Compressed Beam Alignment with Out-of-Band Assistance in Millimeter Wave Cellular Networks

Jie Zhao, *Student Member, IEEE,* Xin Wang, *Member, IEEE,* Harish Viswanathan, *Fellow, IEEE* Arjuna Madanayake, *Member, IEEE*, and Guangxue Yue, *Member, IEEE.*

Abstract—Network transmission over millimeter-wave (mmW) bands has a big potential to provide orders of higher bandwidth. However, beamforming is generally needed to compensate for the high path loss. As mmW antennas have a potentially large number of candidate beamforming directions, to achieve high network throughput, the finding of a high gain direction between a base station and each mobile in the mmW network may involve a large overhead if training signals are directly sent along all possible directions or according to a large volume of codebook. Taking advantage of the block sparse characteristics of the mmW channel and coexistence of legacy antennas, we propose a comprehensive design for more efficient beam direction finding. Different from existing compressive-sensing-based schemes which just take a random subset of directions to measure, taking advantage of the path clustering feature of the mmW channel, we develop a self-adaptive block sparse algorithm which can benefit from preliminary channel estimation during each iteration of the problem solving to significantly improve the overall channel estimation accuracy thus the beam alignment gain. We also explore two methods to exploit co-located legacy antennas to provide further guidance for transmission direction finding. Simulation results indicate that our proposed beam alignment scheme outperforms the baseline and peer schemes in terms of the beamforming gain and training cost. By taking advantage of the block sparse properties of mmW channel, our proposed design is able to achieve the transmission throughput comparable with the exhaustive direction search at much lower overhead.

Index Terms—millimeter wave, beamforming, beam alignment, directional antenna, compressed sensing.

1 Introduction

Millimeter-wave (mmW or mmWave) communication is receiving tremendous interest from academia, industry and federal agencies as a promising technique to provide Gigabit data rate demanded by the exponential growth of various applications in wireless networks. A key challenge of data transmissions in mmW cellular networks is the low signal transmission range (Figure 1). According to Frii's Law, the high frequencies of mmW signals result in a large isotropic path loss.

Fortunately, the small wavelength of mmW signals also enables a large number of antenna elements to be placed in the same small dimension (e.g. at the base station, in the skin of a cellphone, or even within a chip), which provides a high beamforming gain that can compensate for the increase in the isotropic path loss.

Although the use of a large antenna array helps to combat the severe path loss, it also makes it difficult to coordinate network transmissions [1], [2], [3], [4]. Taking the initialization and synchronization of base stations (BSs) and mobiles as an example, using only omni-directional transmissions of synchronization signals would be problematic in the mmW range: the availability of high gain antennas would bring a discrepancy between the range at which a cell can be detected (when signaling messages are transmitted omni-directionally before the correct beamforming directions are found) and the range at which reasonable data rates can be achieved (after beamforming is applied). On the other hand, although a beamed transmission from the base station provides a larger footprint and allows for higher data rate, it is difficult for a mobile to find the base stations (BSs) initially without knowing the correct beamforming directions. To address these issues, a cell search phase is needed where the base station of a mmW network beams towards different directions to facilitate a mobile to find a direction that maximizes its receiving rate.

The antenna gains of the transmitter (TX) and the receiver (RX) have significant impacts on the transmission quality. Simply transmitting signaling messages rotationally along each direction would introduce very high delay and cost for finding the optimal beamforming direction with the maximum gain. An example is when TX and RX each has 64 beam directions (in practical mmW networks this number can be even larger), to exhaustively measure every beam pair, $64 \times 64 = 2^{12}$ measurements are required. The finding of an optimal beam direction may take long time to complete, resulting in a large delay to establish a transmission link. As the channel conditions are dynamic, the direction finding may need to be performed frequently, which would sig-

• *J. Zhao and X. Wang are with the Department of Electrical and Computer Engineering, Stony brook University, Stony Brook, NY 11794 (E-mail: jie.zhao@alumni.stonybrook.edu, x.wang@stonybrook.edu).* *X. Wang's research has been supported by the NSF under grant ECCS-1731238.* *H. Viswanathan is with Nokia Bell Labs, Murray Hill, NJ 07974 (E-mail: harish.viswanathan@alcatel-lucent.com). A. Madanayake is with the Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33174 (E-mail: amadanay@fiu.edu). G. Yue is with the College of Mathematical and Information Engineering, Jiaxing University, Zhejiang, China (E-mail: guangxueyue@163.com).*

nificantly compromise the network capacity. Therefore, it is critical to incorporate an efficient beam searching or beam alignment scheme into the MAC design to establish high quality mmW links at low cost.

As lower-frequency legacy radios are likely to be deployed alongside mmW systems to strengthen wide area signaling control and/or facilitate multi-band communications, out-of-band information from legacy band has been considered to help reduce the overhead of establishing the mmW links [5], [6]. As the co-located low-frequency system and mmW system share similar environment, the spatial characteristics of low-frequency channel and mmW channel have correlation, as confirmed by experimental studies. This however, does not mean the channels at different frequencies have exactly the same characteristics. The work in [5] simply models the channel with a single path without considering the characteristics of mmWave channel [7], [8], [9], while authors in [6] assume paths of legacy channel and mmWave channel have nearly-exact match. These inaccuracies will lead to mis-alignment of mmWave beams, which will further degrade the transmission performance in realistic scenarios. As another source of inaccuracy, the work in [5] exploits model-based estimation to directly infer the transmission directions on the mmWave channel based on measurements from the low-frequency channel, which will make the alignment subject to performance degradation due to difference in channels and failure under high dynamics.

The measurement from low band can only give a rough estimation of transmission direction. To enable high performance transmissions, we need to align the beams with the desired angular resolution. Instead of exhaustively training all possible beam pairs within a range directed by the low-band training, we are motivated to further reduce the number of measurements with compressed channel estimation. Different from the legacy low frequency channels, the recent studies [7], [8], [9], show that wireless mmW channels only have a few dominant paths and the paths are often clustered, thus the channels often present sparse characteristics. Some initial studies [10], [11], [12], [13] have been made to exploit the sparse feature to estimate mmW channel with compressed sensing [14], [15], but they did not take into account the clustering feature of transmission paths that differentiates mmW channel from conventional low frequency channels [16].

In this paper, we concurrently exploit legacy band coarse direction finding and compressed channel estimation to enable more accurate beam alignment between each pair of sender and receiver at low cost. The estimation of channel allows for finding the optimal fine beam direction even when it does not fall into the best coarse beam range of low-frequency channel. Rather than just exploiting the basic compressed sensing technique [7], [8], [9], [17], [18] based on the low rank properties of mmW channel, we propose two major techniques to significantly reduce the training overhead, improve channel estimation performance and increase the quality of beam alignment: 1) We develop a self-adaptive block sparse reconstruction algorithm which learns from the iterative channel estimation results to further improve the estimation efficiency; and 2) We design a beam alignment algorithm which concurrently exploits the coarse channel information from low-frequency antennas and block-based mmW channel estimation to enable high-throughput beamed transmissions. Our design intends to be flexible. The block-based channel estimation can work either stand-alone or jointly with out-of-band assistance.

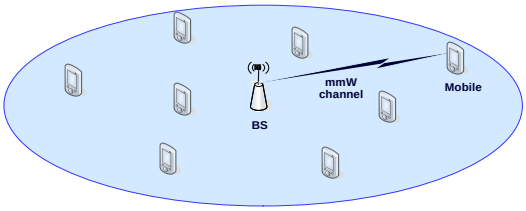

Fig. 1. Millimeter-wave cellular network.

The rest of this paper is organized as follows. After briefly reviewing related work in Section 2, we present compressed sensing preliminaries in Section 3. We then describe the system model and our motivation in Section 4. Block sparse mmW channel estimation is introduced in Section 5, followed by Section 6 that presents our channel reconstruction algorithm and beam alignment design. We analyze the simulation results in Section 7. The paper concludes in Section 8.

2 Related Work

In millimeter-wave (mmW or mmWave) networks, a high beamforming gain is needed to compensate for the large path loss and occlusions in the mmW spectrum range [19]. This requires a joint beamforming (BF) scheme in the MAC protocol to select the best transmission and reception beam directions according to a metric such as signal-to-noise ratio (SNR) [20].

IEEE 802.15.3c [21] and IEEE 802.11ad [22] standards are proposed to enable operation in the 60GHz mmWave band. Multi-level codebook is suggested in both protocols to facilitate the training of beams at hierarchical resolutions of beamwidth. Some other codebook-based beamforming methods are also proposed in [23], [24], where the beam training overhead highly relies on the codebook design thus search space. Although codebook-based schemes reduce the beam search space, the overhead for uplink feedback of beam conditions and selection would be big when the beamwidth is small and there are a large number of directions to search. Another drawback of multi-resolution codebook schemes is that the optimal fine beam direction may be filtered out during high-level coarse direction estimation, as fine beam direction may not exactly align with coarse beam direction. There are also no discussions in existing schemes on how to reduce the large overhead when the

codebook volume is huge, which is often the case in mmW band.

Some existing works attempt to reduce the beamforming overhead by intelligently reducing the search cope (e.g. select only part of the candidate beams to train). In [25], Hur *et al.* propose the use of outdoor millimeter wave communications for backhaul networking to connect cells with the core network and mobile access within a cell between the base station and mobiles. To overcome the outdoor impairments found in millimeter wave propagation, this paper studies beamforming using large arrays. The authors propose an efficient beam alignment technique using adaptive subspace sampling and hierarchical beam codebooks. To perform initial directional cell search in mmW cellular networks, the mobile and base station must jointly search over a potentially large angular directional space to locate a suitable path to initiate communication. Barati *et al.* in [26] propose a directional cell search procedure where a base station periodically transmits synchronization signals in randomly varying directions, and derive detectors for both analog beamforming and digital beamforming. An efficient beam switching technique for the emerging 60GHz wireless personal area networks is proposed in [27], where Li *et al.* formulate a global optimization problem to look for the best beam-pair for data transmissions and adopt a numerical approach to implement the beam searching strategy with divide and conquer used in small regions. Rather than just reducing the beam search space, we also take advantage of the mmW channel feature and the information of trained beams to estimate the mmW channel and guide for better beam selection.

Channel estimation for mmW beamforming is investigated in [28], where Singh *et al.* investigate the feasibility of employing multiple antenna arrays to obtain diversity/multiplexing gains in mmW systems. The authors exploit the sparse multipath property of the mmW channel and propose to reduce the complexity involved in jointly optimizing the beamforming directions across multiple arrays by focusing on a small set of candidate directions. Differently in our work, we not only take advantage of the sparsity of mmW channel, but also employ a useful tool, compressed sensing, to fully extract useful information from mmW channels. In [29], Olfat *et al.* propose to estimate the channel for a frequency-selective millimeter-wave communication system with a minimum number of pilots. A learning-based scheme is taken to find the optimal precoding and combining vectors for transmitting and receiving pilot signals in the face of channel dynamics, where the learning requires the previous knowledge of the channel.

Compressed sensing [14], [15] (CS) has been exploited for channel estimation in [10], [11], [12], [13], [30]. Although these studies exploit the sparse feature of mmW channel, they did not consider the clustering feature of transmission paths, and the path clustering is a distinct characteristic of mmW channel compared to conventional low-frequency channels [16]. In this work, rather than just exploiting the basic low rank properties [7], [8], [9] of the mmW channel based on compressed sensing techniques, we propose a block-sparse channel estimation algorithm that takes full advantage of the path clustering to significantly improve the channel estimation efficiency. In addition, we also investigate the possibility of exploiting the information from co-located legacy antennas (e.g. from 3G cellular networks) to further improve the performances of mmW systems. One major difference of our design from other methodologies on out-of-band assistance for mmW networks [5], [6] is that we take the legacy band information as an optional procedure that can jointly work with mmW channel estimation to enable low cost and high accuracy alignment at the desired beam resolution. These literature studies usually rely on the out-of-band assistance to operate stand-alone and have strong assumptions on the mmWave channel, such as having only single dominating path [5] or having paths matched with those on the legacy channel almost perfectly [6]. These assumptions make the beam alignment more prone to inaccuracy in practical scenarios. In addition, model-based estimation [6] based on coarse measurements from low frequency band can not well capture the channel dynamics at mmWave band to find the fine beam directions with the highest gains.

User motion can make the beamforming issue more challenging in mmW networks. Based on the observation that 60 GHz channel profiles at nearby locations are highly-correlated, Zhou *et al.* in [31] propose a beam-forecast scheme to reconstruct the channel profile and predict new optimal beams. The beam prediction during mobility case is built upon the initial channel estimation and direction finding, and channel scanning will be called for again to realign beams as the prediction errors accumulate over time. Complementary to the work in [31], we propose an efficient scheme to find the optimal beam direction without pre-channel knowledge.

To summarize our differences from existing literature, in this work, we propose to design an intelligent and efficient beam alignment scheme to enable the quick finding of a good transmission direction between a BS and a mobile in mmW cellular networks. Rather than simply and exhaustively measuring all possible beam directions or searching in a large codebook, we can measure a small portion of the beam pairs to achieve comparable performance, thus saving resources like time and power. Specifically, to enable beam alignment with lower training overhead and higher beamforming performances, we propose an efficient CS-based mmW channel estimation methodology that takes full advantage of the block sparse features (due to path clustering) of mmW channel, and the channel estimated serves as a guidance to discover the optimal beam pairs. We further improve the beam alignment performance by gathering information from co-located legacy antennas. Some important issues we consider include: (a) Why does a mmW channel have block-sparse properties? (b) How to take advantage of

the block-sparse properties of mmW channel to perform more efficient channel estimation in order to guide the beam alignment? (c) How to comprehensively design beam alignment schemes to achieve a better network performances? (d) How to exploit the information from legacy antennas?

To answer these questions, we'll first present the system model in the next section.

3 Compressed Sensing Preliminaries

In this work, we will exploit the use of the emerging compressed sensing (CS) technique for more efficient beam alignment in mmWave cellular networks. We first introduce some backgrounds on compressed sensing.

The main idea of compressed sensing (CS) is to take advantage of the sparsity within the signal to significantly reduce the sampling rate. An N-dimensional signal $\mathbf{d}$ is considered to be K-sparse in a domain (also called a dictionary matrix) $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ if there exists an N-dimensional vector $\mathbf{x} \in \mathbb{R}^{N\times 1}$ so that $\mathbf{d} = \mathbf{\Psi x}$ and $\mathbf{x}$ has at most K non-zero entries ($K \ll N$). The CS theory suggests that $\mathbf{d}$ can be fully reconstructed from a sufficient number M ($M \geq cK\log\left(\frac{N}{K}\right)$, where c is a fairly small constant) of linear measurements.

If one performs linear measurements of the signal $\mathbf{d}$ with a measurement matrix $\mathbf{\Phi}$, then one can consider the obtained linear measurements $\mathbf{y}$, possibly affected by noise as:

$$\mathbf{y} = \mathbf{\Phi\Psi x} + \mathbf{n} = \mathbf{Ax} + \mathbf{n}, \tag{1}$$

where the measurements are $\mathbf{y} \in \mathbb{R}^{M\times 1}$, the sparse vector $\mathbf{x} \in \mathbb{R}^{N\times 1}$, the additive noise $\mathbf{n} \in \mathbb{R}^{M\times 1}$, the sensing matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$, and $M < N$. $\mathbf{A}$ is essentially the product of the measurement matrix and the dictionary matrix: $\mathbf{A} = \mathbf{\Phi\Psi}$, where $\mathbf{\Phi} \in \mathbb{C}^{M\times N}$, $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$. We notice that different from the notations above, a small number existing works call $\mathbf{\Phi}$ the sensing matrix. In order to avoid inconsistency and misunderstanding, we will consistently regard $\mathbf{\Phi}$ as the measurement matrix, and $\mathbf{A} = \mathbf{\Phi\Psi}$ the sensing matrix.

Obviously, the number of measurements is smaller than the number of variables in Equation 1, and it indicates this is an under-determined system. Donoho, Candes, Romberg, and Tao show in [32] that the under-determined equation system can be solved as:

$$\min \quad \|\mathbf{x}\|_{l_1} \tag{2}$$

$$\text{s.t.} \quad \|\mathbf{\Phi d} - \mathbf{y}\|_{l_2} \leq \epsilon \tag{3}$$

$$\mathbf{d} = \mathbf{\Psi x} \tag{4}$$

where the parameter ϵ is the bound of the error caused by noise $\mathbf{n}$, l_p means the l_p-norm ($p = 1, 2, ...$). The solution can also be expressed as:

$$\hat{\mathbf{x}} = \underset{\mathbf{u}:\ \|\mathbf{y}-\mathbf{Au}\|_{l_2}\leq\epsilon}{\arg\min} \|\mathbf{u}\|_{l_1}. \tag{5}$$

The signal $\mathbf{d} = \mathbf{\Psi x}$ can then be recovered as $\hat{\mathbf{d}} = \mathbf{\Psi}\hat{\mathbf{x}}$.

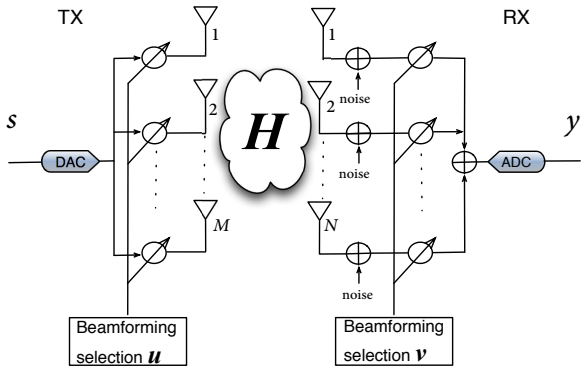

Fig. 2. Beam alignment between TX and RX.

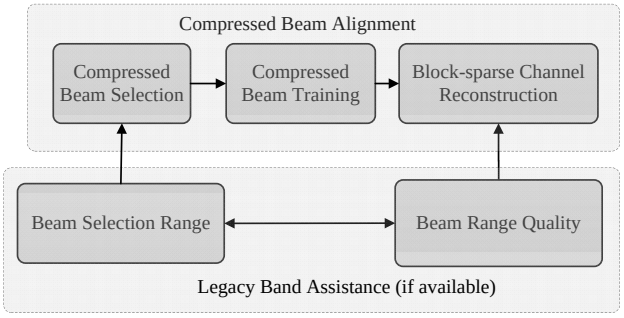

Fig. 3. Framework Overview.

The form of the optimization problem in (5) is known as LASSO [33] or BPDN [34] and also some other variations such as the Dantzig selector [35]. In addition to the convex optimization approach, such as l_1 minimization [36], to reconstruction in compressed sensing, there exist several iterative/greedy algorithms such as IHT [37] and Cosamp [38]. Such convex or greedy approaches are generally called reconstruction algorithms.

4 System Model and Motivation

In order to perform the beam alignment in mmWave cellular networks, a challenge is that the base stations and mobiles need to search for an optimal transmission direction for each transmission pair over a large number of possible beam directions. To frame the problem and the design of our beam-alignment algorithm, we first introduce the basic system model considered in this paper.

4.1 Framework Overview

As shown in Figure 3, our proposed framework consists of two major components: (1) Compressed Beam Alignment and (2) Legacy Band Assistance. The two components can jointly work together. Compressed beam alignment can operate stand-alone if there is no assistance from the legacy band. Our low-cost beam alignment is achieved by reducing the beam search overhead with Compressed Beam Selection and Training, and guided through channel information obtained from Block-Sparse Channel Estimation. Different from the literature work,

block sparsity of the mmW channel is exploited in the modeling and reconstruction of the channel, which improves the accuracy of channel estimation and thus enhances the performance of beam alignment. The impacts of Legacy Band Assistance on Compressed Beam Alignment lie in two aspects: (a) beam selection range suggestion and (b) beam range quality suggestion, where the former helps narrow the angular range of beam search thus reducing the overhead and the channel reconstruction in the latter takes advantage of the suggested beam range qualities to achieve higher accuracy.

4.2 Notations

For a matrix $\mathbf{A}$, we use $\mathbf{A}^T$ to denote its transpose. $\mathbf{A}^*$, $\mathbf{A}^H$, $\mathbf{A}^{-1}$ are its conjugate, Hermitian transpose (conjugate transpose) and inverse, respectively. For a vector $\mathbf{a}$, $\mathrm{diag}(\mathbf{a})$ is a diagonal matrix whose diagonal elements are entries from $\mathbf{a}$. $\mathbf{AB}$ is the matrix product of $\mathbf{A}$ and $\mathbf{B}$. $\mathbf{A} \circ \mathbf{B}$ is the Hadamard product (element-wise product) of $\mathbf{A}$ and $\mathbf{B}$. $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. $\mathbf{A} * \mathbf{B}$ is the Khatri-Rao product of $\mathbf{A}$ and $\mathbf{B}$. $\mathrm{diag}(\mathbf{a})$ is the diagonal matrix of vector $\mathbf{a}$ whose diagonal elements of $\mathrm{diag}(\mathbf{a})$ are entries from $\mathbf{a}$. $\mathrm{vec}(\mathbf{A})$ is vectorized vector of matrix $\mathbf{A}$, and the vec operator creates a column vector from a matrix $\mathbf{A}$ by stacking the column vectors of $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n]$ below one another.

4.3 Channel Model and Downlink Transmission

We will now explain the model using the downlink transmission as an example, where the RX denotes the mobile and the TX means the base station (BS). For a sender and receiver pair, assume there are N_{rx} RX antennas and N_{tx} TX antennas.

The signal transmitted from TX is represented as

$$\mathbf{x}^{N_{tx}\times 1} = \mathbf{u}^{N_{tx}\times 1} \circ \mathbf{s}^{N_{tx}\times 1} \tag{6}$$

where $\mathbf{u}$ is the BF weight vector for the TX antennas and $\mathbf{s}$ is the training signal sent from the TX antennas. The multiplication $\circ$ means the operation between the vectors is element-wise.

The signal received by the RX antennas before the receiver beamforming is

$$\mathbf{y}^{N_{rx}\times 1} = \mathbf{H}^{N_{rx}\times N_{tx}}\mathbf{x}^{N_{tx}\times 1} + \mathbf{e}^{N_{rx}\times 1} \tag{7}$$

where $\mathbf{H}$ is the channel matrix (its element $H_{n,m}$ is the channel coefficient from the TX antenna m to the RX antenna n), $\mathbf{e}$ is AWGN at the RX antennas.

In [16], statistical models of mmW channels are derived from real-world measurements at 28 and 73 GHz in New York City (NYC). The mmW channel is found to be sparse and can be modeled as a small number of clusters each consisting of a small number of subpaths. For ease of presentation, we consider only the azimuth and neglect the elevation in this paper, implying that all scattering happens in the azimuth and that the TX and RX conduct the horizontal beamforming only. Implementations that facilitate both horizontal and vertical beamforming can be built on top of our design. While our proposed design can be used for any kind of antenna arrays, without loss of generality, we adopt uniform linear arrays (ULAs) in this work. Assuming a mmW channel is composed of K clusters within each there are L subpaths, then the mmW channel matrix we consider can be expressed as:

$$\mathbf{H} = \sum_{k=1}^{K}\sum_{\ell=1}^{L} a_{k\ell} \cdot \mathbf{D}_{rx}(\theta_{k\ell}^{rx}) \cdot \mathbf{D}_{tx}^{H}(\theta_{k\ell}^{tx}), \tag{8}$$

where $a_{k\ell}$ is the complex path gain for a path ℓ ($\ell = 1, 2, ..., L$) in the cluster k ($k = 1, 2, ..., K$), with $k\ell$ jointly corresponding to the ℓ-th sub-path in the k-th cluster. For the sake of consistency, in this work, we use the terms path and sub-path interchangeably. $\theta_{k\ell}^{tx}$ and $\theta_{k\ell}^{rx}$ denote the angle of departure (AoD) and the angle of arrival (AoA) for the corresponding path, which throughout this paper indicates the horizontal scattering.

$\mathbf{D}_{tx}(\theta_{k\ell}^{tx})$, the TX antenna's directional response column vector ($N_{tx} \times 1$ dimension) for the sub-path at the angle of departure $\theta_{k\ell}^{tx}$, is expressed as:

$$\begin{aligned}&\mathbf{D}_{tx}(\theta_{k\ell}^{tx})\\ &= \left[D^{(1)}(\theta_{k\ell}^{tx}), D^{(2)}(\theta_{k\ell}^{tx}), ..., D^{(m)}(\theta_{k\ell}^{tx}), ..., D^{(N_{tx})}(\theta_{k\ell}^{tx})\right]\\ &= \left[1, e^{j\cdot 1\cdot w_{k\ell}^{tx}}, e^{j\cdot 2\cdot w_{k\ell}^{tx}}, ..., e^{j\cdot (N_{tx}-1)\cdot w_{k\ell}^{tx}}\right]^T,\end{aligned} \tag{9}$$

where $D^{(m)}(\theta_{k\ell}^{tx})$ is from antenna basics, the spatial frequency $w_{k\ell}^{tx}$ can be written in terms of AoDs, as $w_{k\ell}^{tx} = \frac{2\pi d_t}{\lambda}\sin\theta_{k\ell}^{tx}$. d_t is the distances between two adjacent antenna elements in the ULAs in the TX. $\lambda = \frac{c}{f}$ is wavelength in meters. f is the carrier frequency of the signal in Hz, c is the speed of light (3×10^8 meters/sec).

Similarly, $\mathbf{D}_{rx}(\theta_{k\ell}^{rx})$, the RX antenna's directional response column vector ($N_{rx} \times 1$ dimension) for the path at an angle of arrival $\theta_{k\ell}^{rx}$, is expressed as

$$\begin{aligned}&\mathbf{D}_{rx}(\theta_{k\ell}^{rx})\\ &= \left[D^{(1)}(\theta_{k\ell}^{rx}), D^{(2)}(\theta_{k\ell}^{rx}), ..., D^{(n)}(\theta_{k\ell}^{tx}), ..., D^{(N_{rx})}(\theta_{k\ell}^{tx})\right]\\ &= \left[1, e^{j\cdot 1\cdot w_{k\ell}^{rx}}, e^{j\cdot 2\cdot w_{k\ell}^{rx}}, ..., e^{j\cdot (N_{rx}-1)\cdot w_{k\ell}^{rx}}\right]^T,\end{aligned} \tag{10}$$

where $D^{(n)}(\theta_{k\ell}^{tx})$ is from antenna basics, the spatial frequency $w_{k\ell}^{rx}$ can be written in terms of AoAs, as $w_{k\ell}^{rx} = \frac{2\pi d_r}{\lambda}\sin\theta_{k\ell}^{rx}$. d_r is the distances between two adjacent antenna elements in the ULAs in the RX.

We now use a concatenated column vector $\mathbf{a}$ ($1 \times KL$) to denote the complex path gains. Then

$$\mathbf{a} = [\underbrace{a_{11}, a_{12}, ..., a_{1L}}_{\text{cluster 1}}, \underbrace{a_{21}, a_{22}, ..., a_{2L}}_{\text{cluster 2}}, ..., \underbrace{a_{K1}, a_{K2}, ..., a_{KL}}_{\text{cluster } K}]^T. \tag{11}$$

Note $\mathbf{a}$ is concatenated in a manner that the first L elements are for the first cluster, and the next L elements

are for the second cluster and so on, and there are KL elements in $\mathbf{a}$. The major task of mmW channel estimation in our work is to estimate $\mathbf{a}$ efficiently.

In order to achieve a high receiving gain, a beamforming vector will be applied at RX. The beamforming output at the RX antenna array will be

$$r = \mathbf{v}^H \mathbf{y} + e = \mathbf{v}^H \mathbf{H}\mathbf{u} \circ \mathbf{s} + e, \tag{12}$$

where e is the received noise at the RX. The beamforming process can be presented as in the Figure 2.

We have introduced the model for the downlink transmission, where TX is BS and RX is MS. The uplink communication can be easily derived according to our presentation above by making MS the TX and BS the RX.

4.4 Motivation and Problem

A grand challenge in mmW communications is to efficiently find the optimal beamforming directions for data transmissions in the mmW networks. This requires determining the optimal weights for both the transmitter and the receiver of each transmission pair in the mmW network. To compensate for the path loss in the high frequency mmWave channel, TX and RX potentially need a large number of antennas to achieve a high beamforming gain, which unfortunately also greatly expands the search space of TX and RX antenna directions.

As an example, when TX and RX each has 64 beam directions (in practical mmW networks, this number can be even larger), to exhaustively measure every beam pair, $64 \times 64 = 2^{12}$ measurements are required. Simply transmitting signaling messages rotationally along each direction as suggested by the existing standard [21] would introduce very high delay and cost for finding the optimal beamforming direction with the maximum gain. Therefore, an essentially important question for mmW band beamforming is: how to reduce the search space of TX and RX beam directions in mmW cellular networks while ensuring the high beamforming gain thus transmission rate?

Existing work such as [10], [11], [12], [13] applied compressed sensing (CS) to alleviate this training overhead by randomly training a small subset of beam pairs and then estimating the mmW channel to facilitate beam alignment. However, these studies did not fully exploit the path clustering effect of mmW channels and the resulted block sparsity of the gain coefficients of mmW paths.

Moreover, the small angular spread in the mmW channel indicates that if some directions in the low-frequency channel are dominant, it is also likely that an overlapped mmW channel has a good gain in surrounding angular space. Although transmission paths in the mmW channel and the low-frequency channel may not be exactly identical, the information in the low-frequency channel can give a rough information on which angular range of directions is more likely to have better path gains in the mmW channel. In the existence of some co-located legacy antennas (e.g. from 3G cellular networks) around mmW antennas, more efficient beam alignment schemes can be designed by taking advantage of the information from legacy band. Different from the literature on CS-based mmW channel estimation and beamforming, to support more efficient beam alignment, we aim to answer the following questions in this work:

(a) How to efficiently exploit the block sparse feature of mmW channels to better formulate the mmW channel estimation problem?

(b) How to develop an efficient block sparse CS reconstruction algorithm to achieve more accurate mmW channel estimation?

(c) How can the information from co-located legacy antennas provide more useful guidance for mmW beam alignment?

In response of (a), Section 5 presents a block-sparsity-based mmW channel estimation problem. The solutions for (b) and (c) are presented in Section 6. We propose a self-adaptive weighted algorithm to iteratively learn the weights of channel blocks to increase the CS reconstruction performance. We also develop two methodologies which exploit legacy antennas to facilitate beam alignment, one aiming to help select better beams to be trained and the other aiming to further improve the reconstruction performance of mmW channels.

5 Sparse MmWave Channel Estimation

For each pair of transmitter and receiver in the mmW network, in order to find the best transmission/reception direction, there is a need to estimate the channel between them. This is generally facilitated with the sending of the training signals from the transmitter to the receiver. However, if the channel measurement is performed in a straight-forward way, it would require the transmissions along a large number of directions to find more accurate channel information. In this section, we first discuss the channel features, and then explain how the emerging compressed sensing technique can be applied to reduce the channel measurement overhead.

5.1 Sparse Channel Estimation

From matrix basics we know that the channel matrix $\mathbf{H}$ from (8) can be equivalently written as follows:

$$\mathbf{H} = \mathbf{D}_R \,\mathrm{diag}(\mathbf{a}) \mathbf{D}_T^H \tag{13}$$

where $\mathrm{diag}(\mathbf{a})$ is a diagonal matrix whose diagonal elements are those from $\mathbf{a}$ in order. The matrices $\mathbf{D}_T$ and $\mathbf{D}_R$ contain the TX and RX array response vectors, respectively, in the following form:

$$\begin{aligned}\mathbf{D}_T = [&\mathbf{D}_{tx}(\theta_{11}^{tx}), \mathbf{D}_{tx}(\theta_{12}^{tx}), ..., \mathbf{D}_{tx}(\theta_{1L}^{tx}),\\ &\mathbf{D}_{tx}(\theta_{21}^{tx}), ..., \mathbf{D}_{tx}(\theta_{2L}^{tx}), ...,\\ &\mathbf{D}_{tx}(\theta_{K1}^{tx}), ..., \mathbf{D}_{tx}(\theta_{KL}^{tx})],\end{aligned} \tag{14}$$

$$\begin{aligned}\mathbf{D}_R = [&\mathbf{D}_{rx}(\theta_{11}^{rx}), \mathbf{D}_{rx}(\theta_{12}^{rx}), ..., \mathbf{D}_{rx}(\theta_{1L}^{rx}),\\ &\mathbf{D}_{rx}(\theta_{21}^{rx}), ..., \mathbf{D}_{rx}(\theta_{2L}^{rx}), ...,\\ &\mathbf{D}_{rx}(\theta_{K1}^{rx}), ..., \mathbf{D}_{rx}(\theta_{KL}^{rx})],\end{aligned} \tag{15}$$

The dimensions of matrices $\mathbf{D}_R$, $\mathrm{diag}(\mathbf{a})$ and $\mathbf{D}_T$ are $N_{rx} \times KL$, $KL \times KL$ and $N_{tx} \times KL$, respectively, which corresponds to the fact that $\mathbf{H}$ both (8) and (13) is an $N_{rx} \times N_{tx}$ matrix.

If we take measurements by transmitting the training signals along P directions and letting the receiver to estimate the signals from Q directions, which is achieved through the use of P TX beamforming (BF) vectors ($\mathbf{u}_p$, $p = 1, 2, ..., P$) and Q RX beamforming vectors ($\mathbf{v}_q$, $q = 1, 2, ..., Q$), we have the matrix form

$$\mathbf{R}^{Q \times P} = \mathbf{V}^H \mathbf{H} \mathbf{U} \circ \mathbf{S} + \mathbf{E}, \tag{16}$$

where an element in the location (q, p) in $\mathbf{R}$ denotes the BF output when TX and RX use beamforming weights $\mathbf{u}_p$ and $\mathbf{v}_q$, respectively, $\mathbf{S}$ and $\mathbf{E}$ are respectively the training signals and noise in the matrix form, and

$$\mathbf{V}^{N_{rx} \times Q} = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_q, ..., \mathbf{v}_Q], \tag{17}$$

$$\mathbf{U}^{N_{tx} \times P} = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p, ..., \mathbf{u}_P]. \tag{18}$$

With the training signals transmitted at the power A, we have

$$\mathbf{R}^{Q \times P} = \sqrt{A} \mathbf{V}^H \mathbf{H} \mathbf{U} + \mathbf{E} \tag{19}$$

We can vectorize $\mathbf{R}$,

$$\begin{aligned} \mathbf{r} = \mathrm{vec}(\mathbf{R}) &= \sqrt{A} \mathrm{vec}(\mathbf{V}^H \mathbf{H} \mathbf{U}) + \mathrm{vec}(\mathbf{E}) \\ &\overset{\text{Theorem 1 [39]}}{=\!=\!=} \sqrt{A} (\mathbf{U}^T \otimes \mathbf{V}^H) \mathrm{vec}(\mathbf{H}) + \mathrm{vec}(\mathbf{E}) \\ &\overset{\text{Proposition 1 [40]}}{=\!=\!=} \sqrt{A} (\mathbf{U}^T \otimes \mathbf{V}^H) \mathbf{\Psi} \mathbf{a} + \mathrm{vec}(\mathbf{E}) \\ &= \mathbf{\Phi} \mathbf{\Psi} \mathbf{a} + \mathrm{vec}(\mathbf{E}) = \mathbf{A} \mathbf{a} + \mathrm{vec}(\mathbf{E}), \end{aligned} \tag{20}$$

where $\mathbf{\Psi} = \mathbf{D}_T^* * \mathbf{D}_R$ (Khatri-Rao product), $\mathbf{\Phi} = \sqrt{A}(\mathbf{U}^T \otimes \mathbf{V}^H)$ (Kronecker product) is the measurement matrix (determined by the TX and RX beamforming training directions), $\mathbf{\Psi}$ is the basis matrix that will be introduced in Proposition 1. In the derivation, we have used Theorem 1 [39] and Proposition 1 [40]. For the integrity of presentation, we still provide the proofs below.

Theorem 1. $\mathrm{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{X})$.

Proof. Let $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n]$ (of size $m \times n$) and $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m]$. Then, the k-th column of $\mathbf{AXB}$ is

$$\begin{aligned} (\mathbf{AXB})_{:,k} &= \mathbf{AXb}_k = \mathbf{A} \sum_{i=1}^{m} \mathbf{x}_i b_{i,k} \\ &= [b_{1,k}\mathbf{A}, b_{2,k}\mathbf{A}, ..., b_{m,k}\mathbf{A}] \underbrace{[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m]^T}_{\mathrm{vec}(\mathbf{X})} \\ &= ([b_{1,k}, b_{2,k}, ..., b_{m,k}] \otimes \mathbf{A})\mathrm{vec}(\mathbf{X}) \end{aligned} \tag{21}$$

Stacking the columns together

$$\begin{aligned} \mathrm{vec}(\mathbf{AXB}) &= [(\mathbf{AXB})_{:,1}, \mathbf{AXB})_{:,2}, .., \mathbf{AXB})_{:,n},]^T \\ &= \left[\mathbf{b}_1^T \otimes \mathbf{A}, \mathbf{b}_2^T \otimes \mathbf{A}, ..., \mathbf{b}_n^T \otimes \mathbf{A}\right]^T \mathrm{vec}(\mathbf{X}) \\ &= (\mathbf{B}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{X}) \end{aligned} \tag{22}$$

□

Proposition 1. $\mathrm{vec}(\mathbf{H}) = \mathbf{\Psi}\mathbf{a}$, *where* $\mathbf{\Psi} = \mathbf{D}_R^* * \mathbf{D}_T$ *(Khatri-Rao product).*

Proof. From (13) we know we can estimate channel as follows:

$$\mathbf{H} = \mathbf{D}_R \, \mathrm{diag}(\mathbf{a}) \mathbf{D}_T^H \tag{23}$$

From Theorem 1 we know,

$$\begin{aligned} \mathrm{vec}(\mathbf{H}) &= (\mathbf{D}_T^* \otimes \mathbf{D}_R)\mathrm{vec}(\mathrm{diag}(\mathbf{a})) \\ &= (\mathbf{D}_T^* \otimes \mathbf{D}_R)\mathbf{J}\mathbf{a} \\ &\overset{\text{(a)}}{=\!=} (\mathbf{D}_T^* * \mathbf{D}_R)\mathbf{a} = \mathbf{\Psi}\mathbf{a} \end{aligned} \tag{24}$$

where $\mathbf{\Psi} = \mathbf{D}_T^* * \mathbf{D}_R$ (Khatri-Rao product) defines a basis domain where we can map the channel to. $\mathbf{J}$ is a $K^2L^2 \times KL$ selection matrix, which selects diagnoal elements from $\mathrm{diag}(\mathbf{a})$ to form $\mathbf{a}$. (a) is derived from the relationship between Kronecker product and Khatri-Rao product.

□

To facilitate the channel estimation, we can discretize the channel using an angular grid with the size of $G_{tx} \times G_{rx}$. Then the spatial frequencies $w_{k\ell}^{tx}$ (related to AoDs) and $w_{k\ell}^{rx}$ (related to AoAs) in (9) and (10) can be taken from the uniform grid of G_{tx} and G_{rx} points:

$$\mathbf{D}_T \rightarrow \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j2\pi\frac{1}{G_{tx}}} & \cdots & e^{j2\pi\frac{(G_{tx}-1)}{G_{tx}}} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & e^{j2\pi\frac{1}{G_{tx}}(N_{tx}-1)} & \cdots & e^{j2\pi\frac{(G_{tx}-1)}{G_{tx}}(N_{tx}-1)} \end{bmatrix}. \tag{25}$$

$$\mathbf{D}_R \rightarrow \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j2\pi\frac{1}{G_{rx}}} & \cdots & e^{j2\pi\frac{(G_{rx}-1)}{G_{rx}}} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & e^{j2\pi\frac{1}{G_{rx}}(N_{rx}-1)} & \cdots & e^{j2\pi\frac{(G_{rx}-1)}{G_{rx}}(N_{rx}-1)} \end{bmatrix}. \tag{26}$$

From Equation 24, the channel can be estimated as a vector of the dimension $G_{tx}G_{rx} \times 1$ ($\mathrm{vec}(\mathbf{H})$). In order to differentiate the estimated channel and the actual channel $\mathbf{a}$, we refer the estimated $\mathbf{a}$ as $\mathbf{x}$. For ease of presentation, we will also refer the channel as the estimated path gains of $\mathbf{x}$.

Replacing the vector $\mathbf{a}$ in the equation 20 with $\mathbf{x}$, we have the compressed sensing form $\mathbf{r} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{r}$ is the measurement results, $\mathbf{A}$ is the sensing matrix, $\mathbf{x}$ is a sparse vector to be reconstructed, and $\mathbf{e}$ is the noise. After finding $\mathbf{a}$ with the estimation of $\mathbf{x}$, $\mathbf{H}$ can be estimated from (23).

G_{tx} and G_{rx} determine the angular discretization levels of the channel, which impact the accuracy of

estimating the angle of departure (AoD) and angle of arrival(AoA) (whose values are continuous in reality on each transmission path). They also impact the dimensions of the basis domain (or the size of the reconstruction dictionary) in the channel reconstruction. The larger the values of G_{tx} and G_{rx}, the finer the resolution of the channel is discretized at, the more computational complexity for the channel reconstruction, and the closer the estimated AoD grid and AoA grid are to their actual angles. G_{tx} and G_{rx} can be chosen according to the desired estimation accuracy, conditions of the equipment such as antenna hardware precision and the complexity in the channel estimation when the dimension is large. If $G_{tx} = N_{tx}$ and $G_{rx} = N_{rx}$, we have

$$\mathbf{\Psi} = IDFT^*_{N_{tx}} * IDFT_{N_{rx}}, \tag{27}$$

where $IDFT_N$ denotes an N-dimensional IDFT matrix.

5.2 Sparse Virtual Channel Recovery

The sparseness of mmW channel can be exploited to reduce the number of beamforming pairs to train. To enable the optimal beam alignment, the virtual mmW channel can be more efficiently reconstructed with the training signals sent over a small number of directions. From (20), we know $\mathbf{x}$ (whose elements are the path gain coefficients being vectorized) can be recovered using CS theory, with the following CS components:

1) Basis matrix: $\mathbf{\Psi} = \mathbf{D}^*_R * \mathbf{D}_T$ (Khatri-Rao product);

2) Measurement matrix: $\mathbf{\Phi} = \sqrt{A}(\mathbf{U}^T \otimes \mathbf{V}^H)$ (Kronecker product);

3) Sensing matrix: $\mathbf{A} = \mathbf{\Phi\Psi} = \sqrt{A}(\mathbf{U}^T \otimes \mathbf{V}^H)\mathbf{D}^*_R * \mathbf{D}_T$;

The channel can be estimated by solving the following optimization problem from Equation (20):

$$\min \quad \|\mathbf{x}\|_1, \tag{28}$$

$$\text{subject to} \quad \mathbf{r} = \mathbf{Ax} + \mathbf{e}, \tag{29}$$

where $\|\cdot\|_1$ denotes the ℓ_1-norm, $\mathbf{a}$ is the noise.

To be consistent with the compressed sensing notations introduced earlier, we now use the following notations:

$$\min \quad \|\mathbf{x}\|_1, \tag{30}$$

$$\text{subject to} \quad \mathbf{y} = \mathbf{Ax} + \mathbf{n}, \tag{31}$$

where $\|\cdot\|_1$ denotes the ℓ_1-norm. Note the measurements $\mathbf{r}$ is denoted as $\mathbf{y}$ now.

After recovering $\mathbf{x}$, the virtual channel $\mathbf{H}$ can be estimated as in Equation (13). Existing recovery algorithms of the compressed sensing, such as l_1 minimization [36] and Cosamp [38], can be applied to recover $\mathbf{x}$. However, from the compressed sensing theory, we know that directly recovering vector $\mathbf{x}$ from PQ number of measurements in $\mathbf{y}$ may not be accurate if PQ is not big enough.

6 Block Sparse Channel Reconstruction and Beam-Alignment Design

In the basic channel model we consider, there are K clusters for transmission paths, each containing L paths. K_d out of K clusters are dominating, i.e., concentrating most of the signal power. The path clustering brings the block sparsity characteristics to the coefficient vector $\mathbf{a}$ in (24) for the gain of the path estimated. That is, not only that $\mathbf{a}$ is sparse with only a small fraction of elements non-zero, but also the non-zero elements are clustered. In simulations, we adopt the statistical mmW channel model developed based on real-world measurements in New York City [16] and also select the channel parameters based on empirical values suggested by the NYC model.

In order to improve the recovery performance, we propose to exploit the block property of the sparse vector $\mathbf{x}$ in the reconstruction problem in (30), which translates to the path gain vector $\mathbf{a}$ in Section 5, with the following design goals:

1) Reducing the number of measurements required for the recovery, i.e., reducing the number of training beam pairs (P and Q in (16));
2) Improving the channel reconstruction accuracy for a given number of measurements.

6.1 Block Sparse Property of MmW Channel

In CS-based channel estimation, the vector of the path gain $\mathbf{a}$ of mmW channel is what we are most interested in, where only a small fraction of the elements are non-zero. One of the effects of the path clustering feature of mmW channels is that it results in assembled non-zero elements in the path gain vector (Observation 1) and this additional block sparsity information can be further exploited in CS reconstruction to improve the reconstruction performance.

Observation 1. *The vector of the path gain* $\mathbf{a}$ *of mmW channel presents block sparsity due to the path clustering.*

Remarks: In [16], spatial statistical models of mmW channels are derived from real-world measurements at 28 and 73 GHz in New York City. It indicates that in micro-cell level, receivers in typical measurement locations experience a small number of path clusters, two to three being dominant. Moreover, within each path cluster, the angular spread is relatively small. The covariance matrix of the mmW channel is low-rank in the sense that the paths are clustering into relatively small and narrow beam clusters. The authors in [16] also studied the distribution of energy fraction in spatial directions and the results show that for 28GHz NYC channel, 3 dimensions of spatial directions can capture 95% of the channel energy for a 4×4 uniform planar array (which has a dimension of 16).

6.2 Block-Sparse Channel State Reconstruction

In our framework, the channel states to be recovered by CS (path gains) exhibit the characteristics of the block sparsity, as shown in Figure 4, which motivates us to use the block sparse properties to help reconstruct the channel signal for better beam-alignment.

6.2.1 Weighted Recovery

To exploit the block structure, instead of solving an ℓ_1 optimization problem in (2), we instead solve the following problem:

$$\begin{aligned} \text{min} \quad & \sum_{i=1}^{n} \|\mathbf{X}_i\|_2 \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{y}, \mathbf{x} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n], \end{aligned} \tag{32}$$

where $\mathbf{X}_i = x_{(i-1)d+1:id}$, , as shown in Figure 4.

In [41], an algorithm has been proposed to solve the problem in (32). Instead of equally treating all the channel blocks, we would like to take advantage of the channel estimation and learning after each iteration of problem solving to further increase the reconstruction quality. We give a block with higher channel estimated energy a lower weight so its recovered gain value is less restricted. We solve the following problem in our design:

$$\begin{aligned} \text{min} \quad & \sum_{i=1}^{n} w_i \|\mathbf{X}_i\|_2 \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{y}, \mathbf{x} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n], \end{aligned} \tag{33}$$

where w_i is a weighted factor for block $\mathbf{X}_i$.

6.2.2 Channel State Estimation with Co-located Legacy Antennas

A major challenge in estimating the channels of the mmW band is its narrow angular spread of signal and thus the need of performing the channel measurements in a large number of directions. At the BS end and mobile end of mmW cellular networks, it is likely that there are some co-located legacy antennas (e.g. from 3G cellular networks), which operate in much lower-frequency band than mmW, e.g., L Band: 1 to 2 GHz or Ultra High Frequency: 300MHz to 3GHz (for ease of presentation we will refer to it interchangeably as low-frequency band, legacy band or low-band). Since the number of antennas in the legacy band is typically small, these co-located TX and RX antennas can perform the direction search in a low cost fashion, i.e., even an exhaustive search will not bring unbearable overhead. As the mmW interface and co-located legacy interface share similar spatial environment, which leads to strong correlation between the mmW channel and the legacy channel, it would be helpful if we could exploit the channel information from the low-frequency channels to guide more intelligent beam direction finding in the mmW band.

Due to the small angular spread in the mmW channel, if some clusters (blocks in sparsity) in the low-band

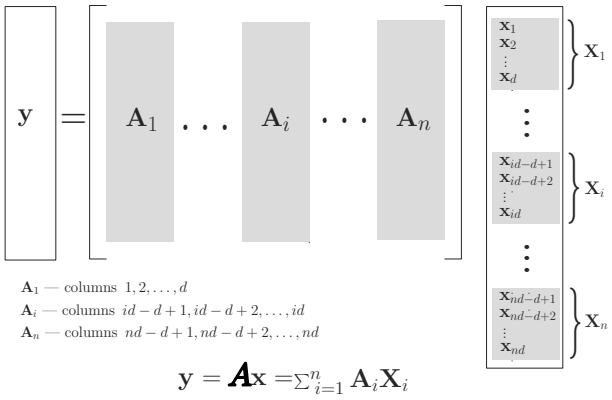

Fig. 4. Block sparse model.

are dominant, it is likely that an overlapped mmW channel has a good channel gain. Although signal paths in the mmW channel and the low-frequency channel may not be exactly the same, the information on the strong channel in the low-frequency band can give a rough information on which antenna direction range is more likely to have better antenna gains in the mmW channel. To provide a guide for more efficient mmW band beam training and channel estimation, we propose to exploit the legacy band information with two techniques: (1) legacy-band-assisted beam selection and (2) legacy-band-assisted channel reconstruction.

Legacy-band-assisted beam selection: We propose to use the low-band information to efficiently compress the initial range of beam selection for beamforming training. From the low-frequency channel information, one can easily find the best low-band direction (beam) pair. For mmW band, instead of training within the whole angular range at high cost, we propose to only randomly select beams within the best direction range detected in low-band training. This will significantly reduce the training overhead and speed up the training process. On the other hand, given the same number of beam directions to train, the beams selected with the guidance of low-band results are likely to have better quality than those randomly selected from the whole angular space, thus improving the channel estimation and beamforming performance.

After selected beams are trained, channel reconstruction is called for to estimate the channel, where legacy band information can be further exploited.

Legacy-band-assisted channel reconstruction: We propose to use the direction-range quality from the low-band as the initial weights w_i in our proposed beam alignment algorithm, and order the directions by their beam alignment gains. If a low-band direction overlapped by a mmW direction is detected to have a good channel quality, the weight of the corresponding mmW channel block is set to a smaller value in Problem (33).

Our weighted-block-sparse recovery algorithm is facilitated with the low-band information, as shown in Algorithm 1. The main idea is to assign a weight to

each block and update the weights through iteratively solving (33). For the blocks that are not likely to contain many non-zero elements, weights are increased, so that the values of these blocks can be reduced when solving the minimization problem in (33). For the blocks that are likely to be non-zero, weights are reduced to relax those blocks. Our algorithm enhances the channel reconstruction performance from two aspects: (a) error/noise reduction by disregarding data from likely-zero blocks as in Algorithm 1; (b) accuracy improvement by updating block weights through each iteration to concentrate the residual information to the likely-nonzero blocks to increase the signal recovery quality.

From Algorithm 1 we can see that it learns from the iterations and updates the weights by self-adaptation, and this contributes to the fast convergence of the algorithms. The computational complexity of Algorithm 1 is $\mathcal{O}(n^3)$, where $n = N/d$ is the block sparsity level (the number of blocks), d is the size of a block, and N is the size of path gain vector $\mathbf{x}$ that we try to recover in this algorithm. In the conventional CS recovery based on min-ℓ_1 [36], a typical convex optimization approach without considering block effect will incur a computational complexity of $\mathcal{O}(N^3) >> \mathcal{O}(n^3)$. Obviously, our block-based CS recovery has much lower complexity, thus can be completed with much shorter time. Our performance studies also demonstrate that our method can better capture the clustering features of mmWave channels to achieve higher recovery accuracy.

6.3 Beam Alignment after Channel Estimation

Optimizing the beamforming gain is a critical goal of beam alignment, since a larger beamforming gain can translate to larger receiving power and better signal-to-noise ratio thus achievable transmission rate. Assume there are I and J possible directions at TX and RX, respectively, after obtaining the channel estimation result, the optimal beam pair $(\mathbf{u}_{opt}, \mathbf{v}_{opt})$ can be determined from Algorithm 2. The optimal transmission and reception directions are the ones that can maximize the beamforming gain. Instead of blind or purely random beam training, the estimated channel information is applied in our scheme to guide the finding of the optimal beam direction.

7 Simulations and Results

In this section, we will perform simulations to show the effectiveness and efficiency of the proposed design.

We will compare the performances of the following schemes:

- *CODEBOOK*: Multi-resolution codebook-based beam searching as in IEEE 802.11ad [22] without CS channel reconstruction.
- *CS*: Use a greedy CS reconstruction algorithm such as [38] without block-sparsity to recover the channel.

Algorithm 1 Reconstruction of block sparse signals

Require:
Initialization.
Measured vector $\mathbf{y}$, size of blocks d, block sparsity level n and measurement matrix $\mathbf{A}$.

Ensure:

1: **Low-band exhaustive search.**
Each block $\mathbf{X}_i$ is given a weight w_i based on the co-located low-land antenna search result.

2: **Solve the following optimization problem using semi-definite programming**

$$\begin{aligned} \min_{\mathbf{x}} \quad & \sum_{i=1}^{n} w_i \|\mathbf{X}_i\|_2 \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{y}, \mathbf{x} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n]. \end{aligned} \quad (34)$$

3: **Reduce noise by updating $\mathbf{y}$.**
Sort $\mathbf{X}_i$ such that $\|\mathbf{X}_{j_1}\|_2 \geq \|\mathbf{X}_{j_2}\|_2 \geq ...\|\mathbf{X}_{j_n}\|_2$.
Update $\hat{\mathbf{A}}$ to be the submatrix of $\mathbf{A}$ containing columns of first $(n-1)d$ rows of $\mathbf{A}$ that corresponds to the blocks $j_1, ..., j_{n-1}$.
Update $\mathbf{y} \leftarrow \hat{\mathbf{A}}\mathbf{x}$ (Disregard the weakest block information from measurements.)

4: **Update weights w_i for each block**
Calculate $\mathbf{x} = \hat{\mathbf{A}}^{-1}\mathbf{y} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n]$.
For each block $\mathbf{X}_i$, count the number of $\mathbf{x}$ entries that are above a predetermined threshold as m_i.
Update the weight of each block, w_i as follows:
$w_i \leftarrow \frac{\sum_{i=1}^{n} m_i}{m_i}$.

5: **Iteration**
If the termination condition (e.g. allowed maximal number of iterations) is not met, go back to **Step 2**. Otherwise algorithm terminates.

6: **Output**
Output the recovered signal $\mathbf{x} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n]$.

Algorithm 2 Beam Alignment

Require: mmW channel estimation result $\mathbf{x}$
Initialization.

Ensure:

1: RX constructs channel $\mathbf{H}$ from $\mathbf{x}$.

2: RX determines the best beam pair as follows and transmit the best beam pair to TX:
$(\mathbf{u}_{opt}, \mathbf{v}_{opt}) = \arg\max_{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_I, \mathbf{v}_1, ..., \mathbf{v}_J} \|\mathbf{v}^H \mathbf{H} \mathbf{u}\|_1$ ($\mathbf{v}_j$ is j-th RX beam direction).

3: TX and RX align their beams according to direction pair $(\mathbf{u}_{opt}, \mathbf{v}_{opt})$.

- *BLOCK*: Use an unweighted block-sparse-based reconstruction algorithm [41] to recover channel.
- PROPOSED1: Use proposed weighted block-sparse-based reconstruction algorithm to recover channel, but without low-band block information (legacy band assistance). Block weights are initialized to the same value for all blocks.
- PROPOSED2: Use proposed weighted block-sparse-

based reconstruction algorithm with legacy-band-assisted reconstruction (but without legacy-band-assisted beam selection) to recover the channel. Block weights are initialized according to the legacy-band information.
- PROPOSED3: PROPOSED2 with legacy-band-assisted beam selection.

7.1 Simulation Settings

In the simulations, we consider a 28GHz mmW cellular network, and assume the TX has 64 $\lambda/2$ uniform linear arrays, and the RX has 16 $\lambda/2$ uniform linear arrays. We also assume the numbers of colocated low-band TX and RX antennas are both 4. The mmW channel is generated based on the model discussed in Section 4. We also select the channel parameters based on empirical values based on real-world measurements in [16], with default $K_d = 2$ dominating clusters, $L = 16$ transmission paths in each cluster. We assume the transmission bandwidth is 0.5 GHz.

To evaluate the effectiveness of different beam-alignment schemes, we compare their losses in link gain from the optimal one with different ratios of directions searched. The optimal beam direction is found through the exhaustive search of all possible beam pairs at high overhead. A bigger loss reflects a larger reduction of the transmission rate from the optimal. To reduce the search range, *CODEBOOK* exploits a multi-level codebook. After each coarse resolution of beam search, only the beam pairs within the range detected to have the best quality will be further trained. Although *CS*, *BLOCK* and our proposed schemes all train a subsect of beam pairs and exploit the channel information to guide the beam training, our proposed schemes further take advantage of the block sparse properties of mmWave channel and the assistance from the legacy band to improve the beam alignment efficiency. We expect our scheme can achieve comparable performances as the optimal one with the searching cost lower than that of other schemes.

The first metric we will use to evaluate the performance of a beam pair is the SNR degradation compared with the SNR value obtained at the optimal beam pair. We define the optimal SNR as $R_{opt} = R(\mathbf{u}_{opt}, \mathbf{v}_{opt})$ and the actual SNR obtained from the estimated best beam pair $(\mathbf{u}, \mathbf{v})$ as $R(\mathbf{u}, \mathbf{v})$. Then the SNR loss for this beam pair in decibels is defined as:

$$\text{Loss}(dB) = -10 \log_{10} \left[\frac{R(\mathbf{u}, \mathbf{v})}{R_{opt}} \right]. \quad (35)$$

The smaller the Loss (Loss is larger than 0 in definition), the better the beam pair selected.

The second metric we will evaluate is the *Search Rate*, which is defined as the number of measured subset of beam pairs (PQ) normalized to all the possible beam pairs $N_{tx}N_{rx}$, that is:

$$\text{Search Rate} = PQ/(N_{tx}N_{rx}). \quad (36)$$

The third metric we adopt is the actual data transmission rate achieved by communicating over the beam pair suggested by beam alignment.

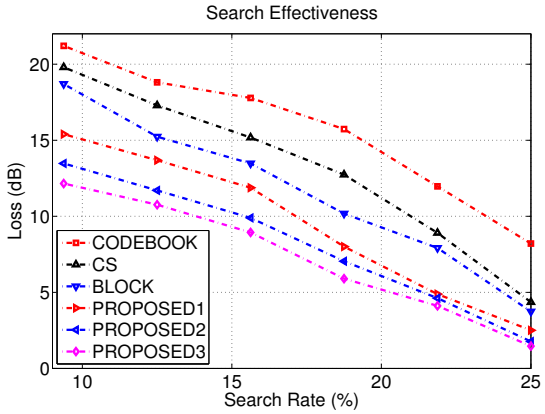

Fig. 5. Search effectiveness: Loss.

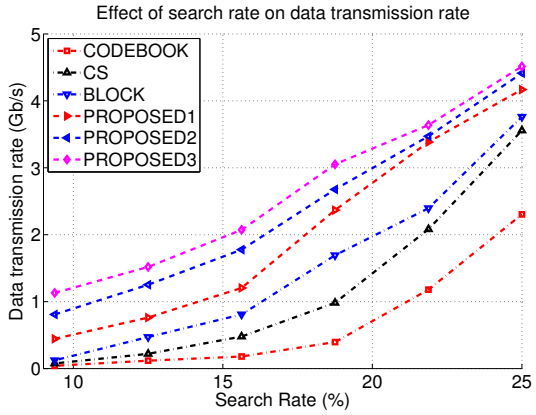

Fig. 6. Search effectiveness: data transmission rate.

7.2 Search Effectiveness

We use two metrics to evaluate the search effectiveness of beam alignment: Loss and data transmission rate.

Figures 5 and 6 show how beam alignment quality changes with different search rates for various beam alignment schemes. In Figure 5, among all the schemes compared, *CODEBOOK* is the most inefficient one with the lowest SNR loss. It may filter out the optimal beam direction in its coarse-level search and end up with the selection of some sub-optimal ones. In contrast, *CS*, *BLOCK* and our proposed schemes can rely on the estimated channel information to infer the optimal beam direction, even if the optimal one hasn't been trained earlier or falls out of the best coarse beam range suggested by the legacy channel. We also see that *BLOCK* performs better than the conventional *CS* recovery algorithm by considering the block properties of the sparse channel. At the same search rate, our three proposed schemes always outperform *CS* and *BLOCK* with lower SNR Loss. PROPOSED2 is able to achieve a small Loss of $1.8dB$ when only measuring one fourth of all the possible beam pairs. In comparison with *CODEBOOK*, and PROPOSED1, PROPOSED2 achieves a Loss reduction by 69.44% and 78.42%, respectively. Compared to *CS*, the Loss reduction of PROPOSED2 is

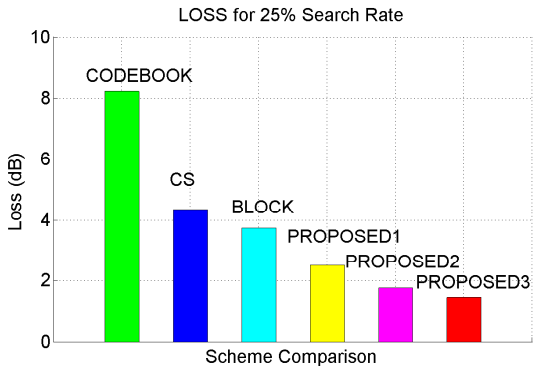

Fig. 7. Cost efficiency: Loss for fixed search rate.

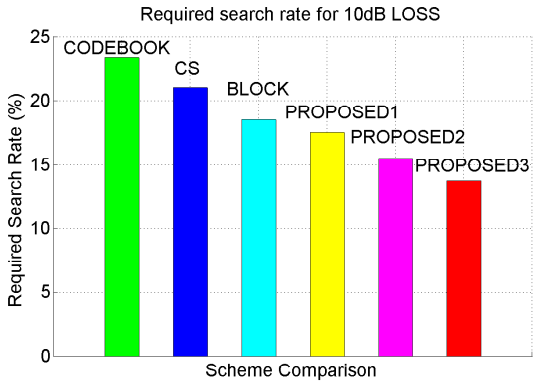

Fig. 8. Cost efficiency: required rate for a given Loss.

59.34%, whereas *BLOCK* and PROPOSED1 reduce the loss by 14.14% and 42.42%, respectively. Our proposed schemes not only take advantage of the sparse properties of mmW channel, but also exploit its block feature and learning from channel estimation in the previous rounds of iteration to further improve the channel estimation accuracy. In addition, PROPOSED2 outperforms PROPOSED1 by making use of the information from the low-frequency channels to facilitate the beam alignment. We also observe an additional improvement of about 20.83% for PROPOSED3, compared with PROPOSED2. This improvement is mainly contributed by the guidance of legacy band for the efficient selection of beams to train, which further increases the accuracy of channel reconstruction. Compared with the scheme using CS directly, the loss reduction from the use of PROPOSED3 is about 66.35% in the case of 25% sampling rate, and the loss reduction is generally above 38% for all tested cases.

In Figure 6, as expected, the transmission rates increase with the search rate. Our proposed schemes achieve a drastic enhancement in the data transmission rates from other methods. Compared with *CODEBOOK*, at the search rate 25%, PROPOSED1, PROPOSED2 and PROPOSED3 obtain an improvement of data rate by 224%, 255% and 267%, respectively. Compared with *CS*, the improvements are 30.73%, 43.06% and 48.17%, whereas the enhancements over *BLOCK* are 18.63%, 29.82% and 34.46%.

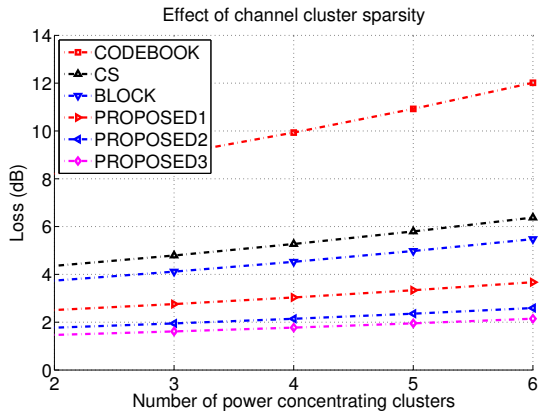

Fig. 9. Effect of channel clusters on Loss.

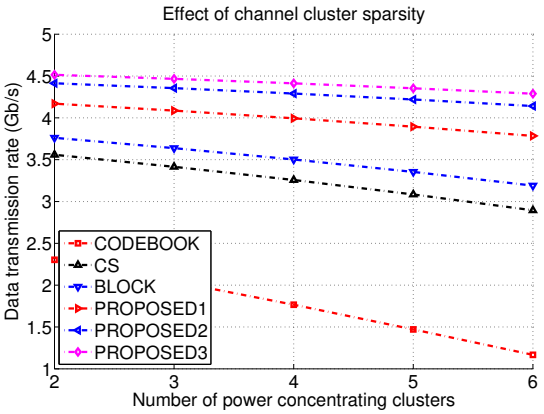

Fig. 10. Effect of channel clusters on transmission rate.

7.3 Cost Efficiency

For a given number of Search Rate (this usually happens in a resource-limited network, where TX and RX may only be able to measure a limited number of beam pairs), the estimated best beam pair quality determines the effectiveness of a beam alignment scheme. For a given SNR Loss, a certain Search Rate requirement will need to be met. The required Search Rate in this case indicates the cost efficiency of the scheme. When more beam pairs are searched, the overhead will become higher (e.g. time, energy, computational complexity).

Figure 7 shows the advantages of our proposed schemes over *CS* and *BLOCK*. For a given number of beam pairs trained, our proposed schemes experience much smaller Loss than *CS* and *BLOCK*. In Figure 8, we compare the search rate needed by different schemes to achieve the same beam alignment quality. Our schemes require a smaller number of training signal transmissions as compared to *CS* and *BLOCK*. At a Loss of $10dB$ achieved by *CS*, the search rate of our PROPOSED3 is only 59% that of *CODEBOOK*, 65.48% that of *CS*, 74.32% that of *BLOCK*. Compared to the exhaustive search, our rate is only 13.75%. The cost reduction can be huge under the circumstance that the number of possible beam directions is large.

7.4 Effect of Channel Clusters

The number of clusters that concentrate most of the signal power in the channel impacts the block sparse level

in the estimated path gain vector. The fewer the number of power concentrating clusters, the sparser the channel. By varying the number of dominating clusters K_d while keeping the total number of paths KL unchanged, we further investigate the effect of channel clustering on the achieved Loss when the Search Rate is kept the same at 25%. The results are shown in Figure 9.

In Figure 9, the cluster number has a significant impact on the channel estimation thus beam alignment performances. When the number of clusters increases (i.e., the channel becomes less sparse), the performances of all schemes degrade with a larger Loss, as 25% Search Rate is not enough. Among all the schemes, *CODEBOOK* is impacted the most, because more clusters introduce more uncertainty. The optimal beam direction has a higher chance of being filtered out during the coarse-level training. The performances of other schemes change less dramatically, because they are able to estimate the actual channel information to guide the beam alignment. When there are 6 channel clusters, compared to *CODEBOOK*, *CS* and *BLOCK* have a loss reduction of 46.93% and 54.43%, respectively, whereas Proposed1, Proposed2 and Proposed 3 reduce the loss at 69.44%, 78.42% and 82.14%, respectively. Again, from PROPOSED1 to PROPOSED2 then to PROPOSED3, we see consistently increase of improvement. Better exploiting the block sparsity, channel knowledge and low-frequency band assistance, our schemes can more effectively select the beams to train.

Figure 10 depicts the effects of clusters on the data transmission rate. We observe that when the channel becomes less sparse, the users will experience lower data transmission rate as a result of the degraded quality of the beam pair found in beam alignment. Compared to *CODEBOOK*, *CS*, *BLOCK*, PROPOSED1, PROPOSED2 and PROPOSED3 achieve rate improvements of 148%, 173%, 224%, 255% and 267%, respectively.

8 Conclusion

This paper presents an efficient beam alignment scheme for the transmitter and receiver to jointly decide the beam directions to combat the large path loss in the mmW cellular networks. Unlike the conventional scheme which searches through all the possible beam pairs or a large volume of codebook at the cost of severe delay and overhead, to enable efficient beam alignment, our algorithm takes advantage of the mmW channel sparsity to enable efficient beam direction matching. Rather than just randomly selecting a small subset of beam directions to train as done in existing CS-based channel measurement, to further reduce the beam alignment overhead and improve the gain, we exploit the block sparse properties and learning of iterative channel estimation to more accurately estimate the mmW channel at lower cost. We also exploit the assistance from co-located low-frequency antennas to guide the beam selection for training and further increase the channel reconstruction accuracy thus beam alignment quality. Simulation results demonstrate the significant advantages of our design in the search effectiveness and cost efficiency.

References

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

[2] S. Singh, R. Mudumbai, and U. Madhow, "Distributed coordination with deaf neighbors: Efficient medium access for 60 ghz mesh networks," in *Proceedings of IEEE INFOCOM*, pp. 1–9, March 2010.

[3] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Transactions on Communications*, vol. 63, pp. 3437–3458, Oct 2015.

[4] S. S. and Vignesh Venkateswaran and Xinyu Zhang and Parmesh Ramanathan, "60 ghz indoor networking through flexible beams: A link-level profiling," in *Proceedings of ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, (New York, NY, USA), pp. 71–84, ACM, 2015.

[5] S. Sur, I. Pefkianakis, X. Zhang, and K.-H. Kim, "Wifi-assisted 60 ghz wireless networks," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, MobiCom '17, (New York, NY, USA), pp. 28–41, ACM, 2017.

[6] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: Mm-wave beam steering without in-band measurement," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, pp. 2416–2424, April 2015.

[7] P. A. Eliasi, S. Rangan, and T. S. Rappaport, "Low-rank spatial channel estimation for millimeter wave cellular systems." arXiv:1410.4831, 2014.

[8] H. Deng and A. Sayeed, "Mm-wave mimo channel modeling and user localization using sparse beamspace signatures," *IEEE SPAWC*, June 2014.

[9] A. Sayeed and J. Brady, "Beamspace mimo for high-dimensional multiuser communication at millimeter-wave frequencies," *IEEE Globecom 2013*, pp. 3785–3789, Dec. 2013.

[10] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 164–174, 2010.

[11] A. Alkhateeby, G. Leusz, 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, IEEE, 2015.

[12] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, "Channel estimation in millimeter wave mimo systems with one-bit quantization," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, pp. 957–961, IEEE, 2014.

[13] 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 *Information Theory and Applications Workshop (ITA), 2015*, pp. 90–97, IEEE, 2015.

[14] E. J. C. and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.

[15] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[16] M. Akdeniz, Y. Liu, M. Samimi, S. Sun, S. Rangan, T. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[17] X. Li, J. Fang, H. Li, and P. Wang, "Millimeter wave channel estimation via exploiting joint sparse and low-rank structures," *IEEE Transactions on Wireless Communications*, vol. 17, pp. 1123–1133, Feb 2018.

[18] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 1499–1513, March 2014.

[19] H. Shokri-Ghadikolaei, C. Fischione, P. Popovski, and M. Zorzi, "Design aspects of short-range millimeter-wave networks: A mac layer perspective," *IEEE Network*, vol. 30, pp. 88–96, May 2016.

[20] L. C. Pansana, "Transmit-receive beamforming for 60 ghz indoor wireless communications." Thesis, Aalto University, 2010.

[21] "Ieee std 802.15.3c-2009 (amendment to ieee std 802.15.3-2003)." IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 15.3: Amendment 2: Millimeter-wave-based Alternative Physical Layer Extension, Oct. 2009.

[22] "Ieee standard 802.11ad-2012." IEEE Standard for Information technology–Local and metropolitan area networks–Specific requirements-Part 11: WLAN MAC and PHY Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band, 2012.

[23] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. A. Rahman, R. Funada, F. Kojima, I. Lakkis, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-gbps millimeter-wave wpan systems," *Proc. IEEE Global Telecommun. Conf.*, 2009.

[24] Z. Yang, P. H. Pathak, Y. Zeng, and P. Mohapatra, "Sensor-assisted codebook-based beamforming for mobility management in 60 ghz wlans," in *Mobile Ad Hoc and Sensor Systems (MASS), 2015 IEEE 12th International Conference on*, pp. 333–341, Oct 2015.

[25] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. on Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[26] C. Barati, S. Hosseini, S. Rangan, T. K. P. Liu, and S. Panwar, "Directional cell search for millimeter wave cellular systems," 2014.

[27] B. Li, Z. Zhou, W. Zou, X. Sun, and G. Du, "On the efficient beamforming training for 60ghz wireless personal area networks," *IEEE Trans. Wireless Comm.*, Feb. 2013.

[28] J. Singh and S. Ramakrishna, "On the feasibility of beamforming in millimeter wave communication systems with multiple antenna arrays," *IEEE Globecom*, 2014.

[29] E. Olfat, H. Shokri-Ghadikolaei, N. Najari-Moghadam, M. Bengtsson, and C. Fischione, "Learning-based pilot precoding and combining for wideband millimeter-wave networks," in *2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 1–5, Dec 2017.

[30] 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*, vol. 8, pp. 831–846, Oct 2014.

[31] A. Zhou, X. Zhang, and H. Ma, "Beam-forecast: Facilitating mobile 60 ghz networks via model-driven beam steering," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, pp. 1–9, May 2017.

[32] E. J. C. and J. Romberg and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[33] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, pp. 267–288, 1996.

[34] S. C. and D. Donoho and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, pp. 33–61, 1998.

[35] E. J. C. and T. Tao, "The dantzig selector: Statistical estimation when p is much larger than n," *Annals Of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

[36] E. J. Candès and J. Romberg, "ℓ_1 -magic: Recovery of sparse signals via convex programming," 2005.

[37] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.

[38] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, May 2009.

[39] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[40] C. G. Khatri and C. R. Rao, "Solutions to some functional equations and their applications to characterization of probability distributions," *Sankhy: The Indian Journal of Statistics, Series A (1961-2002)*, vol. 30, no. 2, pp. 167–180, 1968.

[41] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3075 –3085, 2009.

Jie Zhao received the Ph.D. degree in Electrical Engineering from the State University of New York at Stony Brook, New York, USA, and the B.S. degree in Telecommunications Engineering from Huazhong University of Science and Technology, Wuhan, China. His research interests include millimeter-wave communications, cognitive radio networks, as well as networked sensing and detection.

Xin Wang received the B.S. and M.S. degrees in telecommunications engineering and wireless communications engineering respectively from Beijing University of Posts and Telecommunications, Beijing, China, and the Ph.D. degree in electrical and computer engineering from Columbia University, New York, NY. She is currently an Associate Professor in the Department of Electrical and Computer Engineering of the State University of New York at Stony Brook, Stony Brook, NY. Before joining Stony Brook, she was a Member of Technical Staff in the area of mobile and wireless networking at Bell Labs Research, Lucent Technologies, New Jersey, and an Assistant Professor in the Department of Computer Science and Engineering of the State University of New York at Buffalo, Buffalo, NY. Her research interests include algorithm and protocol design in wireless networks and communications, mobile and distributed computing, as well as networked sensing and detection. She has served in executive committee and technical committee of numerous conferences and funding review panels, and serves as the associate editor of IEEE Transactions on Mobile Computing. Dr. Wang achieved the NSF career award in 2005, and ONR challenge award in 2010.

Harish Viswanathan is a CTO Partner in the Corp CTO organization. As a CTO Partner he advises the Corporate CTO on Technology Strategy through in-depth analysis of emerging technology and market needs. Harish Viswanathan joined Bell Labs in 1997 and has worked on multiple antenna technology for cellular wireless networks, mobile network architecture, and M2M. He received the B. Tech. degree from the Department of Electrical Engineering, Indian Institute of Technology, Chennai, India and the M.S. and Ph.D. degrees from the School of Electrical Engineering, Cornell University, Ithaca, NY. He holds more than 50 patents and has published more than 100 papers. He is a Fellow of the IEEE and a Bell Labs Fellow.

Arjuna Madanayake is an Associate Professor at Florida International University (FIU) in Miami, Florida. He completed his Ph.D. and M.Sc. degrees, both in electrical engineering, from the University of Calgary, in Alberta, Canada. He obtained a B.Sc with Specialization in Electronic and Telecommunication Engineering from the University of Moratuwa, Sri Lanka, in 2002. Dr. Madanayake was a tenured faculty member at the University of Akron, in Akron, Ohio, between 2010-2018 before joining the faculty at FIU in August 2018. His research interests are in array signal processing, circuits, systems, electronics, fast algorithms and computer architecture.

Guangxue Yue is a Professor in the College of Mathematical and Information Engineering, Jiaxing University, Jiaxing, Zhejiang, China. He received the M.S. and Ph.D. degrees from Hunan University in 2004 and 2012, respectively. His main research interests include cloudy convergence and collaborative services, wireless mesh networks and mobile cloud computing.