A SPARSE ENCODING AND PHASELESS DECODING APPROACH FOR FAST MMWAVE BEAM ALIGNMENT

Xingjian Li[1] *Jun Fang*[*1] *Huiping Duan*[1] *Hongbin Li*[2]

[1] University of Electronic Science and Technology of China, JunFang@uestc.edu.cn
[2] Stevens Institute of Technology, Hongbin.Li@stevens.edu

ABSTRACT

The problem of beam alignment for millimeter wave (mmWave) communications is studied in this paper. We show that, by exploiting the sparse scattering nature of mmWave channels, the beam alignment problem can be formulated as a sparse encoding and phaseless decoding problem, which involves finding a sparse sensing matrix and an efficient recovery algorithm to recover the support and magnitude of the sparse signal from compressive phaseless measurements. We develop a general function-Code (GF-Code) algorithm for sparse encoding and phaseless decoding. Simulation results are provided to corroborate the effectiveness of the proposed GF-Code method.

***Index Terms*—** mmWave communications, beam alignment, sparse encoding and phaseless decoding

1. INTRODUCTION

Millimeter wave (mmWave) communication has the potential to offer gigabits-per-second communication data rates by exploiting the large bandwidth available at mmWave frequencies [1–3]. Nevertheless, communication at the mmWave frequency bands suffers from high attenuation and signal absorption [4]. To address this issue, large antenna arrays should be used to provide sufficient beamforming gain for mmWave communications [5]. Due to the narrow beam of the antenna array, beamforming training is required to find the best beamformer-combiner pair that gives the highest beamforming gain. A natural approach is to exhaustively search for all possible beam pairs to identify the best beam alignment. This exhaustive search has a sample complexity of $\mathcal{O}(N^2)$ (N denotes the number of possible beam directions) and usually takes a long time (up to several seconds) to converge. The IEEE 802.11ad standard [6] employs an exhaustive search at the receiver, with the transmitter adopting a quasi-omnidirectional beam pattern. This process is then reversed to have the transmitter sequentially scan the entire space while the receiver uses a quasi-omnidirectional beam shape. This protocol still incurs a considerable delay (hundreds of milliseconds) in practice [7]. To further reduce the training time, adaptive beam alignment algorithms [8–11] and compressed sensing-based approaches [12–20] were proposed recently. The adaptive beam alignment requires to adaptively choose a subcodebook at each stage based on the output of earlier stages. A drawback of this scheme is that the base station has to interact with each user individually, which may not be feasible at the initial channel acquisition stage. Compressed sensing-based methods usually involve a computational complexity that might be too excessive for practical systems. In addition, compressed sensing methods require the knowledge of the phase of the measurements, which may not be available in practice.

In this paper, by exploiting the sparse scattering nature of mmWave channels, we show that the beam alignment problem can be formulated as a sparse encoding and phaseless decoding problem. We propose a general function-Code (GF-Code) algorithm for sparse encoding and phaseless decoding. The proposed algorithm is computationally efficient and noise-robust. Also, it can recover the support and magnitude information of a K-sparse signal with a sample complexity of $\mathcal{O}(K^2)$, thus providing a competitive solution for practical mmWave beam alignment systems.

2. SYSTEM MODEL

Consider a mmWave communication system which consists of a receiver (user) and a transmitter (base station). We assume that a hybrid analog and digital beamforming structure is employed at the transmitter, while the receiver has an omni-directional antenna that receives in all directions. The transmitter is equipped with N antennas and R RF chains, where $R \ll N$. The mmWave channel is characterized by a geometric channel model [13]

$$\boldsymbol{h} = \sum_{p=1}^{P} \alpha_p \boldsymbol{a}_t(\theta_p) \tag{1}$$

where P is the number of paths, α_p is the complex gain associated with the pth path, $\theta_p \in [0, 2\pi]$ is the associated azimuth angle of departure (AoD), and $\boldsymbol{a}_t \in \mathbb{C}^N$ is the transmitter array response vector. Suppose a uniform linear array (ULA) is

*The author would like to thank the support by the National Science Foundation of China under Grants 61829103 and 61871091.

used. Due to the sparse scattering nature of mmWave channels, $\boldsymbol{h}$ has a sparse representation in the beam space (angle) domain: $\boldsymbol{h} = \boldsymbol{D}\boldsymbol{x}$, where $\boldsymbol{D} \in \mathbb{C}^{N\times N}$ is the discrete Fourier transform (DFT) matrix, and $\boldsymbol{x} \in \mathbb{C}^N$ is a K-sparse vector. If the true AoD parameters $\{\theta_p\}$ lie on the discretized grid specified by the DFT matrix, then the number of nonzero entries in the beam space domain equals the number of signal paths, i.e. $K = P$. The objective of beam alignment is to estimate the AoD and the attenuation (in magnitude) of each path, which is equivalent to recover the location indices and the magnitudes of the nonzero entries in $\boldsymbol{x}$. The AoDs of the dominant paths are then reported back to the base station via a control channel for beam alignment. Suppose the transmitter sends a constant signal $s(t) = 1$ to the receiver. The signal received at the tth time instant can thus be expressed as

$$r(t) = \boldsymbol{h}^T\boldsymbol{c}(t)s(t) + w(t) = \boldsymbol{c}^T(t)\boldsymbol{D}\boldsymbol{x} + w(t) \quad (2)$$

where $\boldsymbol{c}(t) \in \mathbb{C}^N$ is the beamforming vector used by the transmitter at the tth time instant, and $w(t)$ denotes the additive complex Gaussian noise with zero mean and variance σ^2. The beamforming vector can be expressed as

$$\boldsymbol{c}(t) = \boldsymbol{F}_{\text{RF}}(t)\boldsymbol{f}_{\text{BB}}(t) \quad (3)$$

in which $\boldsymbol{F}_{\text{RF}}(t) \in \mathbb{C}^{N\times R}$ and $\boldsymbol{f}_{\text{BB}}(t) \in \mathbb{C}^R$ represent the radio frequency (RF) precoding matrix and the baseband (BB) precoding vector, respectively. Specifically, to provide a sufficient beamforming gain for signal reception, the transmitter needs to form multiple beams simultaneously and steers them towards different directions to probe the channel. To this objective, the RF precoding matrix is chosen to be a submatrix of the DFT matrix $\boldsymbol{D}$, i.e.

$$\boldsymbol{F}_{\text{RF}}(t) = \boldsymbol{D}^*\boldsymbol{S}(t) \quad (4)$$

where $\boldsymbol{S}(t) \in \mathbb{R}^{N\times R}$ is a column selection matrix containing only one nonzero entry per column.

Substituting (3)–(4) into (2), we obtain

$$r(t) = \boldsymbol{a}^T(t)\boldsymbol{x} + w(t) \quad (5)$$

where $\boldsymbol{a}(t) \triangleq \boldsymbol{S}(t)\boldsymbol{f}_{\text{BB}}(t)$ is an N-dimensional sparse vector with at most R nonzero elements. It should be noted (5) is an ideal model without taking the carrier frequency offset (CFO) effect into account. In mmWave communications, due to the CFO between the transmitter and the receiver, the measurements $r(t)$ may incur an additional unknown phase shift that varies across time. In this case, only the magnitude information of the measurements $r(t), t = 1, \ldots, T$ is reliable.

Our objective, therefore, is to devise a measurement matrix $\boldsymbol{A} \triangleq [\boldsymbol{a}(1) \ \ldots \ \boldsymbol{a}(T)]^T \in \mathbb{C}^{T\times N}$ and develop an efficient recovery algorithm to recover $\boldsymbol{z} = |\boldsymbol{x}|$ from phaseless measurements:

$$\boldsymbol{y} \triangleq |\boldsymbol{r}| = |\boldsymbol{A}\boldsymbol{x} + \boldsymbol{w}| \quad (6)$$

where $\boldsymbol{r} \triangleq [r(1) \ \ldots \ r(T)]^T$, and $\boldsymbol{w} \triangleq [w(1) \ \ldots \ w(T)]^T$. The measurement matrix $\boldsymbol{A}$ has to satisfy:

C1 $\boldsymbol{A}$ is a sparse matrix with each row of $\boldsymbol{A}$ containing at most R nonzero elements.

For this reason, the design of the measurement matrix $\boldsymbol{A}$ is referred to as sparse encoding.

3. PROPOSED GF-CODE ALGORITHM

In this section, we propose a GF-Code algorithm for sparse encoding and phaseless decoding.

3.1. Sparse Encoding

The proposed GF-Code uses a set of bipartite graphs $\{G_l\}_{l=1}^L$ to encode the sparse signal. Let $\boldsymbol{H}_l \in \{0,1\}^{M\times N}$ denote the binary code matrix associated with G_l with N left nodes and M right nodes. The (i,j)th entry of $\boldsymbol{H}_l$ is given by

$$H_l(i,j) = \begin{cases} 1 & \text{if and only if left node } j \text{ of } G_l \text{ is connected} \\ & \text{to right node } i \text{ of } G_l \\ 0 & \text{otherwise} \end{cases}$$

Given $\{\boldsymbol{H}_l\}$, the measurement matrix $\boldsymbol{A} \in \mathbb{R}^{2ML\times N}$ is devised as

$$\boldsymbol{A} \triangleq \begin{bmatrix} \boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{A}_L \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{H}_1 \odot \boldsymbol{T} \\ \vdots \\ \boldsymbol{H}_L \odot \boldsymbol{T} \end{bmatrix} \quad (7)$$

where $\odot$ denotes the Khatri-Rao product, and $\boldsymbol{T} \in \mathbb{R}^{2\times N}$ is a modulation matrix defined as

$$\boldsymbol{T} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1/N & 2/N & \cdots & 1 \end{bmatrix} \quad (8)$$

For each graph G_l, each of its left node can be deemed as a component of the sparse signal $\boldsymbol{x}$, and each right node of G_l refers to a set of 2 measurements obtained as

$$\boldsymbol{y}_{l,m} = |(\boldsymbol{H}_l[m,:] \odot \boldsymbol{T})\boldsymbol{x} + \boldsymbol{w}_{l,m}| \quad \forall m = 1, \ldots, M \quad (9)$$

where $\boldsymbol{H}_l[m,:]$ denotes the mth row of $\boldsymbol{H}_l$, and $\boldsymbol{w}_{l,m}$ denotes the noise added to the mth right node of G_l. A left node, say node n, is called as active left node if the nth signal component, x_n, is nonzero. For a K-sparse signal $\boldsymbol{x}$, there are K active left nodes in total. A right node is called as a nullton, a singleton or a multiton if:

- Nullton: A right node is a nullton if it is not connected to any active left node.
- Singleton: A right node is a singleton if it is connected to exactly one active left node.

- Multiton: A right node is a multiton if it is connected to more than one active left node.

A bipartite graph is called an NM-graph if

- No-Multiton-graph (NM-graph): A bipartite graph whose right nodes are either singletons or nulltons.

For our proposed GF-Code, the purpose of employing multiple bipartite graphs is to ensure that, with an overwhelming probability, there exists at least an NM-graph, i.e. a bipartite graph whose right nodes are either singletons or nulltons.

The bipartite graphs $\{G_l\}$ with N left nodes and M ($M > K$) right nodes are designed as follows. Assume $r \triangleq N/M$ to be an integer. For each graph, we randomly divide N left nodes into M equal-size, disjoint sets (i.e. each set has r left nodes) and establish a one-to-one correspondence between M sets of left nodes and M right nodes. If N is not an integer multiple of M, we can still divide N left nodes into M disjoint sets, with all sets, except the last one, consisting of $r = \text{floor}(N/M)$ left nodes. For each bipartite graph G_l devised as described, its corresponding binary code matrix $\boldsymbol{H}_l$ has only one nonzero element per column, and at most r nonzero elements per row. Therefore, we can choose $R \geq r$ such that $\boldsymbol{A}$ satisfies constraint C1. Once $\boldsymbol{A}$ is given, the RF precoding matrices $\{\boldsymbol{F}_{\text{RF}}(t)\}$ and baseband precoding vectors $\{\boldsymbol{f}_{\text{BB}}(t)\}$ can be accordingly determined.

3.2. Phaseless Decoding

We next devise a phaseless decoding scheme to estimate $\boldsymbol{z} = |\boldsymbol{x}|$ from noisy measurements $\boldsymbol{y}$. The measurements associated with the bipartite graph G_l are give by $\boldsymbol{y}_l \triangleq |\boldsymbol{A}_l\boldsymbol{x} + \boldsymbol{w}_l|$, and the measurements, $\boldsymbol{y}_{l,m} \in \mathbb{R}^2$, corresponding to the mth right node of G_l are expressed as

$$\boldsymbol{y}_{l,m} = |(\boldsymbol{H}_l[m,:] \odot \boldsymbol{T})\boldsymbol{x} + \boldsymbol{w}_{l,m}| \quad \forall m = 1, \dots, M \tag{10}$$

We first need to decide whether a right node of G_l is a nullton or not. Such a problem can be formulated as a binary hypothesis test problem:

$$\begin{aligned} H_0 &: \; y_{l,m}^{(1)} = |w_{l,m}^{(1)}| \\ H_1 &: \; y_{l,m}^{(1)} = \left| \sum_{m_i \in S} x_{m_i} + w_{l,m}^{(1)} \right| \end{aligned} \tag{11}$$

where $y_{l,m}^{(1)}$ denotes the first entry of $\boldsymbol{y}_{l,m}$, $w_{l,m}^{(1)}$ is the additive complex Gaussian noise with zero mean and variance σ^2, and S denotes the set of indices of those active left nodes that are connected to the mth right node of G_l. A simple energy detector can be used to perform the detection: $y_{l,m}^{(1)} \underset{H_0}{\overset{H_1}{\gtrless}} \epsilon$. It is clear that $y_{l,m}^{(1)}$ under H_0 follows a Rayleigh distribution. Given a prescribed false alarm probability, the threshold $\epsilon > 0$ can be easily determined from the distribution of $y_{l,m}^{(1)}$ under H_0. Such an energy detector is able to yield satisfactory detection performance for a moderate and high signal-to-noise ratio.

To proceed with our decoding scheme, we assume all nullton right nodes of G_l are correctly identified. In this case, we are able to determine whether G_l is an NM-graph or not. Specifically, if G_l is an NM-graph, then it contains $M - K$ nullton right nodes; otherwise the number of nullton right nodes is greater than $M - K$. Although the number of active left nodes, K, is unknown *a priori*, those graphs which have the smallest number of nullton right nodes can be considered as NM-graphs and K can be simply estimated as $\hat{K} = M - J$, where J denotes the smallest number of nullton right nodes among all graphs.

We now perform decoding on those NM-graphs. Suppose G_l is an NM-graph and its mth right node is a singleton. Also, x_{m_i} is the active left node connected to the mth right node. When noise is present, the measurements corresponding to the mth right node of the graph G_l can be expressed as

$$\boldsymbol{y}_{l,m} = \begin{bmatrix} |x_{m_i} + w_{l,m}^{(1)}| \\ \left|\frac{m_i}{N} x_{m_i} + w_{l,m}^{(2)}\right| \end{bmatrix} \triangleq \begin{bmatrix} y_{l,m}^{(1)} \\ y_{l,m}^{(2)} \end{bmatrix} \tag{12}$$

where $y_{l,m}^{(1)}$ and $y_{l,m}^{(2)}$ denote the first and the second entry of $\boldsymbol{y}_{l,m}$, respectively, $w_{l,m}^{(1)}$ and $w_{l,m}^{(2)}$ denote the observation noise added to the first and the second entry of the mth right node, respectively. In this case, the magnitude and the location index of the active left node can be estimated as

$$\begin{aligned} z_{\hat{m}} &= y_{l,m}^{(1)} \\ \hat{m} &= \underset{m \in \{m_1^{(l)}, \dots, m_r^{(l)}\}}{\arg\min} \left| \frac{m}{N} - \frac{y_{l,m}^{(2)}}{y_{l,m}^{(1)}} \right| \end{aligned} \tag{13}$$

where $\{m_1^{(l)}, \dots, m_r^{(l)}\}$ denotes the set of indices of the left nodes connected to the mth right node of G_l. After performing (13) for all singleton right nodes, we are able to obtain an estimate of $\boldsymbol{z} = |\boldsymbol{x}|$. Let $\hat{\boldsymbol{z}}^{(l)}$ denote an estimate of $\boldsymbol{z}$ obtained from the measurements associated with G_l. Since we may have more than one NM-graphs, we are able to collect multiple estimates of $\boldsymbol{z}$. The problem lies in, due to the existence of noise, these multiple estimates, denoted as $\{\hat{\boldsymbol{z}}^{(1)}, \dots, \hat{\boldsymbol{z}}^{(I)}\}$, are not exactly the same, in terms of both the estimated magnitudes and the estimated support sets. In the following, we propose a set-intersection scheme to combine these multiple estimates into a more accurate estimate.

To better illustrate our idea, suppose there are two NM-graphs, say G_i and G_j, and x_n is the only active left node in $\boldsymbol{x}$. Recall that for each bipartite graph, the N left nodes are divided into M disjoint sets, with each set of left nodes connected to an individual right node. Let $S_n^{(i)}$ denote the set of left nodes to which x_n belongs in graph G_i, and $S_n^{(j)}$ denote the set of left nodes to which x_n belongs in graph G_j. Suppose the singleton right nodes in both G_i and G_j are correctly identified. Then we know that x_n belongs to both $S_n^{(i)}$ and $S_n^{(j)}$.

If the intersection of the two sets $S_n^{(i)}$ and $S_n^{(j)}$, $S_n^{(i)} \cap S_n^{(j)}$, contains only one element, then it must be x_n and the location index of x_n can be uniquely determined. Such an idea can be easily extended to the scenario where there are more then two NM-graphs, and for such a case, the set-intersection scheme is more likely to succeed because the more sets are used, the higher the probability of the intersection of these sets containing only one element.

There, however, is a problem for the general case where $\boldsymbol{x}$ contains multiple nonzero components (i.e. multiple active left nodes). In this case, we have no idea which set of left nodes a certain active node belongs to for each NM-graph. As a result, it is impossible to determine which sets should be put together to perform the intersection operation. To overcome this difficulty, we note that the magnitudes of those active left nodes are generally different. Hence the estimated magnitude can be used to identify a certain active left node. Without loss of generality, let $x_1, \ldots, x_K$ denote the nonzero components of $\boldsymbol{x}$ in decreasing order in terms of magnitude, i.e. $|x_1| > \cdots > |x_K| > 0$. For each NM-graph, say graph G_i, we can obtain an estimate of $|\boldsymbol{x}|$, denoted as $\boldsymbol{z}^{(i)}$. Specifically, let $\hat{z}_{i_1} > \cdots > \hat{z}_{i_K} > 0$ represent the nonzero components of $\hat{\boldsymbol{z}}^{(i)}$, then the kth largest element $\hat{z}_{i_k}$ can be regarded as an estimate of $|x_k|$. For each NM-graph, say G_i, the set of left nodes containing x_k can therefore be determined as the set of left nodes containing $\hat{z}_{i_k}$. A set intersection operation can then be performed to yield the final estimate of the location index of x_k. On the other hand, the magnitude of the kth largest component of $\boldsymbol{x}$ can be estimated as the average of all estimates, i.e. $|\hat{x}_k| = \frac{1}{I}\sum_{i=1}^{I} \hat{z}_{i_k}$. Note that if the intersection of the sets contains more than one element, then we randomly pick up an element in the intersection set as the estimate of the location index of x_k. In addition, in case the intersection is an empty set, which is possible due to the incorrect association between $\{x_1, \ldots, x_K\}$ and $\{\hat{z}_{i_1}, \ldots, \hat{z}_{i_K}\}$, we randomly select an estimate from $\{\hat{\boldsymbol{z}}^{(1)}, \ldots, \hat{\boldsymbol{z}}^{(I)}\}$ as the final estimate.

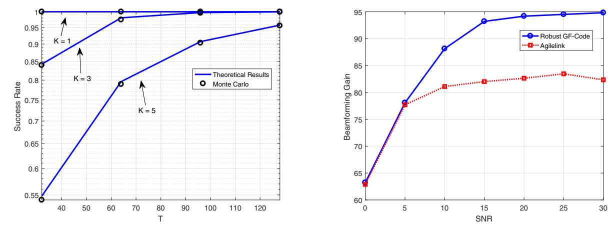
3.3. Theoretical Results for GF-Code

We have the following theoretical guarantee for our proposed GF-Code scheme. To simplify our analysis, we assume a noiseless case and $r \triangleq N/M$ is an integer. The results are summarized as follows.

Theorem 1 *Consider the phaseless decoding problem in (6), where the measurement matrix* $\boldsymbol{A} \in \mathbb{R}^{2ML \times N}$ *is generated according to our proposed sparse encoding scheme. If* $M \geq K$*, then our proposed algorithm can perfectly recover* $\boldsymbol{z} = |\boldsymbol{x}|$ *from phaseless measurements with probability exceeding*

$$p = 1 - (1 - \lambda)^L \quad (14)$$

where λ *is defined as* $\lambda \triangleq r^K C_M^K / C_N^K$*, in which* C_N^K *denotes the number of* K*-combinations from a set with* N *elements.*



(a) Success rates vs. T.

(b) Beamforming gains vs. SNR.

Fig. 1. Success rates vs. T, and beamforming gains vs. SNR.

Due to space limitations, the proof for Theorem 1 is omitted but available at `https://arxiv.org/abs/1811.04775`, where we also show that the sample complexity for our proposed algorithm is of order $\mathcal{O}(K^2)$.

4. SIMULATION RESULTS

In our simulations, the distance between neighboring antenna elements is assumed to be $d = \lambda/2$. The nonzero components of $\boldsymbol{x}$ are assumed to be random variables following a circularly symmetric complex Gaussian distribution $\mathcal{CN}(0, 1)$, and the locations of nonzero entries of $\boldsymbol{x}$ are uniformly chosen at random. Fig. 1(a) depicts the success rates of our proposed algorithm as a function of the total number of measurements $T = 2ML$, where we set $N = 128$, and $M = 16$. The success rate is computed as the ratio of the number of successful trials to the total number of independent runs. A trial is considered successful if $\|\hat{\boldsymbol{z}} - \boldsymbol{z}\|_2^2 / \|\boldsymbol{z}\|_2^2 < 10^{-8}$. The results are averaged over 10^4 independent runs. In each run, $\boldsymbol{x}$ (i.e. $\boldsymbol{h}$) is randomly generated. In the figure, solid lines represent the theoretical performance given in (14), while the circle marks represent the performance obtained via the Monte Carlo experiments. From Fig. 1(a), we see that our theoretical result matches the empirical result very well.

Next, we compare our proposed algorithm with the Agile-Link [21], a beam steering scheme which also relies on the magnitude information of measurements for recovery of signal directions. The beamforming gain defined below is used as a metric to evaluate the performance of respective beam steering schemes $G_{\text{BF}} = E[N|\boldsymbol{a}_t^H(\hat{\theta}_{opt})\boldsymbol{h}|^2 / \|\boldsymbol{h}\|_2^2]$, in which $\hat{\theta}_{opt}$ denotes the estimated optimal beam direction. Fig. 1(b) depicts the beamforming gains of respective algorithms as a function of signal-to-noise ratio (SNR), where the total number of measurements $T = 64$ is set the same for both schemes, and, for a fair comparison, the beamforming vector $\boldsymbol{c}(t)$ used in both schemes is normalized to unit norm. The SNR is defined as $10\log(\|\boldsymbol{h}\|_2^2/(N\sigma^2))$. We see that our proposed method yields a higher beamforming gain than the Agile-Link scheme, which suggests that our proposed method can help find a more accurate optimal beam direction.

5. REFERENCES

[1] Theodore S. Rappaport et al., "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proc. IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug. 2011.

[2] Sundeep Rangan et al., "Millimeter-wave cellular wireless networks: potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, March 2014.

[3] Amitava Ghosh et al., "Millimeter-wave enhanced local area systems: a high-data-rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152–1163, June 2014.

[4] A. Lee Swindlehurst et al., "Millimeter-wave massive MIMO: the next wireless revolution?," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, September 2014.

[5] Ahmed Alkhateeb et al., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, December 2014.

[6] "IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band," *IEEE Std 802.11ad-2012 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012 and IEEE Std 802.11aa-2012)*, pp. 1–628, Dec. 2012.

[7] Sanjib Sur et al., "60 GHz indoor networking through flexible beams: A link-level profiling," *SIGMETRICS Perform. Eval. Rev.*, vol. 43, no. 1, pp. 71–84, June 2015.

[8] J. Song et al., "Adaptive millimeter wave beam alignment for dual-polarized MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6283–6296, Nov. 2015.

[9] Sooyoung Hur et al., "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, October 2013.

[10] C. Liu et al., "Millimeter-Wave small cells: Base station discovery, beam alignment, and system design challenges," *IEEE Wireless Communications*, vol. 25, no. 4, pp. 40–46, August 2018.

[11] S. Noh et al., "Multi-Resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5689–5701, Sept. 2017.

[12] Ahmed Alkhateeb et al., "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?," in *Proc. 40th IEEE Inter. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, Brisbane, Australia, April 19-24 2015, pp. 2909–2913.

[13] Ahmed Alkhateeb et al., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, October 2014.

[14] Philip Schniter and Akbar Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, California, USA, November 2-5 2014, pp. 273–277.

[15] Taejoon Kim and David J. Love, "Virtual AoA and AoD estimation for sparse millimeter wave MIMO channels," in *Proc. 16th IEEE Inter. Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, Stockholm, Sweden, June 28 - July 1 2015, pp. 146–150.

[16] Zhinus Marzi, Dinesh Ramasamy, and Upamanyu Madhow, "Compressive channel estimation and tracking for large arrays in mm-Wave picocells," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 514–527, April 2016.

[17] Zhen Gao, Linglong Dai, Zhaocheng Wang, and Sheng Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169–6183, December 2015.

[18] Xinyu Gao, Linglong Dai, and Akbar M. Sayeed, "Low RF-complexity technologies for 5G millimeter-wave MIMO systems with large antenna arrays," *available at arXiv:1607.04559*, 2016.

[19] Zhou Zhou, Jun Fang, Linxiao Yang, Hongbin Li, Zhi Chen, and Shaoqian Li, "Channel estimation for millimeter-wave multiuser MIMO systems via PARAFAC decomposition," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7501–7516, November 2016.

[20] Zhou Zhou, Jun Fang, Linxiao Yang, Hongbin Li, Zhi Chen, and Rick S. Blum, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE Journal Selected Areas in Communications*, vol. 35, no. 7, pp. 1524–1538, July 2017.

[21] Omid Abari, Haitham Hassanieh, Michael Rodriguez, and Dina Katabi, "Millimeter wave communications: From point-to-point links to agile network connections," in *Proc. 15th ACM Workshop on Hot Topics in Networks*, Atlanta, Georgia, USA, November 9-10 2016, pp. 169–175.