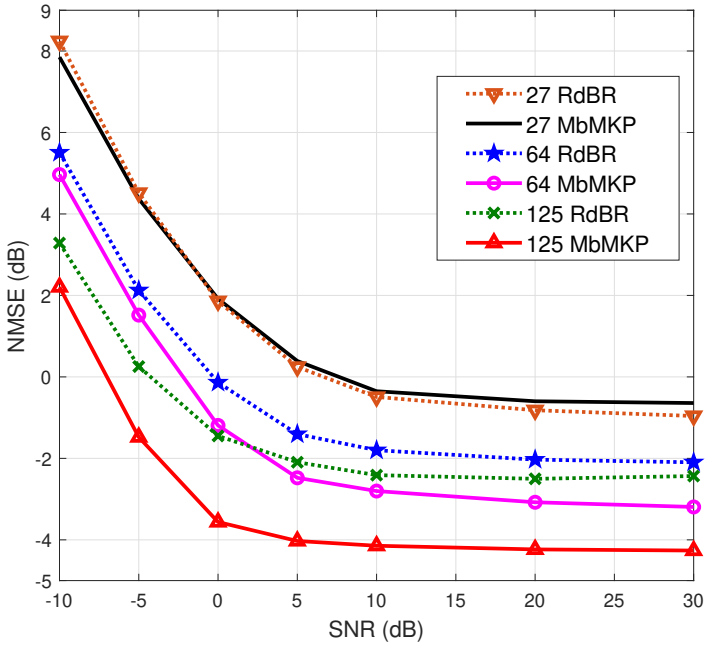


FIGURE 5.6: NMSE vs SNR for MbMKP and RdBR when $L = 6$.

Fig. 5.5 illustrates that the performance of our MbMKP approach becomes more comparable to RdPerm when the number of antennas increases. Also, Fig. 5.6 shows that our MbMKP approach outperforms RdBR when the number of antennas is 64 or 125.

Fig. 5.7 and Fig. 5.8 illustrate the performance of different methods in terms of the number of

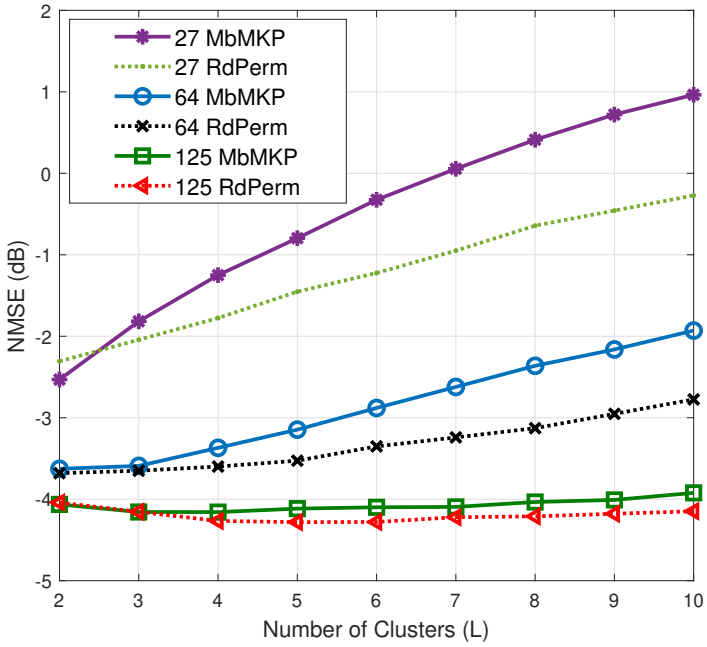


FIGURE 5.7: NMSE vs L for MbMKP and RdPerm when SNR = 10 dB.

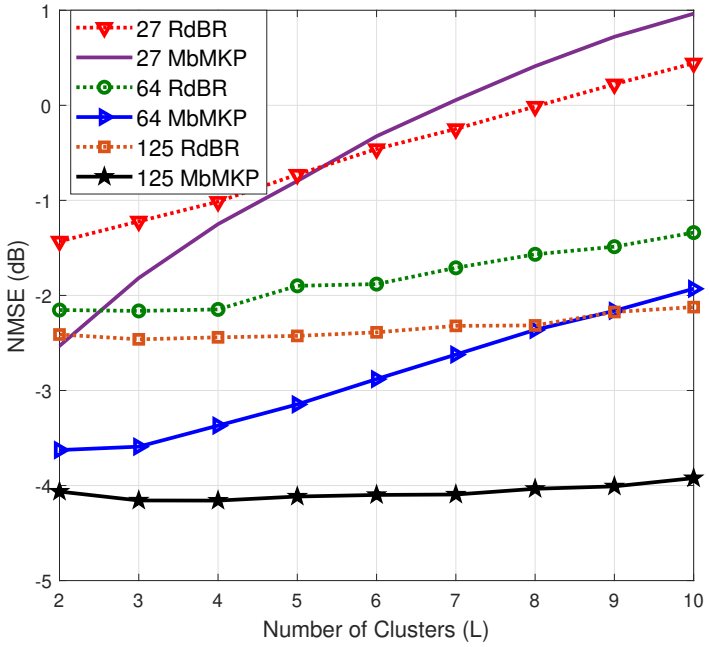


FIGURE 5.8: NMSE vs L for MbMKP and RdBR when SNR = 10 dB.

clusters or L. With increasing the number of antennas the performance of our MbMKP approach becomes closer to RdPerm. Also, as seen in Fig. 5.8, the performance of our MbMKP approach is better than RdBR when the number of antennas is 64 or 125.

Fig. 5.9 shows the results for our RbRKP approach and RdPerm. For this figure, the results are

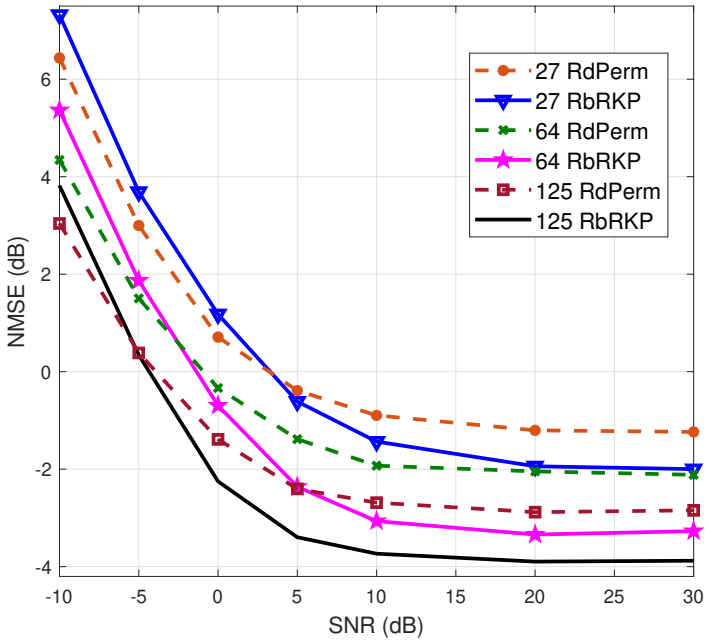


FIGURE 5.9: NMSE vs SNR for RbRKP and RdPerm when $L = 2$.

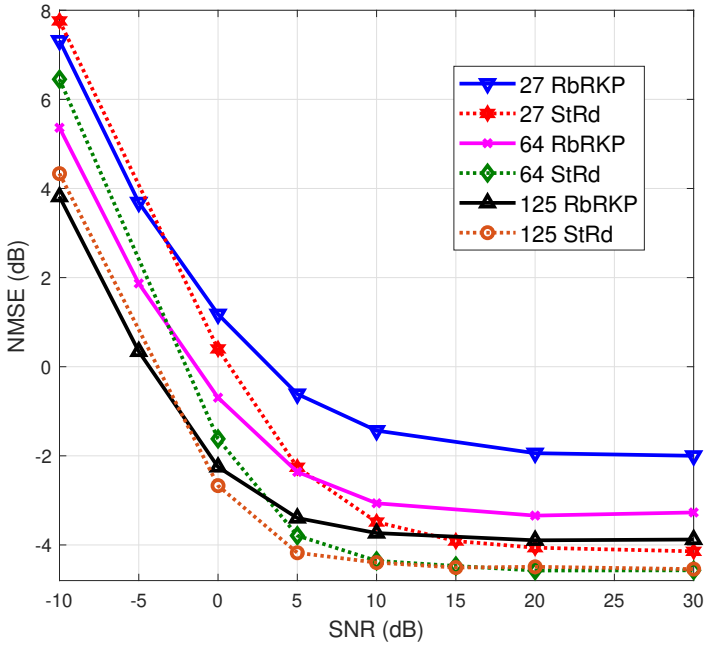


FIGURE 5.10: NMSE vs SNR for RbRKP and StRd when $L = 2$.

obtained for a channel with two propagation paths. In low SNR regimes which are typically the case for mmWave systems, our approach has superior performance in some situations. For example, when the BS and UE are equipped with 125 antennas, our approach shows better performance in the SNR range [-5 dB,0 dB].

Fig. 5.10 illustrates the performance of our RbRKP approach and StRd for the case when the number of clusters is 2. As seen, for low SNR scenarios, our RbRKP approach presents improvement.

Note that for Fig. 5.9 and Fig. 5.10, the number of measurements for the cases with 27 antennas, 64 antennas and 125 antennas are respectively 45, 96 and 175.

Chapter 6

Conclusion

In this thesis, we have studied the beam alignment problem in a mmWave system consisting of a multi-antenna BS and multi-antenna UEs. The BS and UEs are equipped with uniform linear arrays and they have one RF chain. Since a small number of clusters contributes to the propagation paths in the mmWave channels, we have used the clustered physical channel model which is adopted in the literature. Following the standard approach in the literature, which uses the idea of virtual channel model, we have presented the channel estimation problem by a sparse formulation. For sparse formulations, compressed sensing tools are leveraged to reconstruct the sparse vector. The sensing matrix in the sparse formulation has a key role in sparse recovery, and the structure of the sensing matrix in the beam alignment problem is determined based on the beamforming (BF) vectors employed by the BS and UEs. The literature has employed random or structured random sensing matrices which impose significant overhead.

Inspired by existing deterministic sensing matrices from the theory of compressive sensing, we have proposed two novel *deterministic* sensing matrices based on which we have designed the pilot beam patterns.

We have shown that our proposed approach results in significant overhead reduction compared to the proposed approach in [39](the StRd approach) and the approach using a Bernoulli random BF codebook. In terms of performance, our proposed approach presents improvement compared to the proposed method in [36], but our proposed approach is not as good as the StRd approach. The reason is that, in designing the BF vectors, we have used the linear combinations of the columns of the DFT matrix using the coefficients having the value zero or one, whereas in the StRd approach, BF vectors are designed by complex linear combinations of the columns of the DFT matrix using the coefficients which are unit magnitude complex values. Therefore, the StRd approach uses BF vectors which are much closer to the dominant eigenvector of the actual channel matrix, but at the cost of a large overhead. It might not be entirely fair to compare the performance of our proposed

approach with the StRd approach because the StRd approach employs a much wider class of possible BF vectors, but we have compared the performance of our proposed approach with the StRd approach to show the tradeoff between the performance and the overhead. Our proposed approach reduces the overhead significantly at the cost of losing performance.

For future work, it is interesting to see whether it is possible to reduce the overhead significantly and at the same time boost the performance. This can be achieved based on a complex deterministic sensing matrix and BF vectors which are complex linear combinations of the columns of the DFT matrix.

References

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski. *Five disruptive technology directions for 5g*. IEEE Communications Magazine **52**(2), 74 (2014).

[2] B. C.-P. J. Gozalvez. *Device-centric wireless networks for 5g*. In *XXX Simposium Nacional de la Unin Cientfica Internacional de Radio URSI* (2015).

[3] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. *What will 5g be?* IEEE Journal on Selected Areas in Communications **32**(6), 1065 (2014).

[4] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino. *Millimeter-wave massive mimo: the next wireless revolution?* IEEE Communications Magazine **52**(9), 56 (2014).

[5] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. *Millimeter wave mobile communications for 5g cellular: It will work!* IEEE Access **1**, 335 (2013).

[6] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath. *Channel estimation and hybrid precoding for millimeter wave cellular systems*. IEEE Journal of Selected Topics in Signal Processing **8**(5), 831 (2014).

[7] Z. Xiao, T. He, P. Xia, and X. Xia. *Hierarchical codebook design for beamforming training in millimeter-wave communication*. IEEE Transactions on Wireless Communications **15**(5), 3380 (2016).

[8] S. Rangan, T. S. Rappaport, and E. Erkip. *Millimeter-wave cellular wireless networks: Potentials and challenges*. Proceedings of the IEEE **102**(3), 366 (2014).

[9] W. Roh, J. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar. *Millimeter-wave beamforming as an enabling technology for 5g cellular communications: theoretical feasibility and prototype results*. IEEE Communications Magazine **52**(2), 106 (2014).

[10] X. Gao, L. Dai, S. Han, C. I, and R. W. Heath. *Energy-efficient hybrid analog and digital precoding for mmwave mimo systems with large antenna arrays*. IEEE Journal on Selected Areas in Communications **34**(4), 998 (2016).

[11] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip. *Millimeter wave channel modeling and cellular capacity evaluation*. IEEE Journal on Selected Areas in Communications **32**(6), 1164 (2014).

[12] T. S. Rappaport, Y. Qiao, J. I. Tamir, J. N. Murdock, and E. Ben-Dor. *Cellular broadband millimeter wave propagation and angle of arrival for adaptive beam steering systems*. In *2012 IEEE Radio and Wireless Symposium*, pp. 151–154 (2012).

[13] A. M. Sayeed. *Deconstructing multiantenna fading channels*. IEEE Transactions on Signal Processing **50**(10), 2563 (2002).

[14] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak. *Compressed channel sensing: A new approach to estimating sparse multipath channels*. Proceedings of the IEEE **98**(6), 1058 (2010).

[15] N. Czink, X. Yin, H. OZcelik, M. Herdin, E. Bonek, and B. H. Fleury. *Cluster characteristics in a mimo indoor propagation environment*. IEEE Transactions on Wireless Communications **6**(4), 1465 (2007).

[16] L. Vuokko, V. . Kolmonen, J. Salo, and P. Vainikainen. *Measurement of large-scale cluster power characteristics for geometric channel models*. IEEE Transactions on Antennas and Propagation **55**(11), 3361 (2007).

[17] A. Forenza, D. J. Love, and R. W. Heath. *Simplified spatial correlation models for clustered mimo channels with different array configurations*. IEEE Transactions on Vehicular Technology **56**(4), 1924 (2007).

[18] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic. *Millimeter wave mimo channel estimation using overlapped beam patterns and rate adaptation*. IEEE Transactions on Signal Processing **65**(3), 601 (2017).

[19] C. Liu, M. Li, S. V. Hanly, P. Whiting, and I. B. Collings. *Millimeter-wave small cells: Base station discovery, beam alignment, and system design challenges*. IEEE Wireless Communications **25**(4), 40 (2018).

[20] D. Zhang, H. Chen, M. Kokshoorn, Y. Li, N. Wei, and B. Vucetic. *A probe-then-refine beam tracking algorithm for millimeter wave miso systems*. In *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6 (2018).

[21] J. Seo, Y. Sung, G. Lee, and D. Kim. *Training beam sequence design for millimeter-wave mimo systems: A pomdp framework*. IEEE Transactions on Signal Processing **64**(5), 1228 (2016).

[22] C. Liu, M. Li, S. V. Hanly, I. B. Collings, and P. Whiting. *Millimeter wave beam alignment: Large deviations analysis and design insights*. IEEE Journal on Selected Areas in Communications **35**(7), 1619 (2017).

[23] Q. Xue, X. Fang, and M. Xiao. *Beam management for millimeter wave beamspace mu-mimo systems*. In *2017 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–6 (2017).

[24] A. Sayeed and J. Brady. *Beamspace mimo channel modeling and measurement: Methodology and results at 28ghz*. In *2016 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6 (2016).

[25] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath. *Channel estimation and hybrid combining for mmwave: Phase shifters or switches?* In *2015 Information Theory and Applications Workshop (ITA)*, pp. 90–97 (2015).

[26] R. Hu, J. Tong, J. Xi, Q. Guo, and Y. Yu. *Robust channel estimation for switch-based mmwave mimo systems*. In *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–7 (2017).

[27] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank. *Sensitivity to basis mismatch in compressed sensing*. IEEE Transactions on Signal Processing **59**(5), 2182 (2011).

[28] E. J. Candes and Y. Plan. *Matrix completion with noise*. Proceedings of the IEEE **98**(6), 925 (2010).

[29] A. Alkhateeb, G. Leus, and R. W. Heath. *Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?* In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2909–2913 (2015).

[30] J. Lee, G. Gil, and Y. H. Lee. *Channel estimation via orthogonal matching pursuit for hybrid mimo systems in millimeter wave communications*. IEEE Transactions on Communications **64**(6), 2370 (2016).

[31] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath. *Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave mimo systems*. IEEE Transactions on Wireless Communications **17**(5), 2946 (2018).

[32] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath. *Channel estimation for hybrid architecture-based wideband millimeter wave systems*. IEEE Journal on Selected Areas in Communications **35**(9), 1996 (2017).

[33] D. E. Berraki, S. M. D. Armour, and A. R. Nix. *Application of compressive sensing in sparse spatial channel recovery for beamforming in mmwave outdoor systems*. In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 887–892 (2014).

[34] J. Choi. *Beam selection in mm-wave multiuser mimo systems using compressive sensing*. IEEE Transactions on Communications **63**(8), 2936 (2015).

[35] T. Bai and R. W. Heath. *Coverage in dense millimeter wave cellular networks*. In *2013 Asilomar Conference on Signals, Systems and Computers*, pp. 2062–2066 (2013).

[36] X. Song, S. Haghighatshoar, and G. Caire. *A scalable and statistically robust beam alignment technique for millimeter-wave systems*. IEEE Transactions on Wireless Communications **17**(7), 4792 (2018).

[37] E. J. Candes and T. Tao. *Near-optimal signal recovery from random projections: Universal encoding strategies?* IEEE Transactions on Information Theory **52**(12), 5406 (2006).

[38] M. Rani, S. B. Dhok, and R. B. Deshmukh. *A systematic review of compressive sensing: Concepts, implementations and applications*. IEEE Access **6**, 4875 (2018).

[39] C. Tsai and A. Wu. *Structured random compressed channel sensing for millimeter-wave large-scale antenna systems*. IEEE Transactions on Signal Processing **66**(19), 5096 (2018).

[40] J. Romberg. *Compressive sensing by random convolution*. SIAM Journal on Imaging Sciences **2**(4), 1098 (2009). https://doi.org/10.1137/08072975X, URL https://doi.org/10.1137/08072975X.

[41] Y. Shabara, C. E. Koksal, and E. Ekici. *Linear block coding for efficient beam discovery in millimeter wave communication networks*. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, pp. 2285–2293 (2018).

[42] A. Amini and F. Marvasti. *Deterministic construction of binary, bipolar, and ternary compressed sensing matrices*. IEEE Transactions on Information Theory **57**(4), 2360 (2011).

[43] A. Ravelomanantsoa, H. Rabah, and A. Rouane. *Compressed sensing: A simple deterministic measurement matrix and a fast recovery algorithm*. IEEE Transactions on Instrumentation and Measurement **64**(12), 3405 (2015).

[44] S. Jafarpour. *Deterministic compressed sensing*. Ph.D. thesis, Princeton University (2011).

[45] E. J. Candes and T. Tao. *Decoding by linear programming*. IEEE Transactions on Information Theory **51**(12), 4203 (2005).

[46] E. J. Candes, J. Romberg, and T. Tao. *Stable signal recovery from incomplete and inaccurate measurements*. Communications on Pure and Applied Mathematics **59**, 1207–1223 (2006).

[47] E. J. Candes and M. B. Wakin. *An introduction to compressive sampling*. IEEE Signal Processing Magazine **25**(2), 21 (2008).

[48] A. S. Bandeira, E. Dobriban, D. G. Mixon, and W. F. Sawin. *Certifying the restricted isometry property is hard*. IEEE Transactions on Information Theory **59**(6), 3448 (2013).

[49] R. Baraniuk, M. A. Davenport, M. F. Duarte, and C. Hegde. *An introduction to compressive sensing*. (2011). OpenStax-CNX. [Online], URL http://legacy.cnx.org/content/col11133/1.5/.

[50] T. T. Cai and L. Wang. *Orthogonal matching pursuit for sparse signal recovery with noise*. IEEE Transactions on Information Theory **57**(7), 4680 (2011).

[51] J. Bourgain, S. Dilworth, K. Ford, S. Konyagin, and D. Kutzarova. *Explicit constructions of rip matrices and related problems*. Duke Math. J. **159**(1), 145 (2011). URL https://doi.org/10.1215/00127094-1384809.

[52] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. *A simple proof of the restricted isometry property for random matrices*. Constructive Approximation **28**(3), 253 (2008). URL https://doi.org/10.1007/s00365-007-9003-x.

[53] D. L. Donoho. *Compressed sensing*. IEEE Transactions on Information Theory **52**(4), 1289 (2006).

[54] J. . Fuchs. *On sparse representations in arbitrary redundant bases*. IEEE Transactions on Information Theory **50**(6), 1341 (2004).

[55] E. J. Candes, J. Romberg, and T. Tao. *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*. IEEE Transactions on Information Theory **52**(2), 489 (2006).

[56] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee. *Compressive sensing: From theory to applications, a survey*. Journal of Communications and Networks **15**(5), 443 (2013).

[57] X. Song, T. Kühne, and G. Caire. *Fully-/partially-connected hybrid beamforming architectures for mmwave MU-MIMO*. CoRR **abs/1904.10276** (2019). 1904.10276, URL http://arxiv.org/abs/1904.10276.

[58] M. Majidzadeh, A. Moilanen, N. Tervo, H. Pennanen, A. Tölli, and M. Latva-aho. *Hybrid beamforming for single-user mimo with partially connected rf architecture*. In *2017 European Conference on Networks and Communications (EuCNC)*, pp. 1–6 (2017).

[59] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. *Millimeter wave mobile communications for 5g cellular: It will work!* IEEE Access **1**, 335 (2013).

[60] M. F. Duarte and R. G. Baraniuk. *Kronecker product matrices for compressive sensing*. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3650–3653 (2010).

[61] S. Jokar and V. Mehrmann. *Sparse solutions to underdetermined kronecker product systems*. Linear Algebra and its Applications **431**(12), 2437 (2009). Special Issue in honor of Shmuel Friedland, URL http://www.sciencedirect.com/science/article/pii/S0024379509004066.

[62] R. A. DeVore. *Deterministic constructions of compressed sensing matrices*. Journal of Complexity **23**(4), 918 (2007). Festschrift for the 60th Birthday of Henryk Woźniakowski, URL http://www.sciencedirect.com/science/article/pii/S0885064X07000623.

[63] J. Proakis and M. Salehi. *Digital Communications* (McGraw-Hill, 2008). URL https://books.google.com.au/books?id=ABSmAQAACAAJ.

[64] S. Haghighatshoar and G. Caire. *The beam alignment problem in mmwave wireless networks*. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pp. 741–745 (2016).