Millimeter Wave Beam-Selection Using Out-of-Band Spatial Information

Anum Ali, *Student Member, IEEE*, Nuria González-Prelcic, *Member, IEEE*, and Robert W. Heath Jr., *Fellow, IEEE*

***Abstract*—Millimeter wave (mmWave) communication is one feasible solution for high data-rate applications like vehicular-to-everything communication and next generation cellular communication. Configuring mmWave links, which can be done through channel estimation or beam-selection, however, is a source of significant overhead. In this paper, we propose to use spatial information extracted at sub-6 GHz to help establish the mmWave link. Assuming a fully digital architecture at sub-6 GHz; and an analog architecture at mmWave, we outline a strategy to extract spatial information from sub-6 GHz and its use in mmWave compressed beam-selection. Specifically, we formulate compressed beam-selection as a *weighted* sparse signal recovery problem, and obtain the weighting information from sub-6 GHz channels. In addition, we outline a structured precoder/combiner design to tailor the training to out-of-band information. We also extend the proposed out-of-band aided compressed beam-selection approach to leverage information from all active OFDM subcarriers at mmWave. To simulate multi-band frequency dependent channels, we review the prior work on frequency dependent channel behavior and outline a multi-frequency channel model. The simulation results for achievable rate show that out-of-band aided beam-selection can considerably reduce the training overhead of in-band only beam-selection.**

***Index Terms*—Millimeter-wave communications, beam-selection, out-of-band information, weighted compressed sensing, structured random codebooks**

I. Introduction

MILLIMETER WAVE (mmWave) communication systems use large antenna arrays and directional beamforming/precoding to provide sufficient link margin [2], [3]. Large arrays are feasible at mmWave as antennas can be packed into small form factors. Configuring these arrays, however, is not without challenges. First, the high power consumption of RF components makes fully digital baseband precoding difficult [2]. Second, the precoder design usually relies on channel state information, which is difficult to acquire at mmWave due to large antenna arrays and low pre-beamforming signal-to-noise ratio (SNR). Therefore, several approaches have been proposed to rapidly establish mmWave links [4]–[8]. The usual strategy is to exploit some sort of structure in the unknown channel that aids in link establishment, e.g., sparsity [4], [5] or channel dynamics [6].

MmWaves have applications in cellular systems [2], [9], [10], including fixed wireless access [11], backhaul [8], mobile access [2], [10], and even vehicle-to-everything (V2X) communications [12], [13]. The V2X application is of interest as the sensors on next generation intelligent vehicles may generate up to hundreds of Mbps [14], and the current vehicular communication mechanisms do not support such data-rates. MmWave communication has the potential to provide the required data-rates owing to the large bandwidth. Unfortunately, configuring mmWave links in high mobility is challenging as the link configuration could consume a significant fraction of the channel coherence interval, leaving little time for utilization.

We propose to leverage out-of-band information extracted from lower frequency channels to reduce the overhead of establishing a mmWave link. This is relevant as mmWave systems will likely be deployed in conjunction with lower frequency systems: (i) to provide wide area control signals; and/or (ii) for multi-band communications [15], [16]. The use of low-frequency information is feasible as the spatial characteristics of sub-6 GHz and mmWave channels are similar [17]. To motivate this idea, consider the hypothetical power azimuth spectrum (PAS) of sub-6 GHz and mmWave shown in Fig. 1 (a). The PASs are substantially similar and we refer to this similarity as "*spatial congruence*". We can obtain a coarse estimate of the dominant directions from sub-6 GHz and use it at mmWave. Consider an elementary use case where the aim is to choose a suitable directional beam at mmWave from the candidate beams shown in Fig. 1 (b). The directional beams of the sub-6 GHz system in the strong directions of the channel are shown in Fig. 1 (c). The sub-6 GHz system has wider beams due to a small number of antennas. Given the sub-6 GHz spatial lobes, the candidate beams at mmWave can now be restricted only to those beams that overlap with sub-6 GHz spatial lobes as shown in Fig. 1 (d).

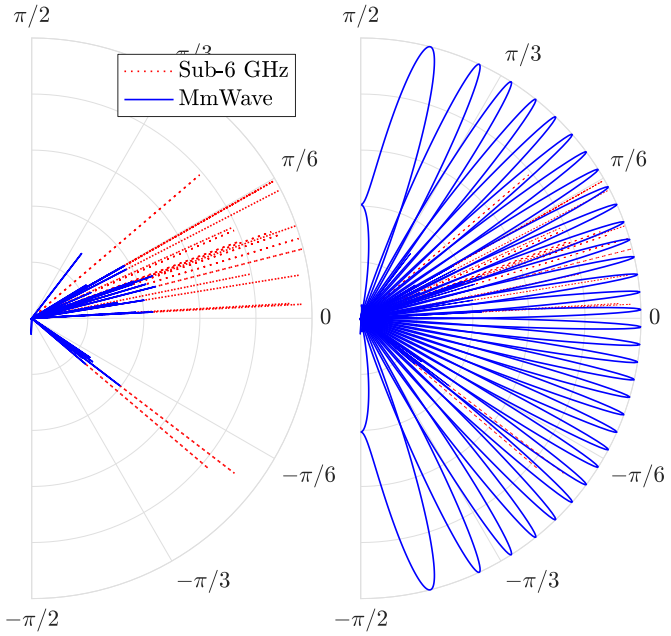
In this work, we use the sub-6 GHz *spatial information* to establish the mmWave link. Specifically, we consider the problem of finding the optimal transmit/receive beam-pair for analog mmWave systems. We assume wideband frequency

This research was partially supported by the U.S. Department of Transportation through the Data-Supported Transportation Operations and Planning (D-STOP) Tier 1 University Transportation Center, by the Texas Department of Transportation under Project 0-6877 entitled "Communications and Radar-Supported Transportation Operations and Planning (CAR-STOP), and by the National Science Foundation under Grant No. 1711702. N. González-Prelcic was supported by the Spanish Government and the European Regional Development Fund (ERDF) under Project MYRADA (TEC2016-75103-C2-2-R).

A. Ali and R. W. Heath Jr. are with the Department of Electrical and Computer Engineering, the University of Texas at Austin, Austin, TX 78712-1687 (e-mail: {anumali,rheath}@utexas.edu).

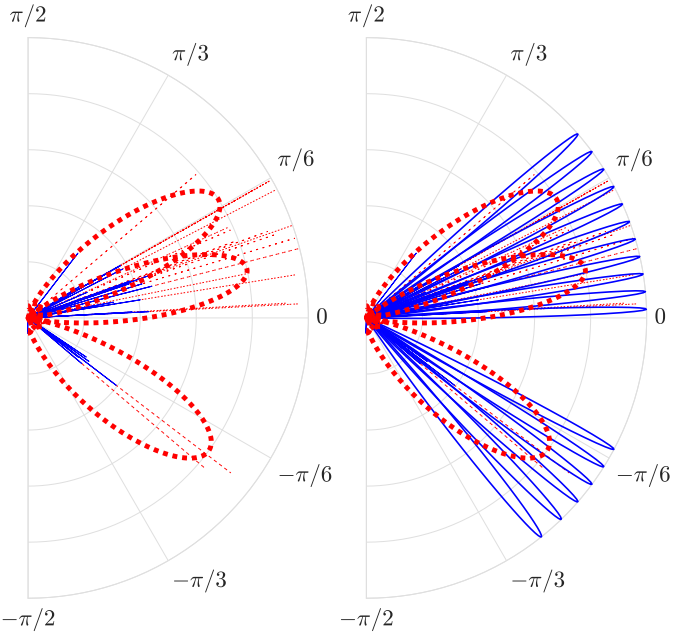
N. González-Prelcic is with the Signal Theory and Communications Department, University of Vigo, Vigo, Spain (e-mail: nuria@gts.uvigo.es).

A preliminary version of this work appeared in the Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, USA, March, 2017 [1].



(a) Dominant paths at sub-6 GHz and mmWave

(b) Candidate beams at mmWave



(c) Beams of sub-6 GHz system in strong channel directions

(d) Candidate beams at mmWave using sub-6 GHz information

Fig. 1: An elementary use case for sub-6 GHz information in mmWave beam-selection.

selective MIMO channels and OFDM signaling for both sub-6 GHz and mmWave systems. The mmWave system uses analog beamforming with quantized phase-shifters, whereas the sub-6 GHz system is fully digital. Both sub-6 GHz and mmWave systems use uniform linear arrays (ULAs) at the transmitter (TX) and the receiver (RX). The main contributions of this work are:

- Exploiting the limited scattering nature of mmWave channels and using the training on one OFDM subcarrier, we formulate the compressed beam-selection problem.
- We outline a strategy to extract spatial information from sub-6 GHz channel. The proposed strategy takes the mmWave beam-codebook design in consideration.
- We use weighted sparse signal recovery [18] to leverage out-of-band information in compressed beam-selection. The weights are chosen based on out-of-band information.
- We propose a structured random codebook design for compressed beam-selection based on out-of-band information. The proposed design enforces the training precoder/combiner patterns to have high gains in the strong channel directions based on out-of-band information.
- We formulate the compressed beam-selection as a multiple measurement vector (MMV) sparse recovery problem [19] to leverage training from all active subcarriers. The MMV based sparse recovery improves the beam-selection by a simultaneous recovery of multiple sparse signals with common support. We extend the weighted sparse recovery and structured codebook design to the MMV case.
- Based on prior work, we draw conclusions about the expected degree of spatial congruence between sub-6 GHz and mmWave channels. Subsequently, we outline a multi-frequency channel model to generate channels that are consistent with frequency-dependent channel behavior observed in prior work. Using this model, we show that the proposed approach can reduce the training overhead of beam-selection considerably.

The rest of the paper is organized as follows: In Section II we review the related work. The system and channel models for mmWave and sub-6 GHz are outlined in Section III. In Section IV, we formulate the compressed beam-selection problem. We outline the proposed out-of-band aided compressed beam-selection approach in Section V. In Section VI, we review the prior work on frequency dependent channel behavior and outline a simulation strategy to generate multi-band frequency dependent channels. The simulation results are presented in Section VII, and Section VIII concludes the paper.

Notation: We use the following notation throughout the paper. Bold lowercase $\mathbf{x}$ is used for column vectors, bold uppercase $\mathbf{X}$ is used for matrices, non-bold letters x, X are used for scalars. $[\mathbf{x}]_m$, $[\mathbf{X}]_{m,n}$, $[\mathbf{X}]_{m,:}$, and $[\mathbf{X}]_{:,n}$, denote mth entry of $\mathbf{x}$, entry at the mth row and nth column of $\mathbf{X}$, mth row of $\mathbf{X}$, and nth column of $\mathbf{X}$, respectively. Superscript T and $*$ represent the transpose and conjugate transpose. $\mathbf{0}$ and $\mathbf{I}$ denote the zero vector and identity matrix respectively. $\mathcal{CN}(\mathbf{x}, \mathbf{X})$ denotes a complex circularly symmetric Gaussian random vector with mean $\mathbf{x}$ and correlation $\mathbf{X}$. We use $\mathbb{E}[\cdot]$, $\|\cdot\|_p$, and $\|\cdot\|_{\mathrm{F}}$ to denote expectation, p norm and Frobenius norm, respectively. $\mathbf{X} \otimes \mathbf{Y}$ is the Kronecker product of $\mathbf{X}$ and $\mathbf{Y}$. Calligraphic letter $\mathcal{X}$ denotes sets and $[X]$ represents the set $\{1, 2, \cdots, X\}$. Finally, $|\cdot|$ is the absolute value of its argument or the cardinality of a set, and $\mathrm{vec}(\cdot)$ yields a vector for a matrix argument. The sub-6 GHz variables are underlined, as $\underline{\mathbf{x}}$, to distinguish them from mmWave.

II. Related Work

Prior work on using out-of-band information in communication systems primarily targets beamforming reciprocity in frequency division duplex (FDD) systems. Based on the observation that the spatial information in the uplink (UL) and downlink (DL) is congruent [20], [21], several strategies were proposed to estimate DL correlation from UL measurements (see e.g., [22], [23] and references therein). The estimated correlation was in turn used for DL beamforming. Along similar lines, in [24] the multi-paths in the UL channel were estimated and subsequently the DL channel was constructed using the estimated multi-paths. In [25], the UL measurements were used as partial support information in compressed sensing based DL channel estimation. The frequency separation between UL and DL is typically small. As an example, there is 9.82% frequency separation between $1935\,\mathrm{MHz}$ UL and $2125\,\mathrm{MHz}$ DL [21]. In essence, the aforementioned strategies were tailored for the case when the percent frequency separation of the channels under consideration is small and spatial information is congruent. We consider channels that can have frequency separation of several hundred percents, and hence some degree of spatial disagreement is expected.

There is some prior work on leveraging out-of-band information for mmWave communications. In [26], the directional information from legacy WiFi was used to reduce the beam-steering overhead of $60\,\mathrm{GHz}$ WiFi. The measurement results presented in [26] confirm the value of out-of-band information for mmWave link establishment. Our work is distinguished from [26] as the techniques developed in this work are applicable to non-line-of-sight (NLOS) channels, whereas [26] primarily considered LOS channels. In [27], the authors study a joint sub-6 GHz-mmWave communication system and solve a scheduling problem over sub-6 GHz and mmWave interfaces to maximize the delay constrained throughput of the mmWave system. Our work, however, focuses on compressed beam-selection in analog mmWave systems using sub-6 GHz information. In [28], the coarse angle estimation at sub-6 GHz followed by refinement at mmWave was pitched. The implementation details and results, however, were not provided. In [29], the mmWave spatial correlation was estimated using sub-6 GHz spatial correlation. The estimated correlation was in turn used for linear channel estimation in SIMO systems. In contrast with [29], we consider MIMO systems and focus on compressed beam-selection in an analog architecture. The concept of radar aided mmWave communication was introduced in [30]. The information extracted from a mmWave radar was used to configure the mmWave communication link. Unlike [30], we use sub-6 GHz communication system's information for mmWave link establishment.

Compressed beam-selection was considered for a narrowband system in [5]. The problem was formulated using codebooks based on sampled array response vectors (i.e., with high-resolution phase-shifters) and did not consider out-of-band information. In contrast, we formulate the compressed beam-selection problem using codebooks based on low-resolution phase-shifters and aid the beam-selection with out-of-band information. We are extending our prior work [1] and provide a detailed treatment of the problem. Herein, we propose and use a frequency dependent clustered channel model, where each cluster contributes multiple rays. The clustered behavior is observed in recent mmWave channel modeling studies [31]. In contrast, [1] assumed that each cluster contributes a single ray. Further, this work considers frequency selective channels for sub-6 GHz and mmWave, whereas [1] assumed narrowband systems. The compressed beam-selection in [1] is based on DFT codebooks, whereas we consider codebooks based on low-resolution phase-shifters. Finally, we present a structured random codebook design that is not present in [1].

III. System and channel model

We consider a multi-band MIMO system shown in Fig. 2, where ULAs of isotropic point sources are used at the TX and the RX. The ULAs are considered for ease of exposition, whereas, the proposed strategies can be extended to other array geometries with suitable modifications. We assume that the sub-6 GHz and mmWave arrays are co-located, aligned, and have comparable apertures. Both sub-6 GHz and mmWave systems operate simultaneously.

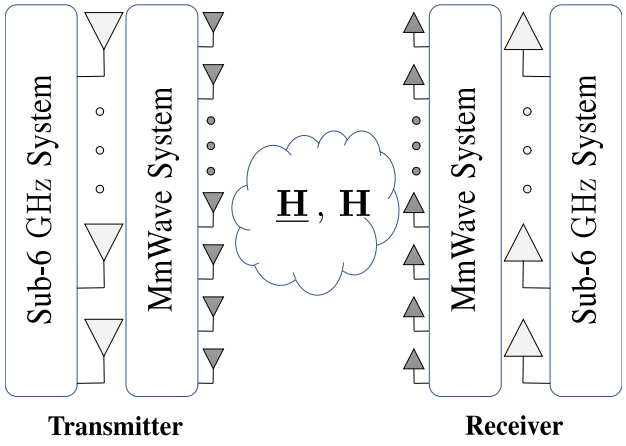


Fig. 2: A multi-band MIMO system with co-located sub-6 GHz and mmWave antenna arrays. The sub-6 GHz channel is $\underline{\mathbf{H}}$ and the mmWave channel is $\mathbf{H}$.

A. Millimeter wave system and channel model

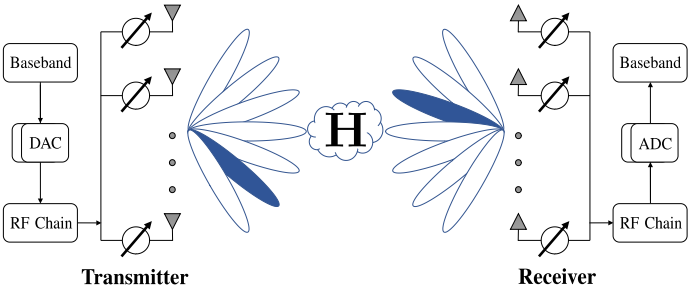


Fig. 3: A mmWave system with phase-shifters based analog beamforming.

The mmWave system is shown in Fig. 3. The TX has M_{TX} antennas and the RX has M_{RX} antennas. Both the TX and

the RX are equipped with a single RF chain, hence only analog beamforming is possible. The idea of using out-of-band information can also be applied to hybrid analog/digital and fully digital low-resolution mmWave architectures, an interesting direction for future work. The mmWave system uses OFDM signaling with K subcarriers. The data symbols $\mathsf{s}[k]$ are transformed to the time-domain using a K-point IDFT. A cyclic prefix (CP) of length L_{c} is then prepended to the time-domain samples before applying the analog precoder $\mathbf{f}$. The length L_{c} CP followed by the K time-domain samples constitute one OFDM block. The effective transmitted signal on subcarrier k is $\mathbf{f}\mathsf{s}[k]$. The data symbols follow $\mathbb{E}[\mathsf{s}[k]\mathsf{s}^*[k]] = \frac{P_{\mathrm{t}}}{K}$, where P_{t} is the total average power in the useful part, i.e., ignoring the CP, per OFDM block. Since $\mathbf{f}$ is implemented using analog phase-shifters, it has constant modulus entries i.e., $|[\mathbf{f}]_m|^2 = \frac{1}{M_{\mathrm{TX}}}$. Further, we assume that the angles of the analog phase-shifters are quantized and have a finite set of possible values. With these assumptions, $[\mathbf{f}]_m = \frac{1}{\sqrt{M_{\mathrm{TX}}}} e^{\mathrm{j}\zeta_m}$, where ζ_m is the quantized angle.

We assume perfect time and frequency synchronization at the receiver. The received signal is first combined using an analog combiner $\mathbf{q}$. The CP is then removed and the time-domain samples are converted back to the frequency-domain using a K-point DFT. If the $M_{\mathrm{RX}} \times M_{\mathrm{TX}}$ MIMO channel at the subcarrier k is denoted as $\mathbf{H}[k]$, the received signal on subcarrier $[k]$ after processing can be expressed as

$$\check{\mathsf{y}}[k] = \mathbf{q}^* \mathbf{H}[k] \mathbf{f} \mathsf{s}[k] + \mathbf{q}^* \check{\mathbf{v}}[k], \tag{1}$$

where $\check{\mathbf{v}}[k] \sim \mathcal{CN}(\mathbf{0}, \sigma_{\check{\mathbf{v}}}^2 \mathbf{I})$.

We adopt a wideband geometric channel model with C clusters. Each cluster has a mean time delay $\tau_c \in \mathbb{R}$ and mean physical AoA/AoD $\{\theta_c, \phi_c\} \in [0, 2\pi)$. Each cluster is further assumed to contribute R_c rays/paths between the TX and the RX. Each ray $r_c \in [R_c]$ has a relative time delay τ_{r_c}, relative AoA/AoD shift $\{\vartheta_{r_c}, \varphi_{r_c}\}$, and complex path gain α_{r_c}. Further, ρ_{pl} denotes the large-scale pathloss and $p(\tau)$ denotes the pulse shaping filter evaluated at point τ. Under this model, the delay-ℓ MIMO channel matrix $\mathbf{H}[\ell]$ can be written as [32]

$$\mathbf{H}[\ell] = \sqrt{\frac{M_{\mathrm{RX}} M_{\mathrm{TX}}}{\rho_{\mathrm{pl}}}} \sum_{c=1}^{C} \sum_{r_c=1}^{R_c} \alpha_{r_c} p(\ell T_{\mathrm{s}} - \tau_c - \tau_{r_c}) \times \mathbf{a}_{\mathrm{RX}}(\theta_c + \vartheta_{r_c}) \mathbf{a}_{\mathrm{TX}}^*(\phi_c + \varphi_{r_c}), \tag{2}$$

where T_{s} is the signaling interval and $\mathbf{a}_{\mathrm{RX}}(\theta)$ and $\mathbf{a}_{\mathrm{TX}}(\phi)$ are the antenna array response vectors of the RX and the TX, respectively. The array response vector of the RX is

$$\mathbf{a}_{\mathrm{RX}}(\theta) = \frac{1}{\sqrt{M_{\mathrm{RX}}}} [1, e^{\mathrm{j}2\pi d \sin(\theta)}, \cdots, e^{\mathrm{j}2\pi (M_{\mathrm{RX}}-1) d \sin(\theta)}]^{\mathsf{T}}, \tag{3}$$

where d is the inter-element spacing in wavelength. The array response vector of the TX is defined in a similar manner. We normalize the array response vectors to have a unit norm and take out the effect of this normalization from the channel (2) by pre-multiplying with $\sqrt{M_{\mathrm{RX}} M_{\mathrm{TX}}}$. This practice ensures that (i) the precoders and combiners based on the array response vector do not need further normalization to be unit norm, and (ii) the channel is normalized irrespective of the number of antennas.

With the delay-ℓ MIMO channel matrix given in (2), the channel at subcarrier k, $\mathbf{H}[k]$ can be expressed as [32]

$$\mathbf{H}[k] = \sum_{\ell=0}^{L-1} \mathbf{H}[\ell] e^{-\mathrm{j}\frac{2\pi k}{K}\ell}, \tag{4}$$

where $L \leq L_{\mathrm{c}} + 1$ is the number of taps in the mmWave channel.

B. *Sub-6* GHz *system and channel model*

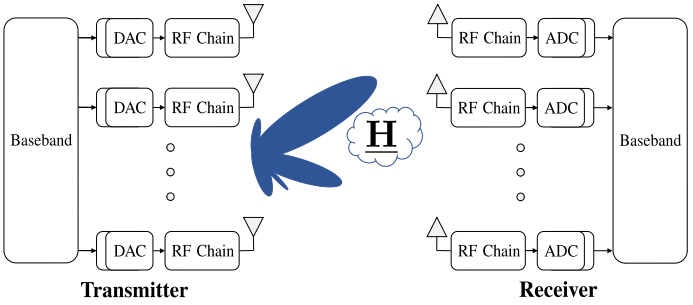


Fig. 4: A sub-6 GHz system with digital precoding.

The sub-6 GHz system is shown in Fig. 4. Note that we underline all sub-6 GHz variables to distinguish them from the mmWave variables. The sub-6 GHz system has one RF chain per antenna and as such, fully digital precoding is possible. The channel model of sub-6 GHz is analogous to mmWave. The sub-6 GHz OFDM system has $\underline{K}$ subcarriers and a CP of length $\underline{L}_{\mathrm{c}}$. The time domain sub-6 GHz channel is thus restricted to have $\underline{L} \leq \underline{L}_{\mathrm{c}} + 1$ taps.

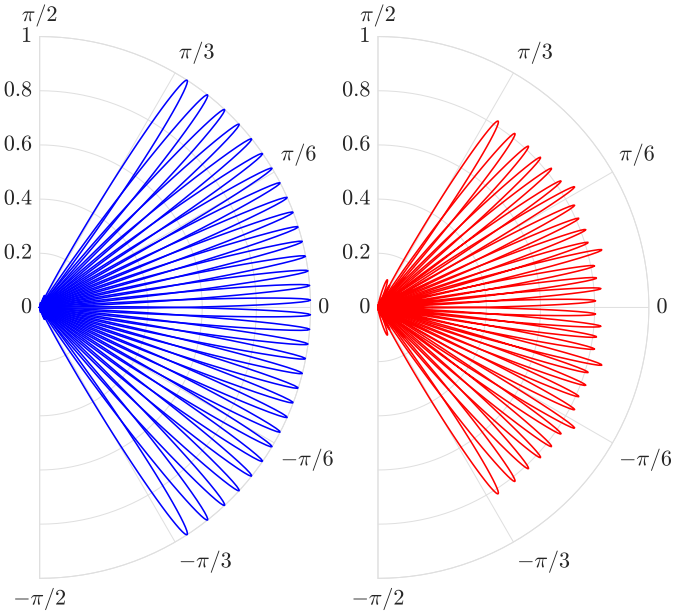
IV. Problem Formulation

In this section, we formulate the compressed beam-selection problem. We discuss the beam codebook design for ULAs using low-resolution phase-shifters. We then present exhaustive beam-selection and the application of sparsity in the beam-selection problem. Subsequently, we formulate the compressed beam-selection problem for the analog mmWave system using training from a single subcarrier. Finally, we extend the proposed formulation to leverage training data from all active subcarriers.

A. *Beam codebook design*

The beam-selection problem for analog mmWave systems is to select the best precoder (and combiner) from the candidate precoders (and combiners). Collectively the candidate precoders are called the precoding codebook. Here we discuss the design of the precoding codebook, but the same design applies to the combining codebook as well. Generating the precoding codebook by sampling the array response vector of the TX array at a few (carefully chosen) directions within the region of interest is a viable choice. Since ULAs produce unequal beamwidth according to the direction - i.e., narrower beams

towards broadside and wider beams towards endfire - separating the sample directions according to the inverse sine is preferable, as it guarantees almost equal gain losses among the adjacent beams [33], [34]. Synthesizing the resulting precoders exactly, however, requires high-resolution phase-shifters. An approximation using D_{TX}-bit phase-shifters can be achieved by quantizing the phase of each element in the precoder to the nearest phase in the set $\{0, \frac{2\pi}{2^{D_{\mathrm{TX}}}}, \cdots, \frac{2\pi(2^{D_{\mathrm{TX}}}-1)}{2^{D_{\mathrm{TX}}}}\}$. Fig. 5 shows the codebook generated using the aforementioned method for a 32 element ULA. The region of interest is a 120° sector spanning the angles $[-\frac{\pi}{3}, \frac{\pi}{3})$ as in [34]. The number of codewords in the codebook is also 32.



(a) Sampled array response vector based codebook (b) 2-bit phase-shifter based approximation

Fig. 5: The codebook based on sampled array response vectors and its approximation using 2-bit phase-shifters.

Here onwards, we will assume that there are G_{TX} precoders in the TX codebook and G_{RX} combiners in the RX codebook. We reserve the notation $\mathbf{w}_n$ for the nth precoder and $\mathbf{z}_m$ for the mth combiner. The $M_{\mathrm{TX}} \times G_{\mathrm{TX}}$ matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{G_{\mathrm{TX}}}]$ collects all the precoders. Similarly, the $M_{\mathrm{RX}} \times G_{\mathrm{RX}}$ matrix $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{G_{\mathrm{RX}}}]$ collects all the combiners. This notational choice for the precoders and combiners is intentionally different from the random precoders and combiners used later in compressed beam-selection.

In the compressed beam-selection problem formulation we will suppose that $\mathbf{W}\mathbf{W}^* \approx \mathbf{I}$, which is true for orthonormal precoders. In Fig. 6 we evaluate this approximation as a function of the number of antennas N_{TX} and the the number of phase-shifter bits D_{TX}, assuming $G_{\mathrm{TX}} = N_{\mathrm{TX}}$. The region of interest is 120° sector spanning the angles $[-\frac{\pi}{3}, \frac{\pi}{3})$. The accuracy of the approximation is tested using the metric $\frac{\|\mathbf{W}\mathbf{W}^* - \mathbf{I}\|_{\mathrm{F}}^2}{\|\mathbf{W}\mathbf{W}^*\|_{\mathrm{F}}^2}$. We can see that the approximation become accurate as D_{TX} increases. The achievable rate results presented in Section VII, however, show good performance even for $D_{\mathrm{TX}} = 2$.

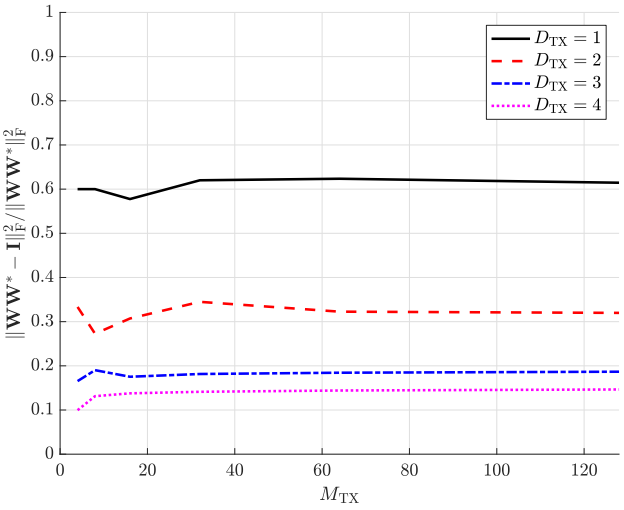


Fig. 6: The metric $\frac{\|\mathbf{W}\mathbf{W}^* - \mathbf{I}\|_{\mathrm{F}}^2}{\|\mathbf{W}\mathbf{W}^*\|_{\mathrm{F}}^2}$ for different phase-shifter bit-resolution and number of antennas M_{TX}.

B. Exhaustive beam-selection

In the training phase, the TX uses a precoding vector $\mathbf{w}_m$ and the RX uses a combining vector $\mathbf{q}_n$. Using the mmWave system model (1), the received signal on the kth subcarrier is

$$\check{\mathsf{y}}_{n,m}^{(\mathrm{E})}[k] = \mathbf{z}_n^* \mathbf{H}[k] \mathbf{w}_m \mathsf{s}_m[k] + \mathbf{z}_n^* \check{\mathbf{v}}_{n,m}^{(E)}[k], \tag{5}$$

where $\mathbf{w}_m \mathsf{s}_m[k]$ is the precoded training symbol on subcarrier k. The superscript (E) on a variable signifies its association to exhaustive beam-selection. The receiver divides through by the training symbol $\mathsf{s}_m[k]$ to get

$$\mathsf{y}_{n,m}^{(\mathrm{E})}[k] = \mathbf{z}_n^* \mathbf{H}[k] \mathbf{w}_m + \mathbf{v}_{n,m}^{(E)}[k], \tag{6}$$

where $\mathbf{v}_{n,m}^{(\mathrm{E})}[k]$ is the post-processing noise after combining and division by the training. The TX transmits the training OFDM blocks on G_{TX} precoding vectors. For each precoding vector, the RX uses G_{RX} distinct combining vectors. The number of total training blocks is $G_{\mathrm{RX}} \times G_{\mathrm{TX}}$. Collecting the signals (6), we get an $G_{\mathrm{RX}} \times G_{\mathrm{TX}}$ matrix

$$\mathbf{Y}^{(\mathrm{E})}[k] = \mathbf{Z}^* \mathbf{H}[k] \mathbf{W} + \mathbf{V}^{(\mathrm{E})}[k]. \tag{7}$$

The largest absolute entry in $\mathbf{Y}^{(\mathrm{E})}[k]$ determines best beam-pair. If we denote $\mathbf{y}^{(\mathrm{E})}[k] = \mathrm{vec}(\mathbf{Y}^{(\mathrm{E})}[k])$, then $r^\star = \underset{1 \le r \le G_{\mathrm{RX}} \times G_{\mathrm{TX}}}{\arg\max} |[\mathbf{y}^{(\mathrm{E})}[k]]_r|$, determines the best beam-pair. Specifically, the best precoder index is $j^\star = \lceil \frac{r^\star}{G_{\mathrm{RX}}} \rceil$, and the best combiner index is $i^\star = r^\star - (j^\star - 1) G_{\mathrm{RX}}$. The receiver needs to feedback the best precoder index to the transmitter, which can be achieved using the active sub-6 GHz link. Note that we did not keep the index $[k]$ with r as the analog precoder and combiner are independent of the subcarrier. Constructing $\mathbf{Y}^{(\mathrm{E})}[k]$ (or $\mathbf{y}^{(\mathrm{E})}[k]$) by exhaustive-search as in (7) incurs a training overhead of $G_{\mathrm{RX}} \times G_{\mathrm{TX}}$ blocks.

C. Sparsity in beam-selection

The crux of compressed beam-selection is to reduce the training overhead of beam-selection by exploiting the spatial

clustering of multi-paths in the channel. To this end, let us re-write the noise free version of (7) as

$$\mathbf{E}[k] = \mathbf{Z}^{*}\mathbf{H}[k]\mathbf{W}. \tag{8}$$

Due to the spatial clustering in the mmWave channel, the matrix $\mathbf{E}[k]$ is sparse. We show an example for a 64×16 MIMO system in Fig. 7. The channel has a single cluster at $\theta = \phi = 0$ with arrival and departure spread of 2°.

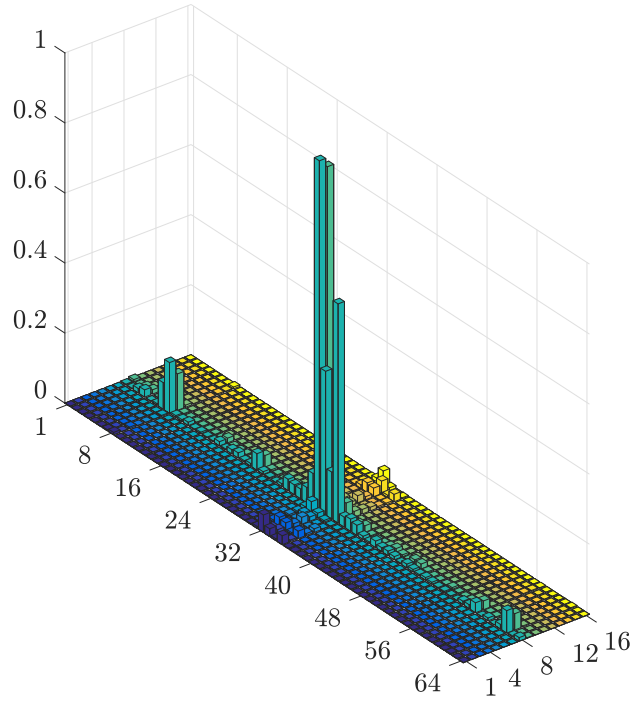


Fig. 7: The matrix $|\mathbf{E}[k]|$ for a 64×16 MIMO system and a single cluster channel with angles $\theta = \phi = 0$ and arrival and departure spread of 2°

Note that depending on the AoA and AoD, even a single path channel will not yield $\mathbf{E}[k]$ with a single non-zero coefficient. This is because any ray illuminates multiple consecutive beams, albeit with reduced power moving from the closest beam to the farthest. With large enough antenna arrays at the transmitter and receiver and a few clusters with small AS in the mmWave channel, the matrix $\mathbf{E}[k]$ can be considered approximately sparse. As such, we can proceed by assuming that $\mathbf{E}[k]$ is a sparse matrix. Note that, in this work, we do not exploit the sparsity in delay-domain, as in e.g., [35]. As such, the time-offset of the incoming rays is irrelevant.

D. Compressed beam-selection

The training burden of beam-selection can be reduced by exploiting the sparsity of $\mathbf{E}[k]$. The resulting framework, called compressed beam-selection, uses a few random measurements to estimate $r^{\star}$. The random training codebooks that respect the analog beamforming constraints were reported in [36], where TX designs its $M_{\mathrm{TX}} \times N_{\mathrm{TX}}$ training codebook such that $[\mathbf{F}]_{n,m} = \frac{1}{\sqrt{M_{\mathrm{TX}}}} e^{\mathrm{j}\zeta_{n,m}}$, where $\zeta_{n,m}$ is randomly and uniformly selected from the set of quantized angles $\{0, \frac{2\pi}{2^{D_{\mathrm{TX}}}}, \cdots, \frac{2\pi(2^{D_{\mathrm{TX}}}-1)}{2^{D_{\mathrm{TX}}}}\}$. The RX similarly designs its $M_{\mathrm{RX}} \times N_{\mathrm{RX}}$ training codebook $\mathbf{Q}$. Similar to (7), we collect all measurements in a $N_{\mathrm{RX}} \times N_{\mathrm{TX}}$ training matrix $\mathbf{Y}[k]$ to get

$$\mathbf{Y}[k] = \mathbf{Q}^{*}\mathbf{H}[k]\mathbf{F} + \mathbf{V}[k], \tag{9}$$

which is further vectorized to set up the following system

$$\mathbf{y}[k] = \mathrm{vec}(\mathbf{Y}[k]) = (\mathbf{F}^{\mathsf{T}} \otimes \mathbf{Q}^{*})\mathrm{vec}(\mathbf{H}[k]) + \mathrm{vec}(\mathbf{V}[k]). \tag{10}$$

At this stage, using the relation $\mathbf{E}[k] = \mathbf{Z}^{*}\mathbf{H}[k]\mathbf{W}$ from (8), and the approximations $\mathbf{Z}\mathbf{Z}^{*} \approx \mathbf{I}$ and $\mathbf{W}^{*}\mathbf{W} \approx \mathbf{I}$, we get $\mathbf{H}[k] \approx \mathbf{Z}\mathbf{E}[k]\mathbf{W}^{*}$. We plug this approximation in (10) to get

$$\begin{aligned}\mathbf{y}[k] &\approx (\mathbf{F}^{\mathsf{T}} \otimes \mathbf{Q}^{*})(\mathbf{W}^{\mathrm{c}} \otimes \mathbf{Z})\mathrm{vec}(\mathbf{E}[k]) + \mathrm{vec}(\mathbf{V}[k]),\\ &\overset{(a)}{=} (\mathbf{F}^{\mathsf{T}} \otimes \mathbf{Q}^{*})(\mathbf{W}^{\mathrm{c}} \otimes \mathbf{Z})\mathbf{e}[k] + \mathrm{vec}(\mathbf{V}[k])\\ &\overset{(b)}{=} \mathbf{\Psi}\mathbf{e}[k] + \mathrm{vec}(\mathbf{V}[k]).\end{aligned} \tag{11}$$

In (11), (a) follows from the notational choice $\mathbf{e}[k] = \mathrm{vec}(\mathbf{E}[k])$ and (b) follows by introducing the sensing matrix $\mathbf{\Psi} = (\mathbf{F}^{\mathsf{T}} \otimes \mathbf{Q}^{*})(\mathbf{W}^{\mathrm{c}} \otimes \mathbf{Z})$. Exploiting the sparsity of $\mathbf{e}$, $r^{\star}$ can be estimated reliably, even when $N_{\mathrm{TX}} \ll G_{\mathrm{TX}}$ and $N_{\mathrm{RX}} \ll G_{\mathrm{RX}}$. The system (11) can be solved for sparse $\mathbf{e}[k]$ using any of the sparse signal recovery techniques. In this work, we use the orthogonal matching pursuit (OMP) algorithm [37]. We outline the working principle of OMP here and refer the interested readers to [37] for details. The OMP algorithm uses a greedy approach in which the support is constructed in an incremental manner. At each iteration, the OMP algorithm adds to the support estimate the column of $\mathbf{\Psi}$ that is most highly correlated with the residual. The measurement vector $\mathbf{y}[k]$ is used as the first residual vector, and subsequent residual vectors are calculated as $\tilde{\mathbf{y}}[k] = \mathbf{y}[k] - \mathbf{\Psi}\hat{\mathbf{e}}[k]$, where $\hat{\mathbf{e}}[k]$ is the least squares estimate of $\mathbf{e}[k]$ on the support estimated so far. As we are interested only in $r^{\star}$, we can find the approximate solution in a single step using the OMP framework, i.e.,

$$r^{\star} = \underset{1 \le r \le G_{\mathrm{RX}}G_{\mathrm{TX}}}{\arg\max} |[\mathbf{\Psi}]_{:,r}^{*}\mathbf{y}[k]|. \tag{12}$$

A single step solution implies low computational complexity of the proposed approach, and makes it suitable for practical implementations.

E. Leveraging data from all active subcarriers

If the unknowns $\mathbf{e}[k]$ were recovered on all subcarriers, a suitable criterion for choosing $r^{\star}$ could be $r^{\star} = \underset{1 \le r \le G_{\mathrm{RX}}G_{\mathrm{TX}}}{\arg\max} \sum_{k \in [K]} |[\mathbf{e}[k]]_r|$. One can recover the vectors $\mathbf{e}[k]$ individually on each subcarrier and then find $r^{\star}$. Instead, we note that the unknown sparse vectors have a similar sparsity pattern i.e., they share an approximately common support [38]. To exploit the common support property, we formulate a joint recovery problem using measurements from all subcarriers. Formally, we collect all vectors $\mathbf{y}[k]$ in a matrix $\bar{\mathbf{Y}}$, which can be written as

$$\begin{aligned}\bar{\mathbf{Y}} &= [\mathbf{y}[1]\ \mathbf{y}[2]\ \cdots\ \mathbf{y}[K]],\\ &= \mathbf{\Psi}\,[\mathbf{e}[1]\ \mathbf{e}[2]\ \cdots\ \mathbf{e}[K]] + [\mathbf{v}[1]\ \mathbf{v}[2]\ \cdots\ \mathbf{v}[K]],\\ &= \mathbf{\Psi}\bar{\mathbf{E}} + \bar{\mathbf{V}}.\end{aligned} \tag{13}$$

The columns of $\bar{\mathbf{E}}$ are approximately jointly sparse, i.e., $\bar{\mathbf{E}}$ has only a few non-zero rows. The sparse recovery problems of the form (13) are referred to as MMV problems. The simultaneous OMP (SOMP) algorithm [19] is an OMP variant tailored for MMV problems. Using SOMP, $r^\star$ can be found as

$$r^\star = \underset{1\leq r\leq G_{\mathrm{RX}}G_{\mathrm{TX}}}{\arg\max} \sum_{k\in[K]} |[\mathbf{\Psi}]^*_{:,r}\mathbf{y}[k]|. \tag{14}$$

The summation over k (i.e., subcarriers) ensures that the measurements from all subcarriers contribute in deciding the best beam-pair.

V. Out-of-band aided compressed beam-selection

The proposed out-of-band aided compressed beam-selection is a two-stage procedure. In the first stage, the spatial information is extracted from sub-6 GHz channel. In the second stage, the extracted information is used for compressed beam-selection.

A. First stage (spatial information retrieval from sub-6 GHz)

The spatial information sought from sub-6 GHz is the dominant spatial directions i.e., AoAs/AoDs. Prior work has considered the specific problem of estimating both the AoAs/AoDs (see e.g., [39]) and the AoAs/AS (see e.g., [40]) from an empirically estimated spatial correlation matrix. The generalization of these strategies to joint AoA/AoD/AS estimation, however, is not straightforward. Further, angle estimation algorithms typically rely on channel correlation knowledge which is difficult to acquire for rapidly varying channels. Therefore, we seek a methodology that can provide reliable spatial information for rapidly varying channels with minimal overhead. For the application at hand, the demand on the accuracy of the direction estimates, however, is not particularly high. Due to the inherent differences between sub-6 GHz and mmWave channels, the extracted spatial information will have an unavoidable mismatch. Consequently, we only need a coarse estimate of the angular information from sub-6 GHz.

We assume that the estimate of the MIMO channel taps $\underline{\hat{\mathbf{H}}}[\ell]$ or equivalently $\underline{\mathbf{H}}[k]$ is available (see e.g., [41] for OFDM channel estimation techniques). The estimate of MIMO channel $\underline{\hat{\mathbf{H}}}[\ell]$ is required for the operation of sub-6 GHz system itself. Hence, the spatial information extraction from the sub-6 GHz channel does not incur any additional training overhead from out-of-band information retrieval point of view. The directional estimate from sub-6 GHz to be used with mmWave beam-selection can be constructed as

$$\underline{\hat{\mathbf{E}}}[k] = \underline{\mathbf{Z}}^*\underline{\mathbf{H}}[k]\,\underline{\mathbf{W}}, \tag{15}$$

where $\underline{\mathbf{W}}$ comprises of sub-6 GHz TX array response vector sampled at the same spatial points as used for mmWave codebook generation. We refer to $|\underline{\hat{\mathbf{E}}}[k]| \in \mathbb{R}^{G_{\mathrm{RX}}\times G_{\mathrm{TX}}}$ as the spatial spectrum. The same procedure is used for constructing $\underline{\mathbf{Z}}$. The spatial spectrum averaged over all sub-6 GHz subcarriers $|\underline{\hat{\mathbf{E}}}|$ is used as out-of-band information in compressed beam-selection. We show the spatial spectrum $|\underline{\hat{\mathbf{E}}}|$ of an 8×2 sub-6 GHz MIMO system for the example considered in Section IV-C in Fig. 8.

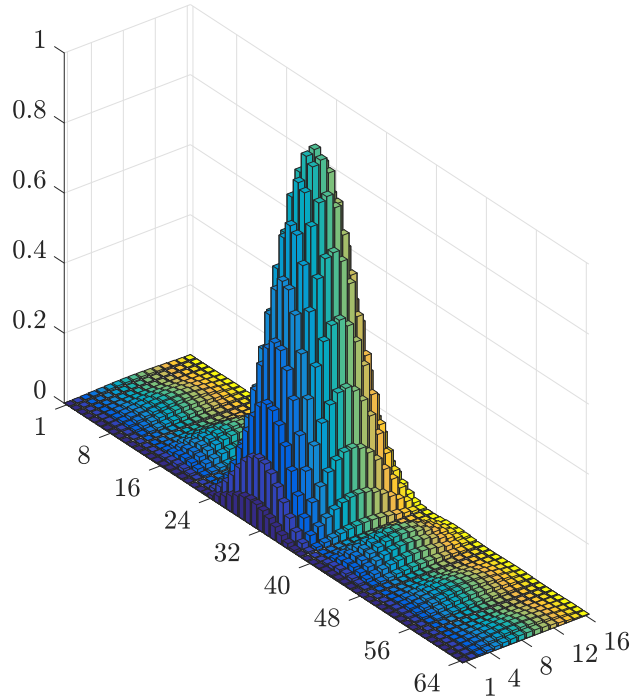


Fig. 8: The 64x16 matrix $|\underline{\hat{\mathbf{E}}}|$ for an 8x2 MIMO channel. The channel has a single cluster channel with angles $\theta = \phi = 0$ and arrival and departure spread of 2°.

The spatial spectrum $|\underline{\hat{\mathbf{E}}}|$ is directly used in weighted sparse signal recovery. The structured random codebook design, however, requires the indices of the dominant sub-6 GHz precoders and combiners. These indices can be easily found by inspecting the spatial spectrum. We collect the indices of the O dominant precoders in the set $\mathcal{J}$ and the set of dominant combiners in the set $\mathcal{I}$.

B. Second stage (out-of-band aided compressed beam-selection)

We explain the out-of-band aided compressed beam-selection in two parts. The first part is the weighted sparse recovery and the second is the structured random codebook design.

Weighted sparse recovery: The OMP based sparse recovery assumes that the prior probability of the support is uniform, i.e., all elements of the unknown can be active with the same probability p. If some prior information about the non-uniformity in the support is available, the OMP algorithm can be modified to incorporate this prior information. In [18] a modified OMP algorithm called logit weighted - OMP (LW-OMP) was proposed for non-uniform prior probabilities. Assume that $\mathbf{p} \in \mathbb{R}^{G_{\mathrm{TX}}G_{\mathrm{RX}}}$ is the vector of prior probabilities. Specifically, the rth element of $\mathbf{e}[k]$ can be active with prior probability $0 \leq [\mathbf{p}]_r \leq 1$. Then $r^\star$ can be found using LW-OMP as

$$r^\star = \underset{1\leq r\leq M_{\mathrm{RX}}M_{\mathrm{TX}}}{\arg\max} |[\mathbf{\Psi}]^*_{:,r}\mathbf{y}[k]| + w([\mathbf{p}]_r), \tag{16}$$

where $w([\mathbf{p}]_r)$ is an additive weighting function. The authors refer the interested reader to [18] for the details of LW-OMP and the selection of $w([\mathbf{p}]_r)$. The general form of $w([\mathbf{p}]_r)$ can be given as $w([\mathbf{p}]_r) = J_{\mathrm{w}} \log \dfrac{[\mathbf{p}]_r}{1-[\mathbf{p}]_r}$, where J_{w} is a constant that depends on sparsity level, the amplitude of the unknown coefficients, and the noise level. In the absence of prior information, (16) can be solved using uniform probability $\mathbf{p} = \delta \mathbf{1}$, where $0 < \delta <= 1$, which is equivalent to solving (12).

The spatial information from sub-6 GHz can be used to obtain a proxy for $\mathbf{p}$. The probability vector $\mathbf{p} \in \mathbb{R}^{G_{\mathrm{RX}} G_{\mathrm{TX}}}$ is obtained using $|\hat{\underline{\mathbf{E}}}| \in \mathbb{R}^{G_{\mathrm{RX}} \times G_{\mathrm{TX}}}$. Let $\hat{\underline{\mathbf{e}}} = \mathrm{vec}(\hat{\underline{\mathbf{E}}})$, then a simple proxy of the probability vector based on the spatial spectrum can be

$$\mathbf{p} = J_{\mathrm{p}} \frac{|\hat{\underline{\mathbf{e}}} - \min(\hat{\underline{\mathbf{e}}})|}{\max(\hat{\underline{\mathbf{e}}}) - \min(\hat{\underline{\mathbf{e}}})}. \tag{17}$$

Initially the spectrum is scaled to meet the probability constraint $0 \leq [\mathbf{p}]_r \leq 1$. The subsequent scaling by$J_{\mathbf{p}} \in (0, 1]$ captures the reliability of out-of-band-information. The reliability is a function of the sub-6 GHz and mmWave spatial congruence, and operating SNR. For highly reliable information, a higher value can be used for J_{p}.

Structured random codebooks: So far we have considered random codebooks that respect the analog hardware constraints, i.e., constant modulus and quantized phase-shifts. The random codebooks used for training, however, can be tailored to out-of-band information. We describe the design of structured codebooks for precoders, but it also applies to the combiners.

Recall that $\mathcal{J}$ is the index set associated with the dominant precoders. Hence $[\mathbf{W}]_{:,\mathcal{J}}$ are the dominant precoders. We construct a super random codebook $\bar{\mathbf{F}}$ containing $\bar{N}_{\mathrm{TX}} \gg N_{\mathrm{TX}}$ codewords according to [36]. The desired random codebook then consists of the N_{TX} codewords from the super codebook that have the highest correlation with the precoder $[\mathbf{W}]_{:,\mathcal{J}}$. The procedure to generate structured precoding codebooks is summarized in Algorithm 1. The LW-OMP algorithm with structured codebooks is referred to as structured LW-OMP.

Algorithm 1 Structured random codebook design

Input: $\mathcal{J}$, $\mathbf{W}$
Output: $\mathbf{F}$

1: Construct a super-codebook $\bar{\mathbf{F}}$ using $\bar{N}_{\mathrm{TX}}$ random codewords generated according to [36].
2: Let $\mathbf{N} = \bar{\mathbf{F}}^* [\mathbf{W}]_{:,\mathcal{J}}$. Populate the index set $\mathcal{M}$ with the indices of N_{TX} rows of $\mathbf{N}$ that have the largest 2-norms.
3: Create the precoding matrix $\mathbf{F} = [\bar{\mathbf{F}}]_{:,\mathcal{M}}$.

The sensing matrices constructed from structured random codebooks and purely random codebooks are expected to have different mutual coherence. Formally, we define mutual coherence of the sensing matrix as $\chi(\mathbf{\Psi}) = \max_{m<n} \frac{|[\mathbf{\Psi}]^*_{:,m} [\mathbf{\Psi}]_{:,n}|}{\|[\mathbf{\Psi}]_{:,m}\|^2 \|[\mathbf{\Psi}]_{:,n}\|^2}$ [42]. We show the mutual coherence as a function of the number of measurements for sensing matrices based on random dictionaries and structured random dictionaries in Fig. 9. We can observe that the mutual coherence of a sensing matrix based on structured random codebooks is higher. From application point of view, however, the structured random codebooks take more meaningful random measurements in the directions that are more likely to be active, and hence can provide gains in compressed beam-selection.

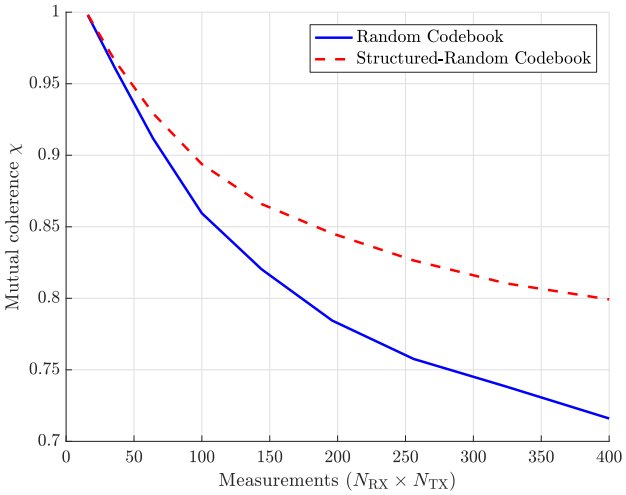


Fig. 9: The mutual coherence χ of sensing matrices based on random codebooks and structured-random codebooks.

Finally, if all active subcarriers are used for the training, out-of-band information can be incorporated in SOMP algorithm via logit weighting and structured codebooks. Specifically, the logit weighted - SOMP (LW-SOMP) algorithm [43] finds $r^\star$ by

$$r^\star = \underset{1 \leq r \leq G_{\mathrm{RX}} G_{\mathrm{TX}}}{\arg\max} \sum_{k \in [K]} |[\mathbf{\Psi}]^*_{:,r} \mathbf{y}[k]| + w([\mathbf{p}]_r). \tag{18}$$

The LW-SOMP algorithm used with structured random codebooks is termed structured LW-SOMP.

VI. Multi-band channel characteristics and simulation

The out-of-band aided mmWave beam-selection strategies proposed in this work rely on the information extracted at sub-6 GHz. Therefore, it is essential to understand the similarities and differences between sub-6 GHz and mmWave channels. Furthermore, to assess the performance of proposed out-of-band aided mmWave link establishment strategies, a simulation strategy is required to generate multi-band frequency dependent channels. In this section, we review a representative subset of prior work to draw conclusions about the expected degree of spatial congruence between sub-6 GHz and mmWave channels. Based on these results, we outline a strategy to simulate multi-band frequency dependent channels.

A. Review of multi-band channel characteristics

The material properties change with frequency, e.g., the relative conductivity and the average reflection increase with frequency [44], [45]. Hence, some characteristics of the channel are expected to vary with frequency. It was reported that

the delay spread decreases [46]–[50], the number of angle-of-arrival (AoA) clusters increase [31], the shadow fading increases [48], and the angle spread (AS) of clusters decreases [47], [49] with frequency. Further, it was observed that the late arriving multi-paths have more frequency dependence due to higher interactions with the environment [51], [52].

Not all channel characteristics vary greatly with frequency. As an example, the existence of spatial congruence between the UL and DL channels is well established [20], [21]. In [20], it was noted that though the propagation channels in UL and DL are not reciprocal, the spatial information is congruent. It was observed in measurements (for $1935\,\mathrm{MHz}$ UL and $2125\,\mathrm{MHz}$ DL) that the deviation in AoAs of dominant paths of UL and DL is small with high probability [21]. Prior work has exploited the spatial congruence between UL and DL channels to reduce/eliminate the feedback in FDD systems, see e.g., [22]–[25].

Some channel characteristics are congruent for larger frequency separations. In [17], the directional power distribution of 5.8 GHz, 14.8 GHz, and 58.7 GHz LOS indoor channels were reported to be almost identical. The number of resolvable paths, the decay constants of the clusters, the decay constants of the subpaths within the clusters, and the number of angle-of-departure (AoD) clusters were found to be similar in 28 and 73 GHz channels [31] in an outdoor scenario. In [53], similar power delay profiles (PDPs) were reported for 10 GHz and 30 GHz indoor channels. The received power as a function of distance was found to be similar for 5.8 GHz and 14.8 GHz in [17]. Only minor differences were observed in the CDFs of delay spread, azimuth AoA/AoD spread, and elevation AoA/AoD spread of six different frequencies between 2 GHz and 60 GHz in the outdoor environments studied in [54].

To the best of authors' knowledge there is no prior work on simultaneous measurements of sub-6 GHz and mmWave vehicular channels. As such, the spatial congruence (or lack of it) for such channels is yet to be established. The existing studies in indoor [17], [53] and outdoor [31], [54], however, confirm that there can be substantial similarity between channels at different frequencies, even with large separations. Hence, it is likely that there is significant, albeit not perfect, congruence between sub-6 GHz and mmWave channels. This observation is leveraged by prior work that used legacy WiFi measurements to configure 60 GHz WiFi links [26].

Due to the differences in the wavelength of sub-6 GHz and mmWave frequencies, it is possible that the Fresnel zone clarity criterion for LOS - e.g., first Fresnel zone 80% obstruction free - is satisfied at sub-6 GHz but, not for mmWave. This would imply that the sub-6 GHz channel is LOS and the mmWave channel is NLOS. It is expected that the out-of-band aided link establishment will not perform well in such scenarios as the spatial information in a LOS sub-6 GHz and NLOS mmWave may be different. In this case, one option is to detect such scenarios and revert to in-band only link establishment. Another option is use machine learning based methods e.g., [55].

B. Simulation of multi-band frequency dependent channels

The following observations are made about the frequency dependent channel behavior from the review of the prior work:

- The channel characteristics differ with frequency, and the differences increase as the percent separation between center frequencies of the channels increase.
- The late arriving multi-paths have more frequency dependence [51], [52].
- Some paths may be present at one frequency but not at the other [56].

The proposed multi-band frequency dependent channel simulation algorithm takes the aforementioned observations into consideration. It takes the parameters of the channels at two frequencies as input and outputs a random realization for each of the two channels. The input parameters include the number of clusters, the number of paths within a cluster, root mean squared (RMS) delay spread, RMS delay spread of the paths within clusters, center frequency, and the RMS AS of the paths within clusters. The output random realizations of the two channels are consistent in the sense that one of the channels is a perturbed version of the other, where the perturbation model respects the frequency dependent channel behavior. Before discussing the proposed simulation algorithm, we present the required preliminaries.

The following exposition is applicable to the channels at two frequencies f_1 and f_2 (not necessarily sub-6 GHz and mmWave). Therefore, we use subscript index $i \in [I]$, where $I = 2$, to distinguish the parameters of the channel at center frequency f_1 from the parameters of the channel at center frequency f_2. We assume that there are C_i clusters in the channel i. Each cluster has a mean time delay $\tau_{c,i}$ and mean physical AoA/AoD $\{\theta_{c,i}, \phi_{c,i}\} \in [0, 2\pi)$. Each cluster c_i is further assumed to contribute $R_{c,i}$ rays/paths between the TX and the RX. Each ray $r_{c,i} \in [R_{c,i}]$ has a relative time delay $\tau_{r_{c,i}}$, relative AoA/AoD shift $\{\vartheta_{r_{c,i}}, \varphi_{r_{c,i}}\}$, and complex path gain $\alpha_{r_{c,i}}$. If $\rho_{\mathrm{pl},i}$ represents the path-loss, then the omni-directional impulse response of the channel i can be written as

$$h_{\mathrm{omni},i}(t,\theta,\phi) = \frac{1}{\sqrt{\rho_{\mathrm{pl},i}}} \sum_{c_i=1}^{C_i} \sum_{r_{c,i}=1}^{R_{c,i}} \alpha_{r_{c,i}} \delta(t - \tau_{c,i} - \tau_{r_{c,i}}) \times \delta(\theta - \theta_{c,i} - \vartheta_{r_{c,i}}) \times \delta(\phi - \phi_{c,i} - \varphi_{r_{c,i}}). \tag{19}$$

The continuous time channel impulse response given in (19) is not band-limited. The impulse response convolved with the pulse shaping filter, however, is band-limited and can be sampled to obtain the discrete time channel as in Section III. Further, in (19) we have only considered the azimuth AoAs/AoDs for simplicity. The general formulation with both azimuth and elevation angles is a straightforward extension, see [31]. A detailed discussion on the choice of the channel parameters is beyond the scope of this paper. The reader is directed to prior work e.g., [57] for discussions on suitable channel parameters. A cursory guideline can be established, however, based on the literature review presented earlier. Assuming $f_1 \leq f_2$ it is expected that $C_1 \geq C_2$ [31],

$R_{c,1} \geq R_{c,2}$ [31], $\tau_{\max,1} \geq \tau_{\max,2}$ [46]–[49], $\sigma_{\vartheta_{c,1}} \geq \sigma_{\vartheta_{c,2}}$, and $\sigma_{\varphi_{c,1}} \geq \sigma_{\varphi_{c,2}}$ [47], [49]. Furthermore, the parameters used for numerical evaluations in Section VII are an example of the parameters that comply with the observations of the prior work.

The proposed channel simulation strategy is based on a two-stage algorithm. In the first stage, the mean time delays $\tau_{c,i}$ and mean AoAs/AoDs $\{\theta_{c,i}, \phi_{c,i}\}$ of the clusters are generated together for both frequencies, while respecting the frequency dependent behavior. In the second stage, the paths within the clusters are generated independently for both frequencies. The first stage of the proposed channel simulation algorithm is outlined in Algorithm 2.

1) First Stage: The number of clusters C_i, the RMS delay spreads $\tau_{\mathrm{RMS},i}$, and the center frequencies f_i for channels $i \in [I]$ are fed to the first-stage of the proposed algorithm as inputs. The first stage has three parts. In the first part, the clusters for both the channels are generated independently. In the second part, we replace several clusters in one of the channels by the clusters of the other channel. The first two parts ensure that there are a few correlated as well as a few independent clusters in the channels. Finally, in the third part frequency dependent perturbations are added to the clusters of one of the channels. This is to imitate the effect that the correlated clusters in the two channels may have a time/angle offset.

Part 1: (Generation) The algorithm initially generates the mean time delays and mean AoAs/AoDs for all the clusters in both channels i.e., $\{\tau_{c,i}, \theta_{c,i}, \phi_{c,i}\}$. The set of the three parameters $\{\tau_{c,i}, \theta_{c,i}, \phi_{c,i}\}$ corresponding to a cluster is referred to as the cluster parameter set. The clusters for both channels are generated independently.

Part 2: (Replacement) We replace several clusters in one channel with the clusters of the other to ensure correlated clusters in the channels. The exact number of replaced clusters varies in each realization. The replacement step is to be carried in accordance with the following observations: (i) the late arriving clustered paths are more likely to fade independently across the two channels [51], [52]; and (ii) independent clustered paths are more likely as the percent frequency separation increases. In other words, for fixed percent frequency separation, the early arriving clustered paths are correlated across the channels with a higher likelihood. We store the indices of correlated clusters in sets $\mathcal{R}_i$, henceforth called the replacement index sets. As an example, the index sets can be created by sorting the cluster parameter sets in an ascending order with respect to $\tau_{c,i}$ and populating $\mathcal{R}_i = \{i : \xi > \frac{|f_i - f_{[[I]\setminus i]}|}{\max(f_i, f_{[[I]\setminus i]})} \frac{\tau_{c,i}}{\tau_{\mathrm{DS},i}}\}$, where $\tau_{\mathrm{DS},i} = \max_c(\tau_{c,i})$ is the delay spread of the channel. Here ξ is a standard Uniform random variable, i.e., $\xi \sim U[0,1]$. For the candidate indices that appear in $\mathcal{R}_1$ and $\mathcal{R}_2$, we replace the corresponding clusters in one channel with those of the other. To be specific, we replace the clusters of the channel with larger delay spread. Hence, we update the cluster parameters sets $\{\tau_{c,b}, \theta_{c,b}, \phi_{c,b}\}$ for all $c_b \in \mathcal{R}_b \cap \mathcal{R}_{[I]\setminus b}$ with $\{\tau_{c,[I]\setminus b}, \theta_{c,[I]\setminus b}, \phi_{c,[I]\setminus b}\}$ for all $c_{[I]\setminus b} \in \mathcal{R}_b \cap \mathcal{R}_{[I]\setminus b}$, where $b = \arg\max_i \tau_{\mathrm{DS},i}$.

Part 3: (Perturbation) So far we have simulated the effect that there will be correlated as well as independent clusters in the channels at two frequencies. Now we need to add frequency dependent perturbation in the clusters of one of the channels to simulate the behavior that correlated clusters can have some time/angle offset. The perturbation should be: (i) proportional to the mean time delay of the cluster [51], [52]; and (ii) proportional to percent center frequency separation. We continue by assuming that the clusters of the channel b are perturbed. A scalar perturbation $\Delta_{c,b}$ is generated for all $c_b \in [C_b]$. The perturbation $\Delta_{c,b}$ is then modified for delays and AoAs/AoDs using deterministic modifiers $g_\tau(\cdot)$, $g_\theta(\cdot)$ and $g_\phi(\cdot)$, respectively. The rationale of using deterministic modifications of the same perturbation for delays and angles is the coupling of these parameters in the physical channels. This is to say that the amount of variation in the mean delay of the cluster, from one frequency to another, is not expected to be independent of the variation in AoA/AoD. Let us define the function

$$q(x,w,y,z) = \begin{cases} 1 & \text{if } x - w < y, \\ -1 & \text{if } x + w > z, \\ \pm 1 \text{ with equal probability} & \text{otherwise.} \end{cases} \tag{20}$$

With this definition, an example perturbation model could be $\Delta_{c,b} \sim U[0,1]$, $g_\tau(\Delta_{c,b}) = q(\tau_{c,b}, \Delta_{c,b}, 0, \tau_{\mathrm{DS},b}) \frac{|f_b - f_{[I]\setminus b}|}{\max(f_b, f_{[I]\setminus b})} \tau_{c,b} \Delta_{c,b}$ and $g_\theta(\Delta_{c,b}) = q(\theta_{c,b}, \Delta_{c,b}, 0, 2\pi) \frac{|f_b - f_{[I]\setminus b}|}{\max(f_b, f_{[I]\setminus b})} \frac{\tau_{c,b}}{\tau_{\mathrm{DS},b}} \Delta_{c,b}$. The modifier g_ϕ can be chosen similar to g_θ. The modified perturbations $g_\tau(\Delta_{c,b})$, $g_\theta(\Delta_{c,b})$, and $g_\phi(\Delta_{c,b})$ are added in $\tau_{c,b}$, $\theta_{c,b}$, and $\phi_{c,b}$, respectively, to obtain the cluster parameters for channel b. Finally, the cluster parameter sets for both channels are returned.

Algorithm 2 Mean time delays $\tau_{c,i}$ and mean AoAs/AoDs $\{\theta_{c,i}, \phi_{c,i}\}$ generation

Input: $C_i, \tau_{\mathrm{RMS},i}, f_i$ for all $i \in [I]$
Output: $\{\tau_{c,i}, \theta_{c,i}, \phi_{c,i}\}$ for all $c_i \in [C_i]$ and $i \in [I]$

1: Draw $\bar{\tau}_{c,i} \sim \tau_{\mathrm{RMS},i} \ln(\mathcal{N}(0,1)), \{\theta_{c,i}, \phi_{c,i}\} \sim U[0, 2\pi)$ for all $c_i \in [C_i]$ and $i \in [I]$. Get $\tau_{c,i} \leftarrow \bar{\tau}_{c,i} - \min_c(\bar{\tau}_{c,i})$. Generate the cluster parameter sets $\{\tau_{c,i}, \theta_{c,i}, \phi_{c,i}\}$ for all $c_i \in [C_i]$ and $i \in [I]$.
2: Populate replacement index sets $\mathcal{R}_i$ for all $i \in [I]$ using a suitable replacement model, and update $\{\tau_{c,b}, \theta_{c,b}, \phi_{c,b}\}$ for all $c_b \in \mathcal{R}_b \cap \mathcal{R}_{[I]\setminus b} \leftarrow \{\tau_{c,[I]\setminus b}, \theta_{c,[I]\setminus b}, \phi_{c,[I]\setminus b}\}$ for all $c_{[I]\setminus b} \in \mathcal{R}_b \cap \mathcal{R}_{[I]\setminus b}$, where $b = \arg\max_i \tau_{\mathrm{DS},i}$.
3: Generate C_b perturbations $\Delta_{c,b}$ and update $\tau_{c,b} \leftarrow \tau_{c,b} + g_\tau(\Delta_{c,b})$, $\theta_{c,b} \leftarrow \theta_{c,b} + g_\theta(\Delta_{c,b})$, and $\phi_{c,b} \leftarrow \phi_{c,b} + g_\phi(\Delta_{c,b})$.

2) Second Stage: Once the parameters $\{\tau_{c,i}, \theta_{c,i}, \phi_{c,i}\}$ for all $c_i \in [C_i]$ and $i \in [I]$ are available, the paths/rays within the clusters are generated independently for both channels in the second stage of the proposed algorithm. The second stage requires the number of paths per cluster $R_{c,i}$, the RMS time spread of the paths within clusters $\sigma_{\tau_{r_{c,i}}}$, and the RMS AS of the relative arrival and departure angle offsets $\{\sigma_{\vartheta_{c,i}}, \sigma_{\varphi_{c,i}}\}$, as input. The output of the second stage are the

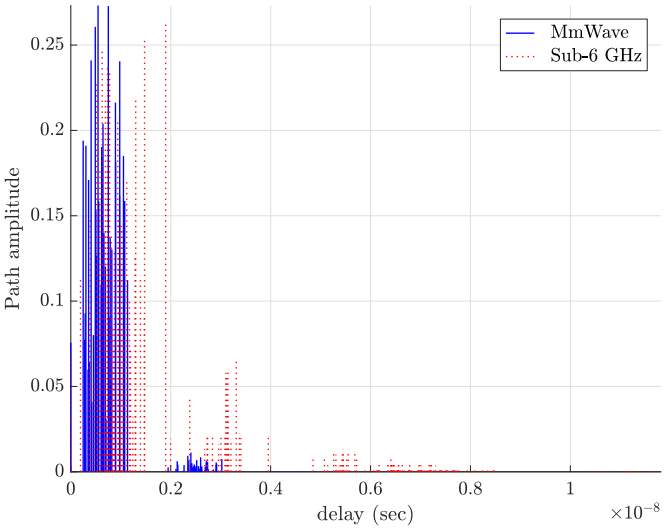
sets $\{\alpha_{r_{c,i}}, \tau_{r_{c,i}}, \vartheta_{r_{c,i}}, \varphi_{r_{c,i}}\}$ for all $r_{c,i} \in R_{c,i}$ and $i \in [I]$. The relative time delays $\tau_{r_{c,i}}$ are generated according to a suitable intra-cluster PDP (e.g., Exponential or Uniform), the relative angle shifts $\{\vartheta_{r_{c,i}}, \varphi_{r_{c,i}}\}$ according to a suitable PAS (e.g., uniform, truncated Gaussian or truncated Laplacian), and the complex coefficients $\alpha_{r_{c,i}}$ according to a suitable fading model (e.g., Rayleigh or Ricean).

VII. Simulation Results

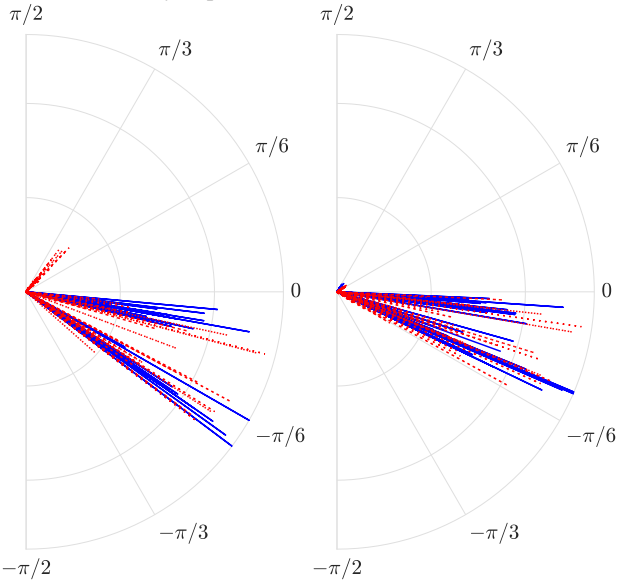
In this section, we present simulation results for the proposed channel simulation strategy and performance of the proposed out-of-band aided mmWave beam-selection strategies. We start by presenting the channel parameters and results, and subsequently present the system parameters and the results for the proposed out-of-band aided mmWave beam-selection strategies.

A. Channel simulation

The sub-6 GHz channel is centered at $\underline{f} = 3.5$ GHz with 150 MHz bandwidth, and the mmWave channel is centered at $f = 28$ GHz with 850 MHz bandwidth. The bandwidths are maximum available bandwidths in the respective bands [58], [59]. The sub-6 GHz and mmWave channels have $\underline{C} = 10$ and $C = 5$ clusters respectively, each contributing $\underline{R}_{\mathrm{c}} = R_{\mathrm{c}} = 20$ rays. The mean AoAs/AoDs of the clusters are limited to $[-\frac{\pi}{3}, \frac{\pi}{3})$. The relative AoA/AoD shifts come from a wrapped Gaussian distribution with with AS $\{\underline{\sigma}_{\underline{\vartheta}_c}, \underline{\sigma}_{\underline{\varphi}_c}\} = 4^\circ$ and $\{\sigma_{\vartheta_c}, \sigma_{\varphi_c}\} = 2^\circ$. As the delay spread of sub-6 GHz channel is expected to be larger than the delay spread of mmWave [46]–[49], we choose $\underline{\tau}_{\mathrm{RMS}} \approx 3.8$ ns and $\tau_{\mathrm{RMS}} \approx 2.7$ ns. The relative time delays of the paths within the clusters are drawn from zero mean normal distributions with RMS AS $\underline{\sigma}_{\underline{\tau}_{r_c}} = \frac{\underline{\tau}_{\mathrm{RMS}}}{10}$ and $\sigma_{\tau_{r_c}} = \frac{\tau_{\mathrm{RMS}}}{10}$. The powers of the clusters are drawn from exponential distributions. Specifically, the exponential distribution with parameter μ is defined as $f(x|\mu) = \frac{1}{\mu} e^{-\frac{x}{\mu}}$. The parameter for sub-6 GHz was chosen as $\underline{\mu} = 0.2$ and for mmWave $\mu = 0.1$. This implies that the power in late arriving multi-paths for mmWave will decline more rapidly than sub-6 GHz. We use the replacement and perturbation models described in Section VI with the angle modifier adjusted to limit the angles in $[-\frac{\pi}{3}, \frac{\pi}{3})$. An example realization of the channel with this configuration is shown in Fig. 10.



(a) Delayed paths at sub-6 GHz and mmWave



(b) AoAs at sub-6 GHz and mmWave (c) AoDs at sub-6 GHz and mmWave

Fig. 10: An example realization generated using the proposed channel generation strategy.

B. Out-of-band aided compressed beam-selection

In this subsection, we present simulation results to test the performance of the proposed out-of-band aided mmWave beam-selection strategies. The sub-6 GHz system has $\underline{M}_{\mathrm{RX}} = 8$ and $\underline{M}_{\mathrm{TX}} = 2$ antennas and the mmWave system has $M_{\mathrm{RX}} = 64$ and $M_{\mathrm{TX}} = 16$ antennas. Both systems use ULAs with half wavelength spacing $\underline{d} = d = 1/2$. The number of sub-6 GHz OFDM subcarriers is $\underline{K} = 32$ and mmWave OFDM subcarriers is $K = 128$. The CP length is quarter quarter the symbol duration for both sub-6 GHz and mmWave. With the chosen operating frequencies, the number of antennas, and the inter-element spacing, the array aperture for sub-6 GHz and mmWave arrays is the same. The transmission power for sub-6 GHz system was set to $\underline{P}_{\mathrm{t}} = 30 - 10\log_{10}(\underline{M}_{\mathrm{TX}})$ dBm per 25 MHz of bandwidth [60]. The subtraction of $10\log_{10}(\underline{M}_{\mathrm{TX}})$ takes care of antenna array gain. The transmission power for mmWave system was set to $P_{\mathrm{t}} = 43 - 10\log_{10}(M_{\mathrm{TX}})$ dBm [61]. These power values are based on Federal Communications Commission (FCC) proposals [60], [61]. The path-loss coefficient at sub-6 GHz and mmWave is 3. The number of taps in sub-6 GHz and mmWave is one more than the length of CP, i.e., $\underline{L} = 9$ and $L = 33$ taps. The raised cosine filter with a roll off factor of 1 is used as a pulse shaping filter.

The metric used for performance comparison is the effective

achievable rate R_{eff} defined as

$$R_{\text{eff}} = \frac{\eta}{EK} \sum_{e=1}^{E} \sum_{k=1}^{K} \log_2 \left(1 + \frac{P_{\text{t}}}{K\sigma_{\check{\mathbf{v}}}^2} |\mathbf{z}_{\hat{i}}^* \mathbf{H}[k] \mathbf{w}_{\hat{j}}|^2 \right), \quad (21)$$

where $\{\hat{i}, \hat{j}\}$ are the estimated transmit and receive codeword indices, E is the number of independent trials for ensemble averaging, $\eta \triangleq \max(0, 1 - \frac{N_{\text{RX}} \times N_{\text{TX}}}{T_{\text{c}}})$, and T_{c} is the channel coherence time in OFDM blocks. With the channel coherence of T_{c} blocks and a training of $N_{\text{RX}} \times N_{\text{TX}}$ blocks, $1 - \frac{N_{\text{RX}} \times N_{\text{TX}}}{T_{\text{c}}}$ is the fraction of time/blocks that are used for data transmission. Thus, η captures the loss in achievable rate due to the training.

In the first experiment, we test the performance of out-of-band aided compressed beam-selection in comparison with in-band only compressed beam-selection. The TX-RX separation for this experiment is fixed at $60\,\text{m}$. The compressed beam-selection is performed using information on a single subcarrier, chosen uniformly at random from the K subcarriers. The number of independent trials is $E = 2000$. The number of measurements for exhaustive-search are fixed at $64 \times 16 = 1024$. The rate results as a function of the number of measurements $N_{\text{RX}} \times N_{\text{TX}}$ are shown in Fig. 11. It can be observed that throughout the range of interest the out-of-band aided compressed beam-selection using structured LW-OMP has a better effective rate in comparison with OMP.

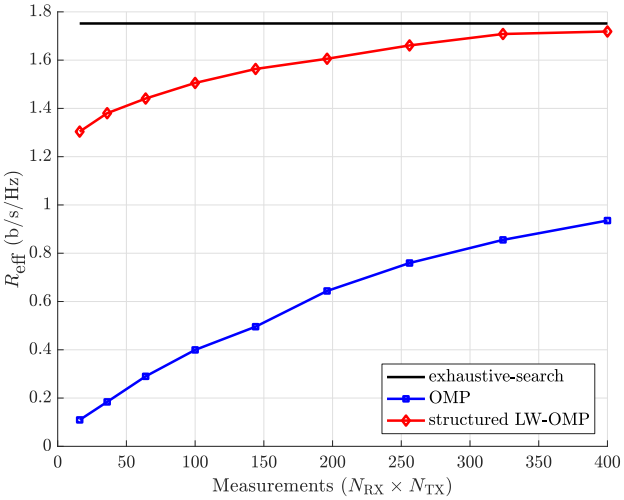


Fig. 11: Effective rate of the structured LW-OMP approach versus the number of measurements $N_{\text{RX}} \times N_{\text{TX}}$ with $60\,\text{m}$ TX-RX separation and $T_{\text{c}} = 128(M_{\text{RX}} \times M_{\text{TX}})$ blocks.

It is observed from Fig. 11 that the effective rate of structured LW-OMP only reaches the rate of exhaustive-search. This, however, is true for large channel coherence T_{c} values. We plot the effective rate of the proposed structured LW-OMP based compressed beam-selection for three channel coherence values in Fig. 12. As the coherence time of the channel decreases, the advantage of the proposed approach becomes significant. As an example, for a medium channel coherence time i.e., $32(M_{\text{RX}} \times M_{\text{TX}})$, the proposed structured LW-OMP based compressed beam-selection can reduce the training overhead of exhaustive-search by over 4x. The gains for smaller channel coherence times are more pronounced. Therefore, the proposed approach is suitable for applications with rapidly varying channels e.g., V2X communications.

To study the fraction of times the proposed approach recovers the best beam-pair, we define and evaluate the success percentage of the proposed approach. The success percentage is defined as

$$\text{SP} = \frac{1}{E} \sum_{e=1}^{E} |\hat{r}^\star \cap \mathcal{B}_N|, \quad (22)$$

where $\hat{r}^\star$ is the index estimated by the proposed approach and $\mathcal{B}_N$ is the set containing the actual indices corresponding to the N best TX/RX beam-pairs. When $N = 1$, the set $\mathcal{B}_1$ has only one element and that is the index corresponding to the beam-pair with the highest receive power. For $N > 1$, the set has N entries that are indices corresponding to the N beam-pairs with the highest receive power. Using a set of indices, instead of the index corresponding to the best beam-pair, generalizes the study and reveals an interesting behavior about selecting one of the better beam-pairs in comparison with selecting the best beam-pair. For now, note that due to the spread of each cluster and the presence of multiple clusters in the mmWave channel, it is possible that the proposed approach does not recover the best beam-pair and still manages to provide a decent effective rate. We populate the set $\mathcal{B}_N$ by performing exhaustive-search in a noiseless channel. We do so as the exhaustive-search in a noisy channel is itself subject to errors. This behavior is revealed in Fig. 13, where the exhaustive-search succeeds $\approx 42\%$ of the times for $\mathcal{B}_1$ and $\approx 58\%$ of the times for $\mathcal{B}_5$. The success percentage of the proposed structured LW-OMP algorithm is $\approx 57\%$ for $\mathcal{B}_1$ and $\approx 30\%$ for $\mathcal{B}_5$. The high success percentage for $\mathcal{B}_5$ is a ramification of having several strong candidate beam-pairs due to cluster spread and the presence of multiple clusters. Note that even though the proposed approach has a (slightly) inferior success percentage for $\mathcal{B}_5$ compared with the exhaustive-search, the training overhead of the proposed approach is significantly lower. With the overhead factored in, the proposed approach is advantageous compared to exhaustive-search as evidenced by the effective rate results in Fig. 12.

Next, we evaluate the performance of structured LW-SOMP based compressed beam-selection using information from all active subcarriers. The TX-RX separation is $60\,\text{m}$. The results of this experiment are shown in Fig. 14. The structured LW-SOMP achieves a better effective rate in comparison with LW-SOMP. Due to the use of training information from all subcarriers, both structured LW-SOMP and LW-SOMP reach the effective rate of exhaustive-search with a handful of measurements. For low channel coherence times T_{c}, the compressed beam-selection approaches, especially out-of-band aided compressed beam-selection, will outperform exhaustive-search.

VIII. Conclusion

In this paper, we used the sub-6 GHz spatial information to reduce the training overhead of beam-selection in an analog mmWave system. We formulated the compressed beam-

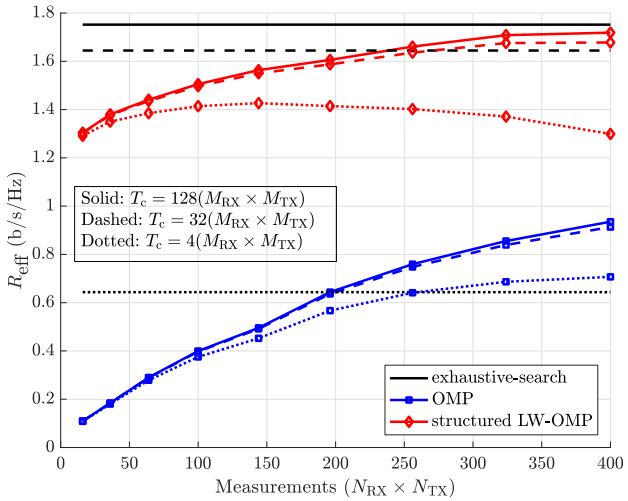


Fig. 12: Effective rate of the structured LW-OMP approach versus the number of measurements $N_{\mathrm{RX}} \times N_{\mathrm{TX}}$ with 60 m TX-RX separation and three different channel coherence times T_{c}.

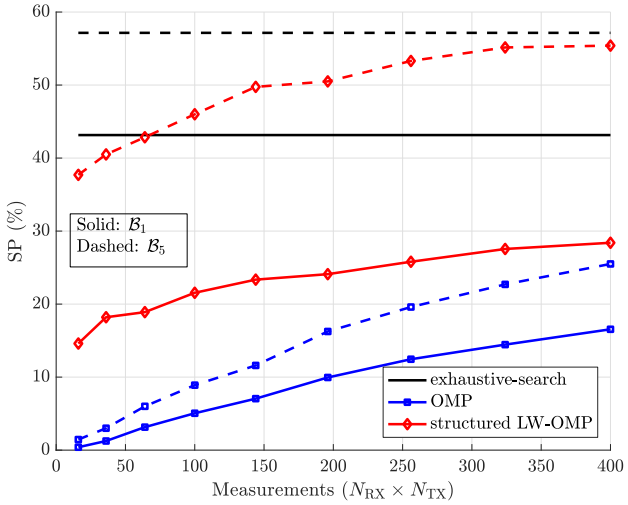


Fig. 13: Success percentage of the structured LW-OMP approach versus the number of measurements $N_{\mathrm{RX}} \times N_{\mathrm{TX}}$ with 60 m TX-RX separation.

selection problem with the codebooks generated from low-resolution phase-shifters. We used a weighted sparse recovery approach with structured random codebooks to incorporate out-of-band information. We proposed a method to generate multi-band frequency dependent channels according to the frequency dependent channel behavior observed in the prior work. We used the proposed multi-band frequency dependent channels to evaluate the achievable rate of the proposed approach. From the rate results, we concluded that the training overhead of in-band only compressed beam-selection can be reduced substantially if out-of-band information is used.

There are several directions for future work. The frequency dependent channel generation strategy can be calibrated with the emerging joint channel modeling results for sub-6 GHz and mmWave. The out-of-band aided beam-selection strategies

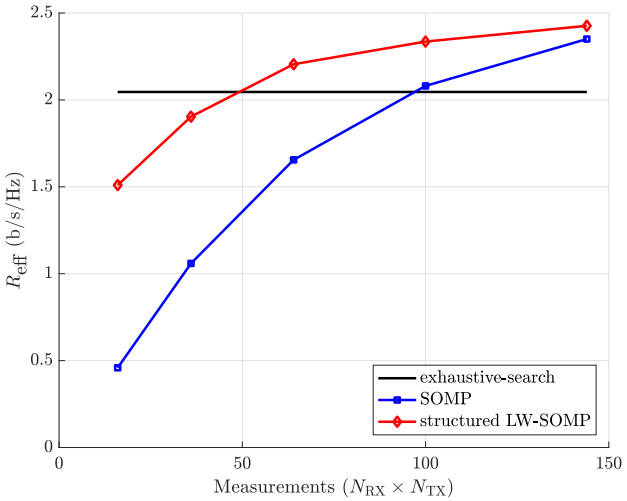


Fig. 14: Effective rate of the structured LW-SOMP approach versus the number of measurements $N_{\mathrm{RX}} \times N_{\mathrm{TX}}$ with 60 m TX-RX separation and $T_{\mathrm{c}} = 32(M_{\mathrm{RX}} \times M_{\mathrm{TX}})$ blocks.

can be explored for arrays other than uniform linear arrays, e.g., circular and planar arrays. Finally, using out-of-band information in hybrid analog/digital and fully digital low-resolution architectures for mmWave systems would also be interesting.

REFERENCES

[1] A. Ali and R. W. Heath Jr., "Compressed beam-selection in millimeter wave systems with out-of-band partial support information," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 3499–3503.

[2] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[3] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[4] A. 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, Oct. 2014.

[5] J. Choi, "Beam selection in mm-Wave multiuser MIMO systems using compressive sensing," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2936–2947, Aug. 2015.

[6] J. Seo *et al.*, "Training beam sequence design for millimeter-wave MIMO systems: A POMDP framework," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1228–1242, Mar. 2016.

[7] J. Wang, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.

[8] S. 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, Oct. 2013.

[9] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[10] T. Bai and R. W. Heath Jr., "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.

[11] Z. Pi, J. Choi, and R. Heath, "Millimeter-wave gigabit broadband evolution toward 5G: fixed access and backhaul," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 138–144, Apr. 2016.

[12] J. Choi *et al.*, "Millimeter wave vehicular communication to support massive automotive sensing," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 160–167, Dec. 2016.

[13] V. Va, J. Choi, and R. W. Heath Jr, "The impact of beamwidth on temporal channel variation in vehicular channels and its implications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5014–5029, Jun. 2017.

[14] A. D. Angelica. (2013) Googles self-driving car gathers nearly 1gb/sec. [Online]. Available: http://www.kurzweilai.net/ googles- self- driving-car- gathers- nearly- 1- gbsec

[15] Y. Kishiyama *et al.*, "Future steps of LTE-A: Evolution toward integration of local area and wide area systems," *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 12–18, Feb. 2013.

[16] R. C. Daniels and R. W. Heath Jr., "Multi-band modulation, coding, and medium access control," IEEE 802.11-07/2780R1, pp. 1–18, Nov. 2007.

[17] M. Peter *et al.*, "Measurement campaigns and initial channel models for preferred suitable frequency ranges," Millimetre-Wave Based Mobile Radio Access Network for Fifth Generation Integrated Communications, Tech. Rep., Mar. 2016.

[18] J. Scarlett, J. S. Evans, and S. Dey, "Compressed sensing with prior information: Information-theoretic limits and practical decoders," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 427–439, Jan 2013.

[19] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, vol. 5, Mar. 2005, pp. 721–724.

[20] T. Asté *et al.*, "Downlink beamforming avoiding DOA estimation for cellular mobile communications," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 1998, pp. 3313–3316.

[21] K. Hugl, K. Kalliola, and J. Laurila, "Spatial reciprocity of uplink and downlink radio channels in FDD systems," May, COST 273 Technical Document TD(02) 066, 2002.

[22] M. Jordan, X. Gong, and G. Ascheid, "Conversion of the spatio-temporal correlation from uplink to downlink in FDD systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2009, pp. 1–6.

[23] A. Decurninge, M. Guillaud, and D. T. Slock, "Channel covariance estimation in massive MIMO frequency division duplex systems," in *Proc. IEEE Glob. Telecom. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[24] D. Vasisht *et al.*, "Eliminating channel feedback in next-generation cellular networks," in *Proc. ACM SIGCOMM*, Aug. 2016, pp. 398–411.

[25] J. Shen *et al.*, "Compressed CSI acquisition in FDD massive MIMO: How much training is needed?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4145–4156, Jun. 2016.

[26] T. Nitsche *et al.*, "Steering with eyes closed: mm-wave beam steering without in-band measurement," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 2416–2424.

[27] M. Hashemi, C. E. Koksal, and N. B. Shroff, "Out-of-band mmwave beamforming and communications to achieve low latency and high energy efficiency in 5G systems," *arXiv preprint arXiv:1701.06241*, 2017.

[28] V. Raghavan *et al.*, "Low-frequency assisted methods for beamforming, timing and frequency offset in mm-wave access systems," May 2016, US Patent 9,337,969. [Online]. Available: https://www.google.com/patents/US9337969

[29] A. Ali, N. González-Prelcic, and R. W. Heath Jr., "Estimating millimeter wave channels using out-of-band measurements," in *Proc. Inf. Theory Appl. (ITA) Wksp*, Feb. 2016, pp. 1–5.

[30] N. González-Prelcic, R. Méndez-Rial, and R. W. Heath Jr., "Radar aided beam alignment in mmwave V2I communications supporting antenna diversity," in *Proc. Inf. Theory Appl. (ITA) Wksp*, Feb. 2016, pp. 1–5.

[31] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 7, pp. 2207–2225, Jul. 2016.

[32] A. Alkhateeb and R. W. Heath Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[33] F. Maschietti *et al.*, "Robust location-aided beam alignment in millimeter wave massive MIMO," *arXiv preprint arXiv:1705.01002*, 2017.

[34] V. Raghavan *et al.*, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.

[35] K. Venugopal *et al.*, "Channel estimation for hybrid architecture based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.

[36] A. Alkhateeb, G. Leus, and R. W. Heath Jr., "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 2909–2913.

[37] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec 2007.

[38] Z. Gao *et al.*, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169–6183, Dec. 2015.

[39] M. L. Bencheikh, Y. Wang, and H. He, "Polynomial root finding technique for joint DOA DOD estimation in bistatic MIMO radar," *Signal Process.*, vol. 90, no. 9, pp. 2723–2730, Sep. 2010.

[40] M. Bengtsson and B. Ottersten, "Low-complexity estimators for distributed sources," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2185–2194, Aug. 2000.

[41] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158–1168, May 2006.

[42] E. J. Candes *et al.*, "Compressed sensing with coherent and redundant dictionaries," *Appl. Computat. Harmonic Anal.*, vol. 31, no. 1, pp. 59–73, 2011.

[43] Z. Li *et al.*, "Compressed sensing reconstruction algorithms with prior information: logit weight simultaneous orthogonal matching pursuit," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, May 2014, pp. 1–5.

[44] ITU, "Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900 MHz to 100 GHz," ITU-R Recommendations, Tech. Rep., 2001.

[45] E. J. Violette *et al.*, "Millimeter-wave propagation at street level in an urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 3, pp. 368–380, May 1988.

[46] R. J. Weiler *et al.*, "Simultaneous millimeter-wave multi-band channel sounding in an urban access scenario," in *Proc. Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2015, pp. 1–5.

[47] A. S. Poon and M. Ho, "Indoor multiple-antenna channel characterization from 2 to 8 GHz," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2003, pp. 3519–3523.

[48] S. Jaeckel *et al.*, "5G channel models in mm-wave frequency bands," in *Proc. Eur. Wireless Conf. (EW)*, May 2016, pp. 1–6.

[49] A. Ö. Kaya, D. Calin, and H. Viswanathan. (2016, Apr.) 28 GHz and 3.5 GHz wireless channels: Fading, delay and angular dispersion.

[50] O. H. Koymen *et al.*, "Indoor mm-Wave channel measurements: Comparative study of 2.9 GHz and 29 GHz," in *Proc. IEEE Glob. Telecom. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[51] R. C. Qiu and I.-T. Lu, "Multipath resolving with frequency dependence for wide-band wireless channel modeling," *IEEE Trans. Veh. Technol.*, vol. 48, no. 1, pp. 273–285, Jan. 1999.

[52] K. Haneda, A. Richter, and A. F. Molisch, "Modeling the frequency dependence of ultra-wideband spatio-temporal indoor radio channels," *IEEE Trans. Antennas Propag.*, vol. 60, no. 6, pp. 2940–2950, Jun. 2012.

[53] D. Dupleich *et al.*, "Simultaneous multi-band channel sounding at mm-Wave frequencies," in *Proc. Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2016, pp. 1–5.

[54] P. Kyösti *et al.*, "Frequency dependency of channel parameters in urban LOS scenario for mmwave communications," in *Proc. Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2016, pp. 1–5.

[55] V. Va *et al.*, "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *arXiv preprint arXiv:1705.05942*, 2017.

[56] K. Haneda, J.-i. Takada, and T. Kobayashi, "Experimental investigation of frequency dependence in spatio-temporal propagation behaviour," in *Proc. Eur. Conf. Antennas Propag. (EuCAP)*, Nov. 2007, pp. 1–6.

[57] V. Nurmela *et al.*, "METIS channel models," Mobile and wireless communications enablers for the twenty-twenty information society, Tech. Rep., Feb. 2015.

[58] [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/DOC-333083A1.pdf

[59] [Online]. Available: http://wireless.fcc.gov/auctions/data/bandplans/lmds.pdf

[60] F. C. Commission *et al.*, "Amendment of the commissions rules with regard to commercial operations in the 3550-3650 MHz band," *GN Docket*, no. 12-354, 2012.

[61] ——, "Fact sheet: spectrum frontiers rules identify, open up vast amounts of new high-band spectrum for next generation (5G) wireless broadband," 2016. [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/DOC-340310A1.pdf