

# Second-Order Statistics-Based Semi-Blind Techniques for Channel Estimation in Millimeter-Wave MIMO Analog and Hybrid Beamforming

Prem Singh, *Student Member, IEEE*, Suraj Srivastava, *Student Member, IEEE*, Aditya K. Jagannatham, *Member, IEEE* and Lajos Hanzo, *Fellow, IEEE*

**Abstract**—Semi-blind (SB) channel estimation is conceived for millimeter wave (mmWave) analog-beamforming (AB) and hybrid-beamforming (HB)-based multiple-input multiple-output (MIMO) systems, which also exploits the data symbols for improving the estimation accuracy. A novel aspect of the proposed framework is that it directly estimates the analog beamformer/ combiner weights without necessitating the estimation of the entire mmWave MIMO channel matrix. By involving powerful matrix perturbation theoretic techniques, a closed-form expression is derived for the mean-squared-error (MSE) of the mmWave-AB-SB algorithm. As a further novelty, our mmWave-HB-SB technique relies on the decomposition of the channel matrix as the product of a decorrelating and a unitary matrix. Subsequently, the former is estimated purely relying on the unknown data symbols, whereas the latter is estimated exclusively from the training vectors. A lower bound on the MSE of the proposed mmWave-HB-SB technique is derived using the constrained Cramér-Rao lower bound (CRLB) framework. Furthermore, the performance gain of our mmWave-HB-SB technique over the conventional purely training-based scheme is also quantified analytically. Our simulation results demonstrate the superiority of the techniques advocated over the existing solutions and also verify the accuracy of our analytical findings.

**Index Terms**—Millimeter wave, MIMO, analog- and hybrid-beamforming, semi-blind channel estimation, CRLB.

## I. INTRODUCTION

MILLIMETER Wave (mmWave) wireless technology has gained significant appeal as a promising candidate for achieving the demanding data rate targets of next-generation wireless networks [1], [2]. The huge bandwidth available in the mmWave frequency band 30–300 GHz [3], together with multiple-input multiple-output (MIMO) technology, facilitates

new applications such as, smart wearables [4], virtual reality (VR) [5] and also augments the capabilities of 5G new radio (5G-NR) devices [2]. However, mmWave signal suffer from the high propagation losses and increased signal blockage. Moreover, the high sampling rate and the related increased power consumption of the radio-frequency (RF) components such as analog-to-digital/ digital-to-analog converters (ADC/ DACs) and the power amplifiers (PAs) prohibit the assignment of an individual RF chain for each antenna [3]. This has motivated researchers to explore practical solutions towards addressing these challenges. A popular low-complexity approach for the design of mmWave transceivers that supports single stream transmission, which is employed in IEEE 802.11 ad systems, is analog beamforming (mmWave-AB) [3]. In such a system, the antenna arrays at the transmitter and receiver are connected to a single RF chain, each via a network of digitally controlled phase-shifters. In order to enhance the capability of mmWave-AB by enabling the transmission of multiple streams, the hybrid beamforming architecture has gained popularity for mmWave MIMO systems [3], [6]–[8]. In such a hybrid system that employs several RF chains, analog precoding/ combining is performed in the RF front-end, while digital processing is carried out in the baseband. Naturally, channel state information (CSI) acquisition plays a key role in the design of both the baseband as well as of the RF precoders and combiners for maximizing the beamforming gain attained [7], [8]. A brief review of the existing contributions on channel estimation for mmWave MIMO systems is presented next.

### A. Review of Existing Works

In the context of beamformer and combiner design for mmWave-AB MIMO systems, beam-training, which is a closed-loop strategy aiming for discovering the angular directions of the strongest signal path between the mmWave receiver and transmitter, has gained widespread popularity. Its low complexity makes it appealing for practical implementation in Wireless HD and IEEE 802.11 ad [9] systems. However, its performance is limited by the number of training beams employed. Moreover, the training overhead of this technique increases linearly with the number of devices, thereby rendering it spectrally inefficient. Similar techniques conceived for mmWave hybrid beamforming (mmWave-HB)

Prem Singh, Suraj Srivastava and Aditya K. Jagannatham are with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, 208016, India (e-mail: {psrawat, srrivast, adityaj}@iitk.ac.in). L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/N004558/1, EP/P034284/1, EP/P034284/1, EP/P003990/1 (COALESCE), of the Royal Society's Global Challenges Research Fund Grant as well as of the European Research Council's Advanced Fellow Grant QuantCom.

This research has been supported by the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, Space Technology Cell, IIT Kanpur, IIMA IDEA Telecom Centre of Excellence, Qualcomm Innovation Fellowship and Arun Kumar Chair Professorship.

MIMO systems initially estimate the channel, followed by the design of the associated analog beamformer and combiner for maximizing the MIMO capacity [10]. A common approach of several treatises in this context is to first estimate the associated parameters such as the angles of arrival (AoAs), angles of departure (AoDs) and the corresponding path-gains of the mmWave MIMO channel. To improve the performance, the authors of [7], [8], [11], [12] proposed sparse signal recovery techniques based on orthogonal matching pursuit (OMP) [11], sparse Bayesian learning and their variants [8] for estimating the mmWave MIMO channel. While generally efficient, an important drawback of these approaches is that their estimation performance is sensitive to the resolution of the grid employed for the quantization of the AoA/ AoD space. Furthermore, the performance of sparse signal recovery schemes is typically also sensitive to the choice of both the sensing matrix as well as of the stopping criterion, which may potentially lead to convergence errors and performance degradation unless suitably chosen. A MUSIC-based algorithm has been proposed in [13] and [14] for mmWave MIMO channel estimation. Although, the scheme therein has been shown to identify multiple paths with a high degree of resolution, its performance depends on factors such as the antenna positions, gains and phase errors. The authors of [15] described an mmWave channel estimation scheme based on estimating the signal parameters via rotational invariance techniques (ESPRIT). To directly apply the ESPRIT method, one has to turn off approximately half of the antennas so that the number of powered antennas is equal to that of the time slots for channel estimation, which will reduce the total transmission power and signal coverage [16]. The authors of [17] proposed an atomic norm minimization (ANM) based semi-blind (SB) (ANM-SB) channel estimator that also exploits the estimated data symbols to enhance the estimation accuracy. This is different from the solutions in [8], [11], which estimate the mmWave MIMO channel by estimating the AoAs/ AoDs and the corresponding complex path gains. However, the scheme of [17] has the following drawbacks: it i) does not exploit the statistical information in the data symbols; ii) requires the knowledge of a bound  $\bar{L}$  on the number of spatial paths; iii) suffers from high computational complexity, which is on the order of  $\mathcal{O}(N_t^2 N_r^2 \bar{L}^2)$ . To reduce the large pilot overhead, the authors of [18], using a low-rank matrix completion technique, proposed the semi-blind data detection and channel estimation methods for a sub-6 GHz multi-user (MU) hybrid massive MIMO systems. Two iterative solutions were presented therein: the regularized alternating least squares (R-ALS) and the bilinear generalized approximate message passing (BiG-AMP). A common limitation of the contributions reviewed above is that they only exploit the training symbols for mmWave MIMO channel estimation, which leads to a reduction in the spectral efficiency. To overcome this, some contributions in the existing literature have described blind channel estimation techniques conceived for mmWave hybrid MIMO systems, which completely eliminate the need for training symbols. For example, reference [19] presented a maximum likelihood formulation aided blind channel estimation scheme using  $l_1$  regularization with ideal as well as one-bit receivers. However, blind schemes are often

computationally complex and may be plagued by convergence problems. A common limitation of the contributions reviewed above is that they only exploit the training symbols for mmWave MIMO channel estimation, which leads to a spectral efficiency reduction. The drawbacks of these techniques motivate us to conceive semi-blind techniques for both mmWave-AB and mmWave-HB MIMO systems, which are capable of significantly enhancing the accuracy of channel estimation by exploiting the statistical properties of the unknown data symbols, along with a limited pilot overhead. A brief summary of our contributions is presented next.

## B. Contributions of the Paper

- We develop a novel semi-blind (SB) channel estimation technique, termed as mmWave-AB-SB, for mmWave-AB MIMO systems communicating over quasi-static channels. In contrast to the conventional ML scheme employed for channel estimation, the proposed mmWave-AB-SB technique directly estimates the dominant left and right singular vectors, without necessitating the estimation of the complete mmWave MIMO channel matrix. Furthermore, the proposed scheme also exploits the statistical properties of the unknown data symbols in addition to a few training symbols, thus leading to a reduced pilot overhead.
- Employing a rigorous matrix perturbation theoretic approach, a closed-form expression is derived for the mean-squared-error (MSE) of the proposed mmWave-AB-SB channel estimator, thereby comprehensively characterizing its performance.
- Next, a semi-blind technique, termed as mmWave-HB-SB, is developed for mmWave-HB MIMO systems communicating over quasi-static channels. The proposed scheme decomposes the channel matrix  $\mathbf{H}$  into a unitary matrix  $\mathbf{U}$  and decorrelating matrix  $\mathbf{D}$ . In contrast to the existing channel estimation solutions, the proposed mmWave-HB-SB scheme offers the following advantages:
  - Unlike the training based conventional ML, and the existing SBL [8], OMP [11] and ANM-SB [17] schemes, i) the proposed scheme exploits both the training as well as the statistical information of the data symbols for the estimation of the component matrices  $\mathbf{U}$  and  $\mathbf{D}$ , respectively, which significantly enhances the accuracy of channel estimation; and ii) the performance of the proposed mmWave-HB-SB scheme also improves with the number of receive antennas.
  - The proposed mmWave-HB-SB technique, i) in contrast to the SBL [8] and OMP [11] sparse channel estimation schemes, does not result in performance degradation in a practical off-grid scenario, where the true AoAs/ AoDs deviate from the set of quantized AoAs/ AoDs corresponding to the beamspace representation; and ii) has a significantly lower computational complexity than the SBL [8], OMP [11] and ANM-SB [17] schemes.

TABLE I: Summary of literature survey on hybrid MIMO channel estimation

	[7]	[11]	[15], [16]	[17]	[18]	[19]	[8]	Proposed
mmWave Band	✓	✓	✓	✓	×	✓	✓	✓
Sub-6 GHz Band	×	×	×	×	✓	×	×	×
AB MIMO	×	×	×	×	×	×	×	✓
HB MIMO	✓	✓	✓	✓	✓	✓	✓	✓
Training-based	✓	✓	✓	✓	✓	✓	✓	✓
Semi-blind	×	×	×	✓	✓	×	×	✓
Blind	×	×	×	×	×	✓	×	×
Theoretical MSE	×	✓	×	×	×	✓	✓	✓
Complexity Analysis	×	×	×	×	✓	×	×	✓
Sensitive to Grid resolution	Yes	Yes	No	No	No	Yes	Yes	No

- The complex constrained Cramér-Rao lower bound (CRLB) is derived for the proposed mmWave-HB-SB channel estimator, which clearly demonstrates that the estimation of the constrained matrix  $\mathbf{U}$  having much fewer free parameters leads to a significantly lower channel estimation MSE for the semi-blind technique. Analytical results are also derived to explicitly characterize the performance gain of our mmWave-HB-SB technique over the conventional training-based ML technique.
- Our simulation results demonstrate the improved performance of the proposed techniques in comparison to that of existing techniques and also validate the accuracy of the analytical results derived.

### C. Organization of the Paper

The next section presents the proposed system model and frame structure, followed by detailing our mmWave-AB-SB scheme conceived for mmWave MIMO scenario. The theoretical analysis characterizing the MSE performance is carried out in Subsection-II-C. Subsequently, Section-III describes our mmWave-HB-SB designed for channel estimation in hybrid beamforming aided mmWave MIMO systems based on a decorrelating-unitary decomposition of the channel matrix. The lower bounds on both the estimation error covariance and on the MSE are also derived in Section-III using the constrained CRLB framework, followed by quantifying the performance gain of the mmWave-HB-SB over the conventional mmWave-HB-ML technique. Our simulation results are provided in Section-V, while Section-VI concludes the paper followed by supplementary proofs of the various results in the Appendices.

### D. Notations

Upper and lower case bold face letters  $\mathbf{A}$  and  $\mathbf{a}$  denote matrices and vectors, respectively. The Frobenius norm of a matrix  $\mathbf{\Delta}$  is denoted as  $\|\mathbf{\Delta}\|_F$ . Furthermore, the notation  $X \sim \mathcal{CN}(0, \sigma^2)$  describes a zero-mean circularly symmetric complex Gaussian random variable  $X$  with mean of zero and variance of  $\sigma^2$ . The matrix Kronecker product is denoted as  $\otimes$ , while the  $\text{vec}(\cdot)$  operator stacks all the columns of a matrix to obtain a single column vector.

## II. MMWAVE ANALOG BEAMFORMING

Consider an mmWave-AB MIMO system with  $N_t$  transmit antennas (TAs) and  $N_r$  receive antennas (RAs), wherein all

the TAs and RAs are connected to a single RF chain via a network of digitally controlled phase-shifters, as shown in Fig. 1. The system model of our mmWave-HB MIMO system relying on hybrid beamforming, and the corresponding channel estimation schemes are described later in Section III. The transmitter is comprised of a cascade of a complex baseband transmit precoder (TPC)  $f_{\text{BB}} \in \mathbb{C}$  and an RF beamformer  $\mathbf{f}_{\text{RF}} \in \mathbb{C}^{N_t \times 1}$ . At the receiver, the signal is first processed using an RF combiner  $\mathbf{w}_{\text{RF}} \in \mathbb{C}^{N_r \times 1}$ , followed by a baseband combiner  $w_{\text{BB}} \in \mathbb{C}$ . Furthermore, as described in [3], the RF beamformer  $\mathbf{f}_{\text{RF}}$  and the RF combiner  $\mathbf{w}_{\text{RF}}$  are implemented using digitally controlled phase-shifters. The elements of the RF precoder and combiner are constrained to have constant magnitude as  $|\mathbf{f}_{\text{RF}}(i)| = \frac{1}{\sqrt{N_t}}$  and  $|\mathbf{w}_{\text{RF}}(j)| = \frac{1}{\sqrt{N_r}}$ ,  $\forall i, j$ . The symbol at the output of the baseband combiner for the system given in Fig. 1 can be mathematically formulated as

$$y = w_{\text{BB}}^* \mathbf{w}_{\text{RF}}^H \mathbf{H} \mathbf{f}_{\text{RF}} f_{\text{BB}} x + v, \quad (1)$$

where  $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$  represents the mmWave MIMO channel matrix,  $x \in \mathbb{C}$  denotes the transmit symbol and  $v \in \mathbb{C}$  is the independent and identically distributed (i.i.d.) additive noise having the distribution of  $\mathcal{CN}(0, \sigma_v^2)$ .

Let the singular value decomposition (SVD) of the mmWave MIMO channel matrix  $\mathbf{H}$  be given as  $\mathbf{H} = \mathbf{S} \mathbf{\Gamma} \mathbf{Q}^H$ , where  $\mathbf{S} \in \mathbb{C}^{N_r \times N_r}$  and  $\mathbf{Q} \in \mathbb{C}^{N_t \times N_t}$  are the left and right singular matrices, respectively, and  $\mathbf{\Gamma} \in \mathbb{C}^{N_r \times N_t}$  contains singular values  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$  along the principal diagonal with  $r = \text{rank}(\mathbf{H})$ . It is widely recognized [20] that the optimal unconstrained digital beamformer and combiner that maximize the SNR at the output, denoted by  $\mathbf{f}^{\text{opt}} \in \mathbb{C}^{N_t \times 1}$  and  $\mathbf{w}^{\text{opt}} \in \mathbb{C}^{N_r \times 1}$ , respectively, are given by  $\mathbf{f}^{\text{opt}} = \mathbf{q}_1$  and  $\mathbf{w}^{\text{opt}} = \mathbf{s}_1$ , where  $\mathbf{s}_1 \in \mathbb{C}^{N_r \times 1}$  and  $\mathbf{q}_1 \in \mathbb{C}^{N_t \times 1}$  are the dominant left and right singular vectors of  $\mathbf{H}$ , and are given by the first columns of the matrices  $\mathbf{S}$  and  $\mathbf{Q}$ , respectively. Hence, in order to maximize the throughput of the system, it is essential to obtain accurate estimates of the singular vectors  $\mathbf{q}_1$  and  $\mathbf{s}_1$ . The conventional procedure for estimation of  $\mathbf{q}_1$  and  $\mathbf{s}_1$  is described next, followed by our more efficient semi-blind technique proposed in this paper.

### A. Conventional Beamformer Estimation

Traditionally, for the purpose of channel estimation,  $M_P$  training beams are transmitted in each frame, as shown in Fig. 2. Note that channel estimation is performed in each frame for both the training and semi-blind schemes. Let the  $m$ th

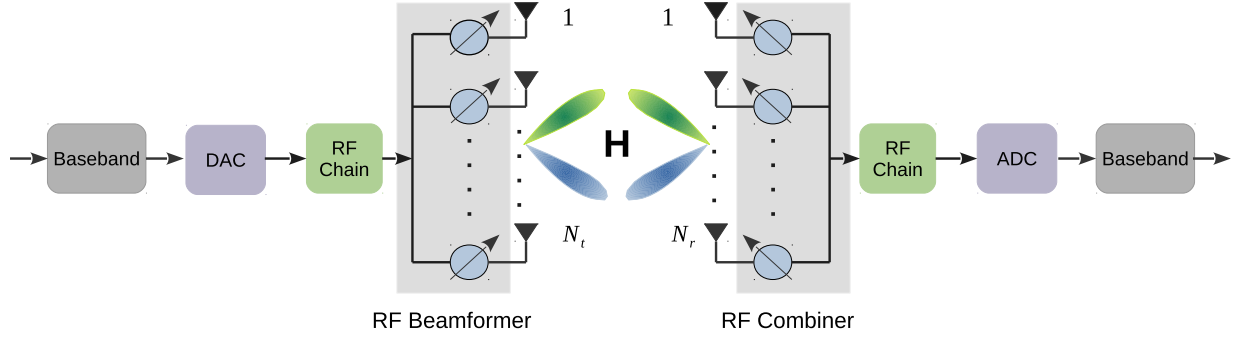


Fig. 1: Schematic of mmWave-AB MIMO Systems.

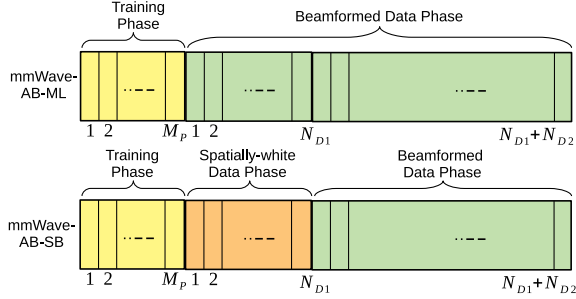


Fig. 2: Frame structure of the mmWave-AB-ML and the proposed mmWave-AB-SB channel estimation schemes.

training beam,  $1 \leq m \leq M_P$ , employ the RF beamformer  $\mathbf{f}_{\text{RF},m} \in \mathbb{C}^{N_t \times 1}$ . The corresponding pilot beam  $\tilde{\mathbf{y}}_m$  reaching the receiver is given by

$$\tilde{\mathbf{y}}_m = \mathbf{H}\mathbf{f}_{\text{RF},m}x_m + \tilde{\mathbf{v}}_m, \quad (2)$$

where  $x_m$  denotes a complex-valued pilot symbol drawn from a phase-shift-keying (PSK) constellation of power  $\mathbb{E}[|x_m|^2] = P_p$ . The received pilot matrix  $\tilde{\mathbf{Y}} \in \mathbb{C}^{N_r \times M_P}$  obtained by concatenating all the  $M_P$  training beams is given by

$$\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_{M_P}] = \mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{X} + \tilde{\mathbf{V}}, \quad (3)$$

where the concatenated RF TPC matrix is  $\mathbf{F}_{\text{RF}} = [\mathbf{f}_{\text{RF},1}, \mathbf{f}_{\text{RF},2}, \dots, \mathbf{f}_{\text{RF},M_P}] \in \mathbb{C}^{N_t \times M_P}$ , and the noise matrix  $\tilde{\mathbf{V}} = [\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \dots, \tilde{\mathbf{v}}_{M_P}] \in \mathbb{C}^{N_r \times M_P}$ . The pilot matrix  $\mathbf{X} \in \mathbb{C}^{M_P \times M_P}$  in (3) is given as  $\mathbf{X} = \text{diag}(x_1, x_2, \dots, x_{M_P})$ . Let  $Q_C$  denote the number of beams combined and  $\mathbf{w}_{\text{RF},q} \in \mathbb{C}^{N_r \times 1}$  represent the RF combiner weights employed in the  $q$ th beam. The output pilot vector  $\tilde{\mathbf{y}}_q \in \mathbb{C}^{1 \times M_P}$  can be determined from (3) as

$$\tilde{\mathbf{y}}_q = \mathbf{w}_{\text{RF},q}^H \mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{X} + \mathbf{w}_{\text{RF},q}^H \tilde{\mathbf{V}}. \quad (4)$$

The output pilot matrix  $\mathbf{Y} = [\tilde{\mathbf{y}}_1^T, \tilde{\mathbf{y}}_2^T, \dots, \tilde{\mathbf{y}}_{Q_C}^T]^T \in \mathbb{C}^{Q_C \times M_P}$  obtained by stacking  $\tilde{\mathbf{y}}_q$  for  $q = 1, 2, \dots, Q_C$ , can be expressed as

$$\mathbf{Y} = \mathbf{W}_{\text{RF}}^H \mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{X} + \mathbf{V}, \quad (5)$$

where the RF combiner matrix  $\mathbf{W}_{\text{RF}} = [\mathbf{w}_{\text{RF},1}, \mathbf{w}_{\text{RF},2}, \dots, \mathbf{w}_{\text{RF},Q_C}] \in \mathbb{C}^{N_r \times Q_C}$  and the stacked

noise matrix  $\mathbf{V} \in \mathbb{C}^{Q_C \times M_P}$  is given by  $\mathbf{V} = \mathbf{W}_{\text{RF}}^H \tilde{\mathbf{V}}$ . The conventional maximum-likelihood (ML) estimate  $\hat{\mathbf{H}}_{c,\text{AB}}$  of the mmWave-AB MIMO channel matrix  $\mathbf{H}$  can be obtained using (5) as

$$\hat{\mathbf{H}}_{c,\text{AB}} = \arg \min_{\mathbf{H}} \|\mathbf{Y} - \mathbf{W}_{\text{RF}}^H \mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{X}\|_F^2. \quad (6)$$

Solving (6) for finding the training-based ML estimate  $\hat{\mathbf{H}}_{c,\text{AB}}$  of  $\mathbf{H}$  may be obtained by the well-known least-squares (LS) solution. In practice, the RF beamformer and the receiver combiner matrices  $\mathbf{F}_{\text{RF}}$  and  $\mathbf{W}_{\text{RF}}$  are frequently set as the submatrices formed by the  $N_t$  and  $N_r$  rows of the discrete Fourier transform (DFT) matrices of size  $M_P \times M_P$  and  $Q_C \times Q_C$ , respectively [8], [11]. Since the elements  $\mathbf{F}_{\text{RF}}$  and  $\mathbf{W}_{\text{RF}}$  are normalized to have magnitude  $\frac{1}{\sqrt{N_t}}$  and  $\frac{1}{\sqrt{N_r}}$  respectively, it follows that  $\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{RF}}^H = \mathbf{I}_{N_t}$  and  $\mathbf{W}_{\text{RF}}\mathbf{W}_{\text{RF}}^H = \mathbf{I}_{N_r}$ . Interestingly, this choice of  $\mathbf{F}_{\text{RF}}$  and  $\mathbf{W}_{\text{RF}}$  satisfies the constraint of constant magnitude elements required for the RF precoder as well as combiner and ensures minimum MSE [21]. The simplified expression of the mmWave-AB-ML estimate for such a setting is given by

$$\hat{\mathbf{H}}_{c,\text{AB}} = \frac{1}{P_p} \mathbf{W}_{\text{RF}} \mathbf{Y} \mathbf{X}^H \mathbf{F}_{\text{RF}}^H. \quad (7)$$

The estimate of the digital beamformer and combiner vectors, denoted as  $\hat{\mathbf{q}}_{1,c}$  and  $\hat{\mathbf{s}}_{1,c}$ , respectively, can now be obtained via the SVD of the mmWave MIMO channel estimate  $\hat{\mathbf{H}}_{c,\text{AB}}$ . It follows from the invariance property of the ML estimate [22] that the estimates  $\hat{\mathbf{q}}_{1,c}$  and  $\hat{\mathbf{s}}_{1,c}$  obtained from  $\hat{\mathbf{H}}_{c,\text{AB}}$  are also the ML estimates of  $\mathbf{q}_1$  and  $\mathbf{s}_1$ , respectively.

While the conventional training-based scheme described above is robust, a drawback of this approach is that it is necessary to estimate the entire mmWave MIMO channel matrix  $\mathbf{H}$  for the sake of estimating the beamforming vectors  $\mathbf{q}_1$  and  $\mathbf{s}_1$ . This assumes greater significance in the context of mmWave MIMO systems deployed in next-generation networks, since the size of the antenna arrays employed is typically large. Furthermore, the conventional techniques for estimating the mmWave MIMO channel employ exclusively training symbols, and do not exploit the statistical information pertaining to the data symbols, which renders them inefficient. To overcome these limitations, a semi-blind technique is proposed next for analog beamforming based mmWave MIMO systems, which yields the estimates of  $\mathbf{q}_1$  and  $\mathbf{s}_1$  without necessitating the

estimation of the complete channel matrix  $\mathbf{H}$ . Furthermore, in addition to training symbols, it also exploits the second-order statistical properties of the data symbols, which leads to a significant improvement in the estimation accuracy.

### B. Proposed mmWave-AB-SB Channel Estimation Scheme

The proposed mmWave-AB-SB scheme divides the data transmission into two phases, wherein the first phase comprises  $N_{D1}$  spatially-white data vectors, followed by  $N_{D2}$  beamformed data vectors in the subsequent phase. Note that  $M_P < N_{D1} \ll N_{D2}$ . Fig. 2 shows the schematic of the proposed frame structure of our mmWave-AB-SB scheme, wherein the training is comprised of  $M_P$  training vectors, similar to the mmWave-AB-ML scheme. The proposed mmWave-AB-SB scheme utilizes the  $M_P$  training vectors for estimating the beamforming vector  $\mathbf{q}_1$ , whereas  $N_{D1}$  spatially-white data vectors are used for estimating the combiner vector  $\mathbf{s}_1$ . Subsequently, the SB estimates of  $\mathbf{q}_1$  and  $\mathbf{s}_1$  are used for the transmission and detection of the  $N_{D2}$  beamformed data vectors. The procedure of the proposed mmWave-AB-SB estimator is described next.

Let  $\mathbf{f}_i = [f_i(1), f_i(2), \dots, f_i(N_t)]^T \in \mathbb{C}^{N_t \times 1}$  denote the  $i$ th spatially-white data vector, where  $i = 1, 2, \dots, N_{D1}$ . The elements  $f_i(p)$  and  $f_i(q)$  of the vector  $\mathbf{f}_i$  are drawn independently from a zero mean  $\mathcal{M}$ -ary PSK constellation, i.e.  $\mathbb{E}[f_i(p)] = 0$ ,  $1 \leq p \leq N_t$ , with average power  $\mathbb{E}[|f_i(p)|^2] = \frac{P_d}{N_t}$ . Thus, these elements satisfy the property

$$\mathbb{E}[f_i(p)f_i^*(q)] = \begin{cases} \mathbb{E}[f_i(p)]\mathbb{E}[f_i^*(q)] = 0, & \text{for } p \neq q \\ \mathbb{E}[|f_i(p)|^2] = \frac{P_d}{N_t}, & \text{for } p = q. \end{cases} \quad (8)$$

Therefore, it follows that  $\mathbb{E}[\mathbf{f}_i\mathbf{f}_i^H] = \frac{P_d}{N_t}\mathbf{I}_{N_t}$ . Recall that since the RF beamformer  $\mathbf{f}_{\text{RF}}$  in (1) is comprised of constant-magnitude phase shifters, its  $j$ th element  $f_{\text{RF}}(j)$ ,  $\forall 1 \leq j \leq N_t$ , can be directly set as  $f_i(j)$  for the transmission of the  $i$ th spatially-white information data vector. Since the MIMO channel is unknown prior to estimation, the RF precoder and combiner are typically reconfigured several times during channel estimation [7], [8], [11] to excite all the angular modes. However, in practice, this might lead to an increased switching delay for reconfiguration of the phase shifters. Therefore, the corresponding output signal  $\mathbf{y}_i \in \mathbb{C}^{Q_c \times 1}$ , similar to (5), can be expressed as

$$\mathbf{y}_i = \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{f}_i + \mathbf{W}_{\text{RF}}^H \mathbf{v}_i, \quad (9)$$

where  $\mathbf{v}_i \in \mathbb{C}^{N_r \times 1}$  is the AWGN noise vector that has a covariance of  $\mathbb{E}[\mathbf{v}_i\mathbf{v}_i^H] = \sigma_v^2\mathbf{I}_{N_r}$ . Employing (8), (9) and exploiting the property of  $\mathbf{W}_{\text{RF}}\mathbf{W}_{\text{RF}}^H = \mathbf{I}_{N_r}$ , the covariance matrix  $\mathbf{R}_y = \mathbb{E}[\mathbf{W}_{\text{RF}}\mathbf{y}_i(\mathbf{W}_{\text{RF}}\mathbf{y}_i)^H] \in \mathbb{C}^{N_r \times N_r}$  of the processed output can be evaluated as

$$\mathbf{R}_y = \frac{P_d}{N_t} \mathbf{H} \mathbf{H}^H + \sigma_v^2 \mathbf{I}_{N_r}. \quad (10)$$

It follows from (10) that the dominant left singular vector  $\mathbf{s}_1$  equals the dominant eigenvector of the covariance matrix  $\mathbf{R}_y$ . The proposed mmWave-AB-SB technique initially obtains the estimate  $\hat{\mathbf{R}}_y$  of the covariance matrix  $\mathbf{R}_y$  from the received  $N_{D1}$  spatially-white data vectors  $\mathbf{y}_i$  as  $\hat{\mathbf{R}}_y =$

$\frac{1}{N_{D1}} \sum_{i=1}^{N_{D1}} \mathbf{W}_{\text{RF}} \mathbf{y}_i (\mathbf{W}_{\text{RF}} \mathbf{y}_i)^H$ . The estimate of  $\mathbf{s}_1$  is in turn evaluated as  $\hat{\mathbf{s}}_{1,s} = \hat{\mathbf{S}}(:, 1)$ , where the estimate  $\hat{\mathbf{S}}$  is derived as  $\hat{\mathbf{S}}\hat{\mathbf{\Gamma}}^2\hat{\mathbf{S}}^H = \text{SVD}(\hat{\mathbf{R}}_y)$ . Thus, the estimate  $\hat{\mathbf{s}}_{1,s}$  can be obtained purely from the blind data symbols. Furthermore, as  $N_{D1}$  increases, a near-perfect estimate of  $\mathbf{s}_1$  can be obtained, since we have  $\hat{\mathbf{R}}_y \rightarrow \mathbf{R}_y$ . Subsequently, the estimate  $\hat{\mathbf{q}}_{1,s}$  of the dominant right singular vector  $\mathbf{q}_1$  can be obtained exclusively from the training-based channel estimation model in (5) as follows. Let the matrix  $\mathbf{Y}$  in (5) be processed as

$$\tilde{\mathbf{y}} = \hat{\mathbf{s}}_{1,s}^H \mathbf{W}_{\text{RF}} \mathbf{Y}.$$

Substituting the output pilot matrix  $\mathbf{Y}$  from (5), one obtains

$$\tilde{\mathbf{y}} = \hat{\mathbf{s}}_{1,s}^H \underbrace{\mathbf{W}_{\text{RF}} \mathbf{W}_{\text{RF}}^H}_{\mathbf{I}_{N_r}} \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X} + \hat{\mathbf{s}}_{1,s}^H \mathbf{W}_{\text{RF}} \mathbf{V}.$$

Replacing the channel  $\mathbf{H}$  with its SVD  $\mathbf{H} = \mathbf{S}\mathbf{\Gamma}\mathbf{Q}^H$ , the vector  $\tilde{\mathbf{y}}$  can be recast as

$$\begin{aligned} \tilde{\mathbf{y}} &= \hat{\mathbf{s}}_{1,s}^H \mathbf{S} \mathbf{\Gamma} \mathbf{Q}^H \mathbf{F}_{\text{RF}} \mathbf{X} + \hat{\mathbf{s}}_{1,s}^H \mathbf{W}_{\text{RF}} \mathbf{V} \\ &= \sigma_1 \mathbf{q}_1^H \mathbf{F}_{\text{RF}} \mathbf{X} + \hat{\mathbf{s}}_{1,s}^H \mathbf{W}_{\text{RF}} \mathbf{V}, \end{aligned} \quad (11)$$

where the simplification  $\hat{\mathbf{s}}_{1,s}^H \mathbf{S} \mathbf{\Gamma} \mathbf{Q}^H = \sigma_1 \mathbf{q}_1^H$  above assumes that the blind estimate  $\hat{\mathbf{s}}_{1,s}$  is sufficiently close to  $\mathbf{s}_1$ , which is valid when  $N_{D1}$  is sufficiently large. The estimate  $\hat{\mathbf{q}}_{1,s}$  can now be obtained as the solution of the constrained optimization problem

$$\hat{\mathbf{q}}_{1,s} = \arg \min_{\|\mathbf{q}_1\|=1} \|\tilde{\mathbf{y}} - \sigma_1 \mathbf{q}_1^H \mathbf{F}_{\text{RF}} \mathbf{X}\|^2. \quad (12)$$

The estimate  $\hat{\mathbf{q}}_{1,s}$  is given by the result below.

*Lemma 1:* For  $\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{RF}}^H = \mathbf{I}_{N_t}$  and  $\mathbf{W}_{\text{RF}}\mathbf{W}_{\text{RF}}^H = \mathbf{I}_{N_r}$ , the solution to the above constrained optimization problem is given by

$$\hat{\mathbf{q}}_{1,s} = \frac{\mathbf{F}_{\text{RF}} \mathbf{X} \tilde{\mathbf{y}}^H}{\|\mathbf{F}_{\text{RF}} \mathbf{X} \tilde{\mathbf{y}}^H\|}. \quad (13)$$

*Proof:* Given in Appendix A.

Thus, the mmWave-AB-SB estimator exploits the unknown data symbols in addition to the known training symbols, making it semi-blind in nature. It is also worth noting that the proposed mmWave-AB-SB technique only utilizes the  $M_P$  training beams for estimating the complex vector  $\mathbf{q}_1$ , which has much fewer parameters than the channel matrix  $\mathbf{H}$ . This naturally leads to a significantly improved estimation accuracy in comparison to its conventional mmWave-AB-ML counterpart, as also supported by the rigorous MSE analysis presented next. Finally, the semi-blind estimate of the RF beamformer and combiner, denoted by  $\hat{\mathbf{q}}_{1,s}^{\text{con}}$  and  $\hat{\mathbf{s}}_{1,s}^{\text{con}}$  respectively, can be derived using the hybrid design by alternating minimization (HD-AM) technique of [23] as

$$\hat{\mathbf{q}}_{1,s}^{\text{con}} = \hat{\mathbf{q}}_{1,s} \circ |\hat{\mathbf{q}}_{1,s}|, \quad \text{and} \quad \hat{\mathbf{s}}_{1,s}^{\text{con}} = \hat{\mathbf{s}}_{1,s} \circ |\hat{\mathbf{s}}_{1,s}|, \quad (14)$$

where the operation  $\circ$  represents the normalization of each entry of the vector by its magnitude resulting in a vector of constant-magnitude elements.

### C. MSE for the mmWave-AB-SB Channel Estimation Scheme

This section presents an outline of the procedure and the key results pertaining to the MSE in the estimate  $\hat{\mathbf{q}}_{1,s}$ . Due to the intractability of the MSE analysis of the beamformer estimation in the mmWave-AB-SB scheme, this section employs key approximation results from rigorous matrix perturbation theory described in [24]. For ease of reading, the technical details of the various derivations are given in Appendices B, C and D. We begin by noting that there are two sources of error in the estimate of  $\mathbf{q}_1$ . First, the noise in the observed training output matrix  $\mathbf{Y}$  in (5), and second, the error in the imperfect estimate of the vector  $\mathbf{s}_1$  due to the noisy data vectors  $\mathbf{y}_i$  in (9). Let the estimate  $\hat{\mathbf{q}}_{1,s}$  be expressed as the function  $\hat{\mathbf{q}}_{1,s} = g(\mathbf{Y}, \hat{\mathbf{s}}_{1,s})$ . Note that we have  $\mathbf{q}_1 = g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \mathbf{s}_1)$ . Using the first-order Taylor series expansion, it follows that

$$\begin{aligned} \hat{\mathbf{q}}_{1,s} - \mathbf{q}_1 &= g(\mathbf{Y}, \hat{\mathbf{s}}_{1,s}) - g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \mathbf{s}_1) \\ &\approx \left[ g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \hat{\mathbf{s}}_{1,s}) - g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \mathbf{s}_1) \right] \\ &\quad + \left[ g(\mathbf{Y}, \mathbf{s}_1) - g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \mathbf{s}_1) \right]. \end{aligned} \quad (15)$$

The first term in the above expression represents the error in  $\hat{\mathbf{q}}_{1,s}$  arising from the noisy training outputs with perfect knowledge of  $\mathbf{s}_1$ . On the other hand, the second term above can be attributed to the estimation error in  $\hat{\mathbf{q}}_{1,s}$  with noise-free training outputs. Furthermore, these two terms are statistically independent, since the pilot output noise in (5) and the estimation error in  $\hat{\mathbf{s}}_{1,s}$  arising from the noise in the blind outputs are assumed to be statistically independent, having come from temporally separated noise vectors. Upon exploiting the above property, we have

$$\mathbb{E} \left[ \|\hat{\mathbf{q}}_{1,s} - \mathbf{q}_1\|^2 \right] = E_1 + E_2, \quad (16)$$

where  $E_1 = \mathbb{E} \left[ \|g(\mathbf{Y}, \mathbf{s}_1) - g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \mathbf{s}_1)\|^2 \right]$  and

$$E_2 = \mathbb{E} \left[ \|g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \hat{\mathbf{s}}_{1,s}) - g(\mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}, \mathbf{s}_1)\|^2 \right].$$

The expressions of  $E_1$  and  $E_2$  are simplified next.

1) *MSE in  $\hat{\mathbf{q}}_{1,s}$  with Perfect Knowledge of  $\hat{\mathbf{s}}_{1,s}$* : Let  $\tilde{\mathbf{q}}_{1,s}$  be defined as

$$\tilde{\mathbf{q}}_{1,s} \triangleq \mathbf{F}_{\text{RF}} \mathbf{X} \mathbf{Y}^H \mathbf{W}_{\text{RF}}^H \mathbf{s}_1 / \sigma_1 P_p.$$

Employing the relationship  $\text{SVD}(\mathbf{H}) = \mathbf{S} \mathbf{\Gamma} \mathbf{Q}^H$  and (5), the above expression can be simplified to

$$\tilde{\mathbf{q}}_{1,s} = \mathbf{q}_1 + \frac{\mathcal{E}^H \mathbf{s}_1}{\sigma_1}, \quad (17)$$

where the matrix  $\mathcal{E}$  is defined as

$$\mathcal{E} = \frac{\mathbf{W}_{\text{RF}} \mathbf{V} \mathbf{X}^H \mathbf{F}_{\text{RF}}^H}{P_p}. \quad (18)$$

From (13), it follows that the estimate  $\hat{\mathbf{q}}_{1,s} = \tilde{\mathbf{q}}_{1,s} / \|\tilde{\mathbf{q}}_{1,s}\|$ . Ignoring the second-order terms, the quantity  $\|\tilde{\mathbf{q}}_{1,s}\|^2$  can be approximated as

$$\|\tilde{\mathbf{q}}_{1,s}\|^2 \simeq 1 + \frac{\mathbf{s}_1^H \mathcal{E} \mathbf{q}_1 + \mathbf{q}_1^H \mathcal{E}^H \mathbf{s}_1}{\sigma_1}. \quad (19)$$

Given (19) and  $\hat{\mathbf{q}}_{1,s} = \tilde{\mathbf{q}}_{1,s} / \|\tilde{\mathbf{q}}_{1,s}\|$ , the expression of  $\hat{\mathbf{q}}_{1,s}$  can be approximated as

$$\hat{\mathbf{q}}_{1,s} \simeq \left( \mathbf{q}_1 + \frac{\mathcal{E}^H \mathbf{s}_1}{\sigma_1} \right) \left[ 1 - \frac{1}{2\sigma_1} \left( \mathbf{s}_1^H \mathcal{E} \mathbf{q}_1 + \mathbf{q}_1^H \mathcal{E}^H \mathbf{s}_1 \right) \right]. \quad (20)$$

Using the above expression, as shown in Appendix C, the MSE in the estimate  $\hat{\mathbf{q}}_{1,s}$  for this scenario can be evaluated as

$$\begin{aligned} E_1 &= \mathbb{E} \left[ \|\hat{\mathbf{q}}_{1,s} - \mathbf{q}_1\|^2 \right] \\ &= \mathbb{E} \left[ \frac{\|\mathcal{E}^H \mathbf{s}_1\|^2}{\sigma_1^2} + \frac{|\mathbf{s}_1^H \mathcal{E} \mathbf{q}_1 + \mathbf{q}_1^H \mathcal{E}^H \mathbf{s}_1|^2}{4\sigma_1^2} \right] \\ &= \frac{\sigma_v^2}{2P_p \sigma_1^2} (2N_t - 1). \end{aligned} \quad (21)$$

2) *MSE in  $\hat{\mathbf{q}}_{1,s}$  with Noise-Free Training*: Let us now consider the noise-free training scenario, i.e.,  $\mathbf{V} = \mathbf{0}$  in (5). Since the columns of the matrix  $\mathbf{S}$  form an orthonormal-basis for the space  $\mathbb{C}^{N_r}$ , one can express  $\hat{\mathbf{s}}_{1,s} = \mathbf{S} \mathbf{c}$ , for some  $\mathbf{c} = [1 + \Delta c_1, \Delta c_2, \dots, \Delta c_{N_r}]^T \in \mathbb{C}^{N_r \times 1}$ . Let the vector  $\check{\mathbf{q}}_{1,s}$  be defined as  $\check{\mathbf{q}}_{1,s} \triangleq \frac{1}{P_p} \mathbf{F}_{\text{RF}} \mathbf{X} \mathbf{y}^H$ . Substituting  $\mathbf{V} = \mathbf{0}$  in (11), and replacing  $\hat{\mathbf{s}}_{1,s}$  by  $\mathbf{S} \mathbf{c}$ , the expression of the vector  $\check{\mathbf{q}}_{1,s}$  may be simplified to  $\check{\mathbf{q}}_{1,s} = \mathbf{Q} \mathbf{\Gamma} \mathbf{c}$ . Thus, the unit-norm vector  $\hat{\mathbf{q}}_{1,s}$  can be expressed as  $\hat{\mathbf{q}}_{1,s} = \frac{\check{\mathbf{q}}_{1,s}}{\|\check{\mathbf{q}}_{1,s}\|} = \mathbf{Q} \tilde{\mathbf{c}}$ , where  $\tilde{\mathbf{c}} \triangleq [\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_{N_t}]^T = \frac{\mathbf{\Gamma} \mathbf{c}}{\sqrt{\mathbf{c}^H \mathbf{\Gamma}^2 \mathbf{c}}}$ . As shown in Appendix D, the squared-error in  $\hat{\mathbf{q}}_{1,s}$  can be written as

$$\begin{aligned} \|\hat{\mathbf{q}}_{1,s} - \mathbf{q}_1\|^2 &= |\tilde{c}_1 - 1|^2 + |\tilde{c}_2|^2 + \dots + |\tilde{c}_{N_t}|^2 \\ &\approx 2[1 - \Re(\tilde{c}_1)] \end{aligned} \quad (22)$$

$$\approx \sum_{i=2}^{N_r} \frac{\sigma_i^2}{\sigma_1^2} |\Delta c_i|^2, \quad (23)$$

where  $\Re(\cdot)$  represents the real part of the quantity. Finally, using the result for  $\mathbb{E}[|\Delta c_i|^2]$  from Appendix D, the MSE component  $E_2$  may be simplified to

$$\begin{aligned} E_2 &= \mathbb{E} \left[ \|\hat{\mathbf{q}}_{1,s} - \mathbf{q}_1\|^2 \right] = \sum_{i=2}^{N_r} \frac{\sigma_i^2}{N_{D1} (\sigma_1^2 - \sigma_i^2)^2 \sigma_1^2} \\ &\quad \left( \frac{\sigma_i^2 \sigma_i^2}{N_{D1}} + \frac{N_t \sigma_v^2 (\sigma_1^2 + \sigma_i^2)}{P_d} + \frac{N_t^2 \sigma_v^4}{P_d^2} \right). \end{aligned} \quad (24)$$

It follows from (24) that the MSE in  $\hat{\mathbf{q}}_{1,s}$  with noise-free training is inversely proportional to the number of data symbols  $N_{D1}$ . Therefore, as  $N_{D1} \rightarrow \infty$ , the MSE in (16) approaches the result derived in (21). Finally, substituting the expressions for  $E_1, E_2$  from (21), (24), respectively, into (16), yields the expression for the net MSE corresponding to the SB estimate of  $\hat{\mathbf{q}}_{1,s}$ .

### III. HYBRID BEAMFORMING

This section presents a semi-blind estimation scheme for our mmWave-HB MIMO system having  $N_t$  TAs and  $N_r$  RAs, as shown schematically in Fig. 3. The system comprises  $N_{\text{RF}} \ll \min(N_t, N_r)$  RF chains both at the transmitter

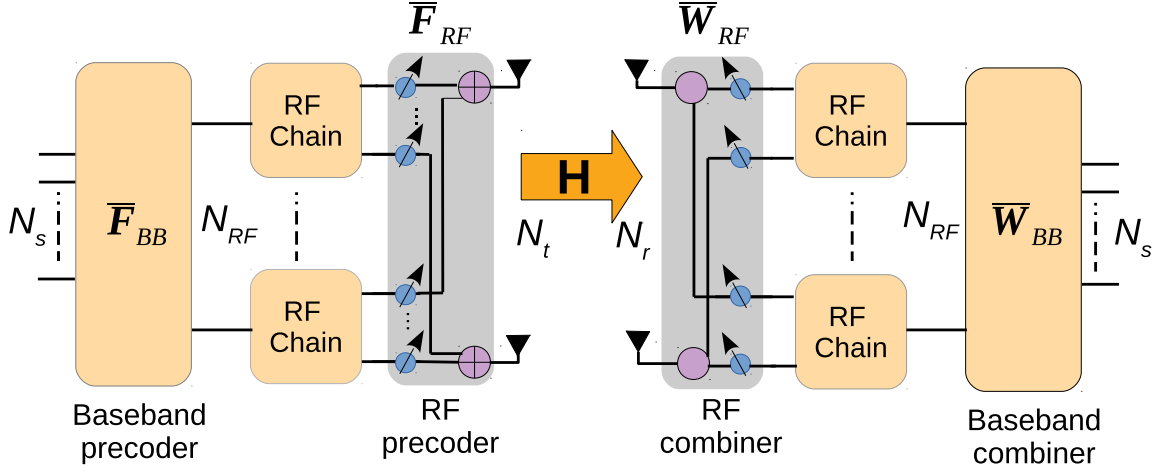


Fig. 3: Schematic Diagram for mmWave-HB MIMO systems.

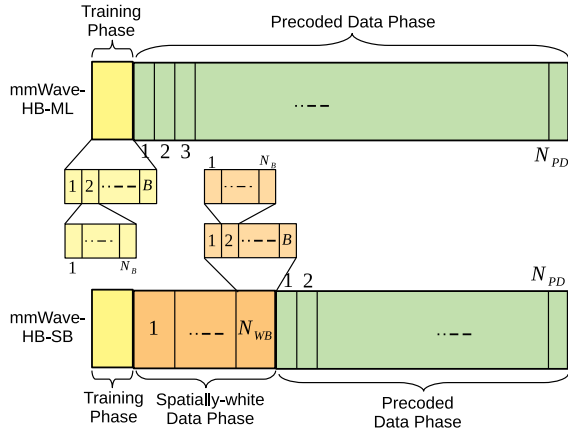


Fig. 4: Frame structures of the mmWave-HB-ML and the proposed mmWave-HB-SB channel estimation schemes.

and receiver for conveying  $N_s \leq N_{RF}$  data streams. The transmitter employs the baseband TPC  $\bar{\mathbf{F}}_{BB} \in \mathbb{C}^{N_{RF} \times N_s}$  and RF TPC  $\bar{\mathbf{F}}_{RF} \in \mathbb{C}^{N_t \times N_{RF}}$  for hybrid beamforming. The receiver processes the received signal using a cascade of a baseband combiner  $\bar{\mathbf{W}}_{BB} \in \mathbb{C}^{N_{RF} \times N_s}$  and an RF combiner  $\bar{\mathbf{W}}_{RF} \in \mathbb{C}^{N_r \times N_{RF}}$ . Let  $\bar{\mathbf{x}} \in \mathbb{C}^{N_s \times 1}$  denote the transmit signal vector. The output signal vector  $\bar{\mathbf{y}} \in \mathbb{C}^{N_s \times 1}$  after baseband combining is

$$\bar{\mathbf{y}} = \bar{\mathbf{W}}_{BB}^H \bar{\mathbf{W}}_{RF}^H \mathbf{H} \bar{\mathbf{F}}_{RF} \bar{\mathbf{F}}_{BB} \bar{\mathbf{x}} + \bar{\mathbf{v}}, \quad (25)$$

where  $\bar{\mathbf{v}}$  denotes the i.i.d. AWGN vector that is distributed as  $\mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{N_s})$ . Similar to analog beamforming, the elements of the hybrid RF TPC and combiner matrices satisfy  $|\mathbf{F}_{RF}(i, j)| = \frac{1}{\sqrt{N_t}}$  and  $|\mathbf{W}_{RF}(i, j)| = \frac{1}{\sqrt{N_r}}, \forall i, j$ .

#### A. Conventional mmWave-HB-ML Channel Estimation

For the purpose of channel estimation, let us consider the transmission of  $B = \frac{N_t}{N_{RF}}$  pilot blocks, where each block comprises  $N_B$  pilot beams, as shown in the training phase of the frame-structure in Fig. 4. Note that both the training

and semi-blind schemes estimate channel in each frame. Let  $\mathbf{X}_{p,b} \in \mathbb{C}^{N_{RF} \times N_B}, \forall 1 \leq b \leq B$ , denote the baseband transmit pilot matrix for the  $b$ th pilot block so that

$$\mathbf{X}_{p,b} \mathbf{X}_{p,b}^H = P_p N_B \mathbf{I}_{N_{RF}}, \quad (26)$$

where the quantity  $P_p$  denotes the pilot power. Let the matrix  $\mathbf{F}_{RF,b} \in \mathbb{C}^{N_t \times N_{RF}}$  denote the RF TPC in the  $b$ th pilot block. The received pilot block  $\bar{\mathbf{Y}}_{p,b} \in \mathbb{C}^{N_r \times N_B}$  is obtained as

$$\bar{\mathbf{Y}}_{p,b} = \mathbf{H} \mathbf{F}_{RF,b} \mathbf{X}_{p,b} + \bar{\mathbf{V}}_{p,b}, \quad (27)$$

where  $\bar{\mathbf{V}}_{p,b} \in \mathbb{C}^{N_r \times N_B}$  is the AWGN noise matrix. Furthermore, the concatenated output matrix  $\bar{\mathbf{Y}}_p = [\bar{\mathbf{Y}}_{p,1}, \bar{\mathbf{Y}}_{p,2}, \dots, \bar{\mathbf{Y}}_{p,B}] \in \mathbb{C}^{N_r \times N_B B}$  across all the  $B$  pilot blocks is given by

$$\bar{\mathbf{Y}}_p = \mathbf{H} \mathbf{F}_{RF} \mathbf{X}_p + \bar{\mathbf{V}}_p, \quad (28)$$

where  $\mathbf{F}_{RF} = [\mathbf{F}_{RF,1}, \mathbf{F}_{RF,2}, \dots, \mathbf{F}_{RF,B}] \in \mathbb{C}^{N_t \times N_t}$  and  $\bar{\mathbf{V}}_p = [\bar{\mathbf{V}}_{p,1}, \bar{\mathbf{V}}_{p,2}, \dots, \bar{\mathbf{V}}_{p,B}] \in \mathbb{C}^{N_r \times N_B B}$  are the concatenated TPC and noise matrices, respectively. The block-diagonal pilot matrix  $\mathbf{X}_p \in \mathbb{C}^{N_t \times N_B B}$  is obtained as  $\mathbf{X}_p = \text{blkdiag}(\mathbf{X}_{p,1}, \mathbf{X}_{p,2}, \dots, \mathbf{X}_{p,B})$ . Let the matrix  $\mathbf{W}_{RF,q} \in \mathbb{C}^{N_r \times N_{RF}}$  denote the  $q$ th RF combiner, where  $1 \leq q \leq Q_{CB}$ , and  $Q_{CB} = \frac{N_r}{N_{RF}}$  denotes the number of combining matrices. The output pilot matrix  $\tilde{\mathbf{Y}}_{p,q} = \mathbf{W}_{RF,q}^H \bar{\mathbf{Y}}_p \in \mathbb{C}^{N_{RF} \times N_B B}$ , after combining with the  $q$ th RF combiner, is obtained as

$$\tilde{\mathbf{Y}}_{p,q} = \mathbf{W}_{RF,q}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{X}_p + \tilde{\mathbf{V}}_{p,q}, \quad (29)$$

where the noise at the output of the combiner obeys  $\tilde{\mathbf{V}}_{p,q} = \mathbf{W}_{RF,q}^H \bar{\mathbf{V}}_p \in \mathbb{C}^{N_{RF} \times N_B B}$ . Similar to  $\mathbf{F}_{RF}$ , let the RF combiner matrix be  $\mathbf{W}_{RF} = [\mathbf{W}_{RF,1}, \mathbf{W}_{RF,2}, \dots, \mathbf{W}_{RF,Q_{CB}}] \in \mathbb{C}^{N_r \times N_r}$ . The matrix  $\mathbf{Y}_p \in \mathbb{C}^{N_r \times N_B B}$ , obtained by stacking  $\tilde{\mathbf{Y}}_{p,q}$  for all the  $Q_{CB}$  combiners, is obtained as

$$\mathbf{Y}_p = \mathbf{W}_{RF}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{X}_p + \mathbf{V}_p, \quad (30)$$

where  $\mathbf{V}_p \in \mathbb{C}^{N_r \times N_B B}$  is the corresponding stacked noise matrix. From (30) above, the training-based ML estimate of the mmWave-HB MIMO channel matrix  $\mathbf{H}$  is obtained as

$$\hat{\mathbf{H}}_{c,HB} = \mathbf{W}_{RF} \mathbf{Y}_p (\mathbf{F}_{RF} \mathbf{X}_p)^\dagger = \mathbf{H} + \bar{\mathbf{V}}_p (\mathbf{F}_{RF} \mathbf{X}_p)^\dagger, \quad (31)$$

where  $(\cdot)^\dagger$  denotes the pseudo-inverse of the corresponding matrix [25]. Note that since  $\mathbf{W}_{\text{RF}}$  is a DFT matrix, the noise samples at the output of the RF combiner are also i.i.d. complex-Gaussian. Therefore,  $\hat{\mathbf{H}}_{\text{c,HB}}$  also denotes the ML estimate of the channel matrix  $\mathbf{H}$ . Again, the ML estimate described above suffers from the drawback that it does not exploit the unknown data, which in turn leads to higher pilot overheads. Hence, in the next subsection we design a novel semi-blind scheme for mmWave-HB MIMO systems that exploits both the pilots as well as the unknown data symbols for improved estimation accuracy.

### B. Proposed mmWave-HB-SB Channel Estimation Scheme

In contrast to the conventional mmWave-HB-ML estimation technique, the proposed mmWave-HB-SB solution segments data transmission into two phases, as shown in Fig. 4. The first phase processes  $N_{WB}$  spatially-white data blocks, where each data block is further divided into  $B$  sub-blocks with each sub-block containing  $N_B$  data beams. The average power of each element in a data beam is  $P_d$ . This is followed by  $N_{PD}$  precoded data vectors in the second phase. Note that  $BN_B < N_{WB}BN_B \ll N_{PD}$ . Let channel matrix  $\mathbf{H}$ , with  $N_r \geq N_t$ , be decomposed as

$$\mathbf{H} = \mathbf{D}\mathbf{U}^H, \quad (32)$$

where  $\mathbf{D} \in \mathbb{C}^{N_r \times N_t}$  and  $\mathbf{U} \in \mathbb{C}^{N_t \times N_t}$  are termed the decorrelating and unitary matrices, respectively. Note that it is always possible to decompose the mmWave MIMO channel matrix  $\mathbf{H}$  into the decorrelating and unitary matrices. For example, given  $\text{SVD}(\mathbf{H}) = \mathbf{S}\mathbf{T}\mathbf{Q}^H$ , we can set  $\mathbf{D} = \mathbf{S}\mathbf{T}$  and  $\mathbf{U} = \mathbf{Q}$ . The proposed mmWave-HB-SB scheme is based on the principle that the decorrelating matrix  $\mathbf{D}$  is estimated by exploiting the second-order statistical properties of the unknown data symbols, whereas the unitary matrix  $\mathbf{U}$  is estimated exclusively from the training beams. Since the unitary matrix  $\mathbf{U}$  is comprised of significantly fewer parameters than  $\mathbf{H}$ , the estimation accuracy of the proposed semi-blind scheme is significantly better than that of the conventional training-based technique. It is worth noting that since the channel  $\mathbf{H}$  is quasi-static i.e. constant over a frame, it follows that the component matrices  $\mathbf{D}$  and  $\mathbf{U}$  are also quasi-static in nature. To begin with, blind estimation of the decorrelating matrix  $\mathbf{D}$  can be carried out as follows.

Upon exploiting the property  $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}_{N_t}$ , it follows from (32) that

$$\mathbf{H}\mathbf{H}^H = \mathbf{D}\mathbf{D}^H. \quad (33)$$

For the  $i$ th data block of the frame in Fig. 4, the received symbol matrix  $\mathbf{Y}_{d,i} \in \mathbb{C}^{N_r \times N_B B}$  is expressed, similar to (30), as

$$\mathbf{Y}_{d,i} = \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{X}_{d,i} + \mathbf{V}_{d,i}, \quad (34)$$

where  $\mathbf{X}_{d,i} \in \mathbb{C}^{N_t \times N_B B}$  is the block-diagonal matrix comprising spatially-white data symbol matrices in the  $i$ th block, and satisfies  $\mathbb{E}[\mathbf{X}_{d,i} \mathbf{X}_{d,i}^H] = N_B P_d \mathbf{I}_{N_t}$ . The matrix  $\mathbf{V}_{d,i} \in \mathbb{C}^{N_r \times N_B B}$  is the additive noise matrix for the  $i$ th block with its elements distributed as  $\mathcal{CN}(0, \sigma_v^2)$ . Let the

correlation matrix  $\mathbf{R}_Y \in \mathbb{C}^{N_r \times N_r}$  be defined as  $\mathbf{R}_Y \triangleq \mathbb{E}[\mathbf{W}_{\text{RF}} \mathbf{Y}_{d,i} (\mathbf{W}_{\text{RF}} \mathbf{Y}_{d,i})^H]$ . Following (34) and exploiting the above properties, the output covariance matrix  $\mathbf{R}_Y$  can be obtained as

$$\mathbf{R}_Y = N_B P_d \mathbf{H}\mathbf{H}^H + N_B B \sigma_v^2 \mathbf{I}_{N_r}. \quad (35)$$

It therefore follows that,

$$\mathbf{D}\mathbf{D}^H = \mathbf{H}\mathbf{H}^H = \frac{\mathbf{R}_Y - N_B B \sigma_v^2 \mathbf{I}_{N_r}}{N_B P_d}. \quad (36)$$

The estimate  $\hat{\mathbf{R}}_Y$  of  $\mathbf{R}_Y$  can be obtained from the outputs corresponding to the blind data symbols as  $\hat{\mathbf{R}}_Y = \frac{1}{N_{WB}} \sum_{i=1}^{N_{WB}} \mathbf{W}_{\text{RF}} \mathbf{Y}_{d,i} \mathbf{Y}_{d,i}^H \mathbf{W}_{\text{RF}}^H$  [26]. Exploiting the relationship in (33), the blind estimate of the decorrelating matrix  $\mathbf{D}$  can now be determined as

$$\hat{\mathbf{D}} = \hat{\mathbf{\Phi}} \hat{\mathbf{\Sigma}}^{1/2}, \quad (37)$$

where the matrices  $\hat{\mathbf{\Phi}}$  and  $\hat{\mathbf{\Sigma}}^{1/2}$  are obtained from the SVD

$$\hat{\mathbf{\Phi}} \hat{\mathbf{\Sigma}} \hat{\mathbf{\Phi}}^H = \text{SVD} \left( \frac{\hat{\mathbf{R}}_Y - N_B B \sigma_v^2 \mathbf{I}_{N_r}}{N_B P_d} \right). \quad (38)$$

The unitary matrix  $\mathbf{U}$  is then estimated by solving the following constrained optimization problem

$$\begin{aligned} \min_{\mathbf{U}} \quad & \left\| \mathbf{W}_{\text{RF}} \mathbf{Y}_p - \hat{\mathbf{D}} \mathbf{U}^H \mathbf{F}_{\text{RF}} \mathbf{X}_p \right\|_F^2 \\ \text{s.t.} \quad & \mathbf{U}\mathbf{U}^H = \mathbf{I}_{N_t}. \end{aligned} \quad (39)$$

For an orthogonal pilot beam matrix  $\mathbf{F}_{\text{RF}} \mathbf{X}_p$  obeying  $\mathbf{F}_{\text{RF}} \mathbf{X}_p \mathbf{X}_p^H \mathbf{F}_{\text{RF}}^H = N_B P_p \mathbf{I}_{N_t}$ , the estimate  $\hat{\mathbf{U}}$  is given by

$$\hat{\mathbf{U}} = \hat{\mathbf{V}}_Q \hat{\mathbf{U}}_Q^H, \quad (40)$$

where the matrices  $\hat{\mathbf{V}}_Q$  and  $\hat{\mathbf{U}}_Q$  are obtained from the SVD as follows:

$$\hat{\mathbf{U}}_Q \hat{\mathbf{\Sigma}}_Q \hat{\mathbf{V}}_Q^H = \text{SVD} \left( \hat{\mathbf{D}}^H \mathbf{W}_{\text{RF}} \mathbf{Y}_p \mathbf{X}_p^H \mathbf{F}_{\text{RF}}^H \right).$$

Finally, the semi-blind estimate of the channel matrix is obtained as

$$\hat{\mathbf{H}}_{\text{SB}} = \hat{\mathbf{D}} \hat{\mathbf{U}}^H = \hat{\mathbf{\Phi}} \hat{\mathbf{\Sigma}}^{1/2} \hat{\mathbf{U}}_Q \hat{\mathbf{V}}_Q^H. \quad (41)$$

The CRLB of the proposed mmWave-HB-SB estimator, which lower bounds its MSE, is derived next using the constrained CRLB framework.

### C. Constrained CRLB for the mmWave-HB-SB Channel Estimation Scheme

The mmWave-HB-SB channel estimation approach outlined above necessitates the estimation of a unitary matrix  $\mathbf{U}$  that is constrained as  $\mathbf{U}^H \mathbf{U} = \mathbf{I}_{N_t}$ . Hence, the constrained CRLB framework developed in [27] is naturally suited for deriving the pertinent CRLB for such an estimator. Let the decorrelating matrix  $\mathbf{D}$  be assumed to be perfectly known, which is a reasonable assumption since  $\mathbf{D}$  can be estimated with a sufficiently high level of accuracy by transmitting a limited number of spatially-white data symbols, as it will be evidenced by the results in Section-V. Furthermore, the estimation of  $\mathbf{D}$



does not add to the pilot overhead of the system. Let us define the desired parameter vector  $\boldsymbol{\rho} \in \mathbb{C}^{2N_t^2 \times 1}$  to be estimated as

$$\boldsymbol{\rho} = \left[ \text{vec}(\mathbf{U})^T \quad \text{vec}(\mathbf{U}^*)^T \right]^T. \quad (42)$$

The construction of the parameter  $\boldsymbol{\rho}$  above follows from [28], which demonstrates that for each of the complex parameters, we also have to include its complex conjugate. Upon exploiting that  $\text{SVD}(\mathbf{H}) = \mathbf{S}\boldsymbol{\Gamma}\mathbf{U}^H$ , (30) can be recast as

$$\underbrace{\mathbf{S}^H \mathbf{W}_{\text{RF}} \mathbf{Y}_p}_{\tilde{\mathbf{Y}}_p} = \boldsymbol{\Gamma} \mathbf{U}^H \mathbf{F}_{\text{RF}} \mathbf{X}_p + \underbrace{\mathbf{S}^H \mathbf{W}_{\text{RF}} \mathbf{V}_p}_{\tilde{\mathbf{V}}_p}. \quad (43)$$

Vectorizing the model in (43), we obtain

$$\text{vec}(\tilde{\mathbf{Y}}_p^H) = \left[ \boldsymbol{\Gamma} \otimes \mathbf{X}_p^H \mathbf{F}_{\text{RF}}^H \right] \text{vec}(\mathbf{U}) + \text{vec}(\tilde{\mathbf{V}}_p^H). \quad (44)$$

The unconstrained Fisher information matrix (FIM) for the estimation of the parameter vector  $\text{vec}(\mathbf{U})$  in (42) is now obtained as [22]

$$\mathbf{C} = \frac{1}{\sigma_v^2} \left[ \boldsymbol{\Gamma}^H \boldsymbol{\Gamma} \otimes \mathbf{F}_{\text{RF}} \mathbf{X}_p \mathbf{X}_p^H \mathbf{F}_{\text{RF}}^H \right] = \frac{N_B P_p}{\sigma_v^2} \left[ \boldsymbol{\Gamma}^2 \otimes \mathbf{I}_{N_t} \right], \quad (45)$$

where the properties of  $\mathbf{X}_p \mathbf{X}_p^H = N_B P_p \mathbf{I}_{N_t}$ ,  $\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{RF}}^H = \mathbf{I}_{N_t}$  and  $\boldsymbol{\Gamma}^H \boldsymbol{\Gamma} = \boldsymbol{\Gamma}^2$  have been exploited in the above simplification. For an orthogonal pilot matrix  $\mathbf{X}_p$ , we can verify that the vector  $\text{vec}(\mathbf{U}^*)$  in (42) has an identical FIM. Thus, it follows that for the unconstrained estimation of  $\mathbf{U}$ , the FIM  $\mathbf{C}_\rho \in \mathbb{C}^{2N_t^2 \times 2N_t^2}$  for the parameter vector  $\boldsymbol{\rho}$  in (42) is given by

$$\mathbf{C}_\rho = \mathbf{I}_2 \otimes \mathbf{C} = \frac{N_B P_p}{\sigma_v^2} \left[ \mathbf{I}_2 \otimes \boldsymbol{\Gamma}^2 \otimes \mathbf{I}_{N_t} \right]. \quad (46)$$

Let  $\mathbf{u}_i$  denote the  $i$ th column of the unitary matrix  $\mathbf{U}$ . Employing the constraints  $\mathbf{u}_i^H \mathbf{u}_j = 1$  if  $i = j$  and 0 otherwise, the set of non-redundant constraints  $\mathbf{f}(\boldsymbol{\rho})$  for the constrained estimation problem can be constructed as

$$\mathbf{f}(\boldsymbol{\rho}) = \left[ \mathbf{u}_1^H \mathbf{u}_1 - 1, \mathbf{u}_1^H \mathbf{u}_2, \mathbf{u}_2^H \mathbf{u}_1, \mathbf{u}_1^H \mathbf{u}_3, \dots, \right. \\ \left. \mathbf{u}_{N_t}^H \mathbf{u}_{N_t} - 1 \right]^T = \mathbf{0}. \quad (47)$$

The constrained CRLB for the parameter vector  $\boldsymbol{\rho}$  can now be derived as [29]

$$\text{CRLB}(\boldsymbol{\rho}) = \mathbf{B}(\mathbf{B}^H \mathbf{C}_\rho \mathbf{B})^{-1} \mathbf{B}^H, \quad (48)$$

where the matrix  $\mathbf{B} \in \mathbb{C}^{2N_t^2 \times N_t^2}$  is an orthonormal basis for the null-space of the matrix

$$\mathbf{J}(\boldsymbol{\rho}) \triangleq \left[ \frac{\partial \mathbf{f}(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}, \frac{\partial \mathbf{f}(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}^*} \right]. \quad (49)$$

The procedure of deriving the matrix  $\mathbf{B}$  from  $\mathbf{J}(\boldsymbol{\rho})$  is given in Appendix-E. Since the size of  $\boldsymbol{\rho}$  is  $2N_t^2 \times 1$  and  $\mathbf{f}(\boldsymbol{\rho})$  comprises  $N_t^2$  constraints, the matrix  $\mathbf{J}(\boldsymbol{\rho}) \in \mathbb{C}^{N_t^2 \times 2N_t^2}$  has a rank  $2N_t^2 - N_t^2$  [27]. Let us now define the parameter vector  $\mathbf{h} = \text{vec}(\mathbf{H}^T) = [\mathbf{D} \otimes \mathbf{I}_{N_t}] \text{vec}(\mathbf{U}^*)$  for the mmWave MIMO channel matrix  $\mathbf{H}$ . Let the orthonormal basis matrix  $\mathbf{B}$  be

partitioned as  $[\mathbf{B}_1^T, \mathbf{B}_2^T]^T$  with  $\mathbf{B}_1, \mathbf{B}_2 \in \mathbb{C}^{N_t^2 \times N_t^2}$ . The CRLB for the covariance of the estimation of  $\mathbf{h}$  is now obtained as

$$\mathbb{E} \left[ (\hat{\mathbf{h}} - \mathbf{h})(\hat{\mathbf{h}} - \mathbf{h})^H \right] \geq \frac{\sigma_v^2}{P_p N_B} [\mathbf{S}\boldsymbol{\Gamma} \otimes \mathbf{I}_{N_t}] \mathbf{B}_1^* \tilde{\boldsymbol{\Gamma}}^{-1} \mathbf{B}_1^T \\ \times \mathbf{S}\boldsymbol{\Gamma} \otimes \mathbf{I}_{N_t}]^H = \mathbf{C}_H,$$

where the matrix  $\tilde{\boldsymbol{\Gamma}} = \text{diag}[2\sigma_1^2, \sigma_1^2 + \sigma_2^2, \sigma_2^2 + \sigma_1^2, 2\sigma_2^2, \dots] \in \mathbb{R}^{r^2 \times r^2}$ . The proof of the above result is given in Appendix-F. Finally, the MSE arising from the semi-blind estimation of each element  $\mathbf{H}(k, l)$  of the mmWave MIMO channel matrix  $\mathbf{H}$  is given by the element  $\mathbf{C}_H[(k-1)N_t + l, (k-1)N_t + l]$  of the matrix  $\mathbf{C}_H$ , which equals

$$E[|\hat{\mathbf{H}}_{s,\text{HB}}(k, l) - \mathbf{H}(k, l)|^2] \geq \frac{\sigma_v^2}{N_B P_p} \sum_{i=1}^r \sum_{j=1}^r \frac{\sigma_i^2}{\sigma_j^2 + \sigma_i^2} \\ \times |\mathbf{S}(k, i)|^2 |\mathbf{U}(l, j)|^2, \quad (50)$$

where the quantities  $\mathbf{S}(k, i)$  and  $\mathbf{U}(l, j)$  denote the  $(k, i)$ th and  $(l, j)$ th elements of the matrices  $\mathbf{S}$  and  $\mathbf{U}$ , respectively. It is worth noting that the weighting factor  $\sigma_i^2/(\sigma_j^2 + \sigma_i^2)$  in (50) yields the net reduction in the estimation error in each term in comparison to the training-based ML channel estimator. In fact, it is easy to verify that setting this equal to 1 yields the MSE bound for the conventional ML scheme described in (31).

#### D. MSE Performance Improvement of the Proposed mmWave-HB-SB Scheme

Let the parameter vector  $\bar{\mathbf{h}}$  be defined as  $\bar{\mathbf{h}} = [\text{vec}(\mathbf{H})^T, \text{vec}(\mathbf{H}^*)^T]^T \in \mathbb{C}^{2N_r N_t \times 1}$ . Note that the number of unconstrained or free real parameters in the mmWave MIMO channel matrix  $\mathbf{H}$  is  $2N_r N_t$ , which corresponds to the  $N_r N_t$  complex parameters. It can be readily verified that the unconstrained FIM matrix for the parameter vector  $\bar{\mathbf{h}}$  can be evaluated as  $\mathbf{C}_{\bar{\mathbf{h}}} = \frac{\sigma_v^2}{N_B P_p} \mathbf{I}_{2N_r N_t}$ . Thus, the MSE bound for the conventional ML or unconstrained estimation of  $\mathbf{H}$ , as performed by the mmWave-HB-ML estimator, can be evaluated as

$$\mathbb{E} \left[ \left\| \hat{\mathbf{H}}_{c,\text{HB}} - \mathbf{H} \right\|_F^2 \right] = \frac{1}{2} \text{Tr}[\mathbf{C}_{\bar{\mathbf{h}}}] = \frac{\sigma_v^2 N_t N_r}{P_p N_B}. \quad (51)$$

The conventional ML estimator utilizes the training matrix  $\mathbf{X}_p$  for the estimation of the complete matrix  $\mathbf{H}$ , which has  $2N_r N_t$  real parameters. However, the estimation of the decorrelating matrix  $\mathbf{D}$  using the statistical information of data symbols allows the proposed mmWave-HB-SB scheme to utilize the training matrix  $\mathbf{X}_p$  for exclusively estimating the unitary matrix  $\mathbf{U}$ , which has only  $N_t^2$  real parameters [28]. This reduction in the number of free parameters significantly enhances the channel estimation accuracy of the proposed scheme. Let us now consider the constrained estimate obtained via the mmWave-HB-SB technique as  $\hat{\mathbf{H}}_{s,\text{HB}} = \mathbf{D} \hat{\mathbf{U}}^H$ . Following our procedure from the previous section, the estimation error  $\boldsymbol{\epsilon}_{\bar{\mathbf{h}}} = \hat{\bar{\mathbf{h}}} - \bar{\mathbf{h}}$  of the parameter  $\bar{\mathbf{h}}$  for this case is bounded as

$$\mathbb{E}[\boldsymbol{\epsilon}_{\bar{\mathbf{h}}} \boldsymbol{\epsilon}_{\bar{\mathbf{h}}}^H] \geq \bar{\mathbf{B}}(\bar{\mathbf{B}}^H \mathbf{C}_{\bar{\mathbf{h}}} \bar{\mathbf{B}})^{-1} \bar{\mathbf{B}}^H, \quad (52)$$

where the matrix  $\bar{\mathbf{B}}$  denotes an orthonormal basis for the matrix  $\mathbf{J}(\bar{\mathbf{h}}) \in \mathbb{C}^{N_t^2 \times 2N_r N_t}$ . Note that the rank of  $\mathbf{J}(\bar{\mathbf{h}})$  is  $2N_r N_t - N_t^2$  [27]. Hence, it follows that the matrix  $\bar{\mathbf{B}}$  has a size of  $2N_r N_t \times N_t^2$ . Substituting  $\mathbf{C}_{\bar{h}}$  in (52) and then applying the trace operator to both sides, we obtain

$$\mathbb{E} \left[ \left\| \hat{\mathbf{H}}_{s,\text{HB}} - \mathbf{H} \right\|_F^2 \right] \geq \frac{1}{2} \text{Tr} \left[ (\bar{\mathbf{B}}^H \mathbf{C}_{\bar{h}} \bar{\mathbf{B}})^{-1} \right] = \frac{\sigma_v^2}{2N_B P_p} N_t^2, \quad (53)$$

which can be seen to be proportional to the number of free parameters of the unitary matrix  $\mathbf{U}$ , namely to  $N_t^2$ . It follows from (51) and (53) that the MSE improvement of the proposed mmWave-HB-SB scheme over the training-based mmWave-HB-ML scheme is

$$\mathcal{G} \leq \mathbb{E} \left[ \left\| \hat{\mathbf{H}}_{c,\text{HB}} - \mathbf{H} \right\|_F^2 \right] / \mathbb{E} \left[ \left\| \hat{\mathbf{H}}_{s,\text{HB}} - \mathbf{H} \right\|_F^2 \right] = \frac{2N_r}{N_t}. \quad (54)$$

It becomes immediately clear from the above expression that when  $N_r = N_t$ , the proposed mmWave-HB-SB scheme outperforms the training-based mmWave-HB-ML scheme by 3 dB. Furthermore, the gain  $\mathcal{G}$  of mmWave-HB-SB is seen to increase with the number of RAs  $N_r$ , which leads to its significantly improved performance in comparison to conventional estimation. Finally, it is also worth noting that the per parameter MSE bound for the mmWave-HB-SB estimator is  $\geq \frac{\sigma_v^2 N N_r^2}{2P_p N_B} \left( \frac{1}{N_t N_r} \right)$ , which is seen to decrease with the number of RAs  $N_r$ , while that of the competing mmWave-HB-ML is  $\frac{\sigma_v^2 N_t N_r}{P_p N_B} \left( \frac{1}{N_t N_r} \right)$ , which can be seen to remain constant with  $N_t$  and  $N_r$ .

#### IV. COMPLEXITY OF THE PROPOSED AND THE EXISTING SCHEMES

This section presents a brief analysis of the computational complexities of the proposed mmWave-AB-SB and mmWave-HB-SB schemes, and the existing mmWave-AB-ML, mmWave-HB-ML, SBL [8], OMP [11] and ANM-SB [17] schemes. The complexity of each scheme is quantified in terms of the number of real floating point operations (flops), where a real flop denotes either a real addition, multiplication or division [30], [31]. Due to lack of space, the detailed derivations of the computational complexity have been relegated to our technical report in [32]. As described in [32], both the conventional mmWave-AB-ML and mmWave-HB-ML schemes cost  $\mathcal{O}(N_t^2 N_r + N_r^2 N_t + N_t^3)$  flops. Next, the state-of-the-art SBL [8] and OMP [11] schemes require  $\mathcal{O}(G^2 N_r N_t + N_r^3 N_t^3)$  and  $\mathcal{O}(G^6 + G^4 N_r N_t)$  flops, respectively, where, typically, the grid-size  $G \geq \min(N_t, N_r)$  [31]. The proposed mmWave-AB-SB and mmWave-HB-SB schemes fare  $\mathcal{O}(N_r^2 N_{D1} + N_r^3 + N_t^2 N_r + N_r^2 N_t)$  and  $\mathcal{O}(N_r^2 N_t N_{WB} + N_t^3 + N_t N_r^2 + N_r N_t^2 + N_r^3)$  flops, respectively. The terms  $N_r^2 N_{D1}$  and  $N_r^2 N_t N_{WB}$  above arise due to the estimation of the covariance matrices  $\mathbf{R}_y$  and  $\mathbf{R}_Y$  using  $N_{D1}$  and  $N_{WB}$  spatially-white data symbols in the proposed mmWave-AB-SB and mmWave-HB-SB schemes, respectively. The cubic order terms in the complexity of the proposed schemes appear due to the associated SVD operations to exploit the statistical characteristics in data symbols. The existing ANM-SB scheme [17] has a complexity

$\mathcal{O}(N_r^2 N_t^2 \bar{L}^2)$ , where the quantity  $\bar{L}$  denotes a bound on the number of spatial paths. With  $G = N_r = N_t$ , it can be seen that the complexity of the above mentioned schemes follow the order  $\mathcal{O}(\text{SBL [8]}) > \mathcal{O}(\text{OMP [11]}) > \mathcal{O}(\text{ANM-SB [17]}) > \mathcal{O}(\text{mmWave-HB-SB}) > \mathcal{O}(\text{mmWave-AB-SB}) > \mathcal{O}(\text{mmWave-AB-ML}) = \mathcal{O}(\text{mmWave-HB-ML})$ . It can be observed that the proposed schemes have significantly lower complexity than the existing OMP, SBL and ANM-SB schemes. A visual comparison of the complexities is given in Section-V.

#### V. SIMULATION RESULTS

As described in several related contributions [7], [8], [33], the narrowband clustered channel model based mmWave MIMO channel matrix  $\mathbf{H}$  is expressed as

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{N_{ray}}} \sum_{i=1}^{N_{cl}} \sum_{j=1}^{N_{ray,i}} \alpha_{ij} \mathbf{a}_r(\theta_{ij}^r) \mathbf{a}_t^H(\theta_{ij}^t), \quad (55)$$

where the 3-tuple  $(\alpha_{ij}, \theta_{ij}^r, \theta_{ij}^t)$  denotes the complex path gain  $\alpha_{ij}$ , angle of arrival (AoA)  $\theta_{ij}^r$ , and angle of departure (AoD)  $\theta_{ij}^t$  associated with the  $j$ th ray in the  $i$ th cluster. The quantity  $N_{cl}$  denotes the total number of clusters, each of which contributes  $N_{ray,i}$  spatial multipath components, and  $N_{ray} = \sum_{i=1}^{N_{cl}} N_{ray,i}$  denotes the total number of rays. The vector  $\mathbf{a}_t(\theta_{ij}^t) = \frac{1}{\sqrt{N_t}} [1, e^{-j\frac{2\pi}{\lambda} d_T \cos \theta_{ij}^t}, \dots, e^{-j\frac{2\pi}{\lambda} (N_t-1) d_T \cos \theta_{ij}^t}]^T \in \mathbb{C}^{N_t \times 1}$  denotes the transmit array response and the vector  $\mathbf{a}_r(\theta_{ij}^r) = \frac{1}{\sqrt{N_r}} [1, e^{-j\frac{2\pi}{\lambda} d_R \cos \theta_{ij}^r}, \dots, e^{-j\frac{2\pi}{\lambda} (N_r-1) d_R \cos \theta_{ij}^r}]^T \in \mathbb{C}^{N_r \times 1}$  is the receive array response, corresponding to the  $j$ th ray in the  $i$ th cluster. Here the quantity  $\lambda$  denotes the carrier's wavelength and  $d_R$  and  $d_T$  are the RA and TA spacings, respectively. The sparse nature of the mmWave MIMO channel above renders it highly correlated, as can also be seen from the beamspace model described in [8], [11]. The various simulation parameters have been chosen as follows, unless stated otherwise. The mmWave MIMO channel comprises of  $N_{cl} \in \{4, 6\}$  clusters and  $N_{ray,i} \in \{4, 6\}$  rays per cluster. The mean-angles of the clusters are assumed to be uniformly distributed between  $(0, \pi)$ , and the  $N_{ray,i}$  rays associated with the  $i$ th cluster are assumed to have a Laplacian distribution with an angular spread of 0.1 radian around the mean-angle of the cluster. The corresponding path gains are generated as i.i.d.  $\mathcal{CN}(0, 1)$ . The inter-antenna spacings of the TA and RA arrays are fixed as  $d_T = d_R = \frac{\lambda}{2}$ . The SNR in decibels (dB) is defined as  $\text{SNR (dB)} = 10 \log_{10} \frac{P_p}{\sigma_v^2}$ . The normalized mean square error (NMSE) is defined as  $\left\| \hat{\mathbf{H}} - \mathbf{H} \right\|_F^2 / \left\| \mathbf{H} \right\|_F^2$ . The corresponding theoretical NMSEs have also been similarly normalized.

##### A. mmWave-AB Beamformer Estimation

Fig. 5(a) shows the NMSE versus SNR performance of both the conventional mmWave-AB-ML and of the proposed mmWave-AB-SB schemes, with perfect  $\mathbf{s}_1$ , for channel estimation in mmWave-AB MIMO systems. Since our solution successfully leverages both the training as well as the

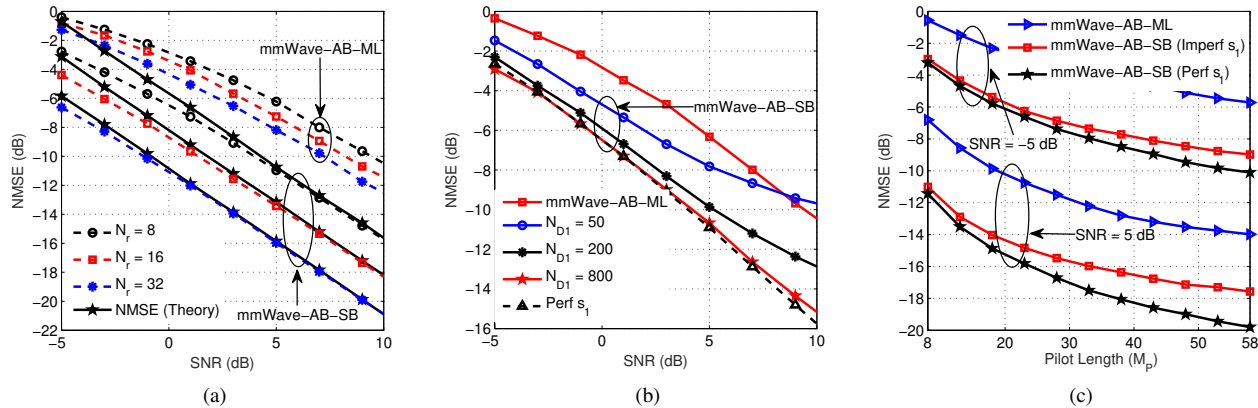


Fig. 5: Performance comparison of the mmWave-AB-ML and mmWave-AB-SB schemes: (a) NMSE versus SNR with perfect  $s_1$  for  $N_r \in \{8, 16, 32\}$ ,  $N_t = 8$  and  $M_P = N_t$ , along with theoretical NMSE from Eq. (21); (b) NMSE versus SNR with perfect  $s_1$ , imperfect  $s_1$  with  $N_{D1} \in \{50, 200, 800\}$  and  $N_t = N_r = M_P = 8$ ; and (c) NMSE versus pilot length  $M_P$  for  $\text{SNR} \in \{-5, 5\}$  dB,  $N_t = N_r = 8$  and  $N_{D1} = 800$ .

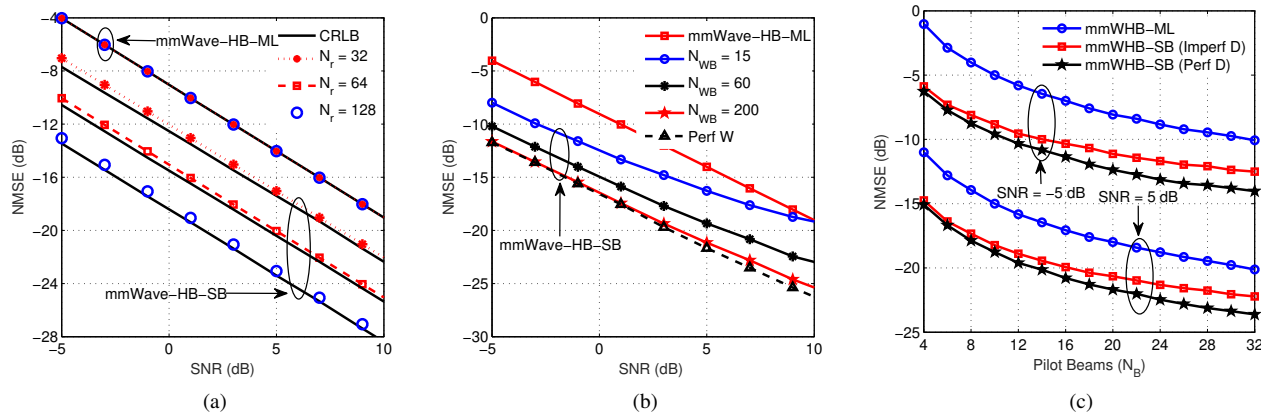


Fig. 6: Performance comparison of the mmWave-HB-SB and mmWave-HB-ML schemes: (a) NMSE versus SNR for  $N_r \in \{32, 64, 128\}$  with perfect  $\mathbf{D}$ ,  $N_t = 32$ ,  $N_{RF} = N_B = 8$  and  $N_{ray,i} = N_{cl} = 4$ , along with CRLB from (53); (b) NMSE versus SNR for  $N_{WB} \in \{15, 60, 200\}$  with imperfect and perfect  $\mathbf{D}$ ,  $N_t = 16$ ,  $N_r = 32$ ,  $N_{RF} = 8$  and  $N_{ray,i} = N_{cl} = 4$ ; and (c) NMSE versus number of pilot beams  $N_B$  for  $\text{SNR} \in \{-5, 5\}$  dB with imperfect and perfect  $\mathbf{D}$ ,  $N_t = N_r = 8$ ,  $N_{RF} = 4$ ,  $N_{WB} = 200$  and  $N_{ray,i} = N_{cl} = 4$ .

second-order statistics of the data symbols, it significantly outperforms the former technique for a fixed number of pilot beams  $M_P = N_t$ . Since the ML technique estimates the complete mmWave MIMO channel matrix  $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ ,  $2N_r N_t$  number of parameters has to be estimated, which increases both with  $N_t$  and  $N_r$ . By contrast, the proposed mmWave-AB-SB scheme directly estimates the dominant right singular vector  $\mathbf{q}_1 \in \mathbb{C}^{N_t \times 1}$ , which only has  $2N_t$  parameters that does not increase with  $N_r$ , thus rendering it efficient. This explains the progressively improving performance of the mmWave-AB-SB technique in comparison to the mmWave-AB-ML scheme upon increasing  $N_r$ . It can also be observed that the experimental MSE values obtained for the proposed mmWave-AB-SB scheme closely agree with the corresponding values obtained using the analytical MSE expression derived in (21).

Fig. 5(b) shows the effect of increasing the number of data

symbols  $N_{D1}$  on the semi-blind estimate. It can be observed that when  $N_{D1}$  is on the order of a few tens of data symbols, the proposed scheme fails to approach the conventional ML scheme at high SNR values, since the error arising in the estimate of  $\mathbf{R}_y$  results in an increased error in the estimation of the left singular vector  $\mathbf{s}_1$ , which results in an NMSE floor for the proposed schemes. However, this NMSE floor disappears as  $N_{D1}$  is increased to a few hundred symbols and the proposed mmWave-AB-SB scheme significantly outperforms the mmWave-AB-ML scheme for the entire SNR range. Furthermore, the NMSE approaches that of the scenario having perfect  $s_1$  as  $N_{D1}$  increases. Fig. 5(c) demonstrates the NMSE performance of both the mmWave-AB-ML and the proposed mmWave-AB-SB schemes versus the pilot length  $M_P$  at different SNR values for both  $N_{D1} = 800$  and for a large number of data symbols having perfect  $s_1$  estimates. As expected, the performance of both the ML as well as of the

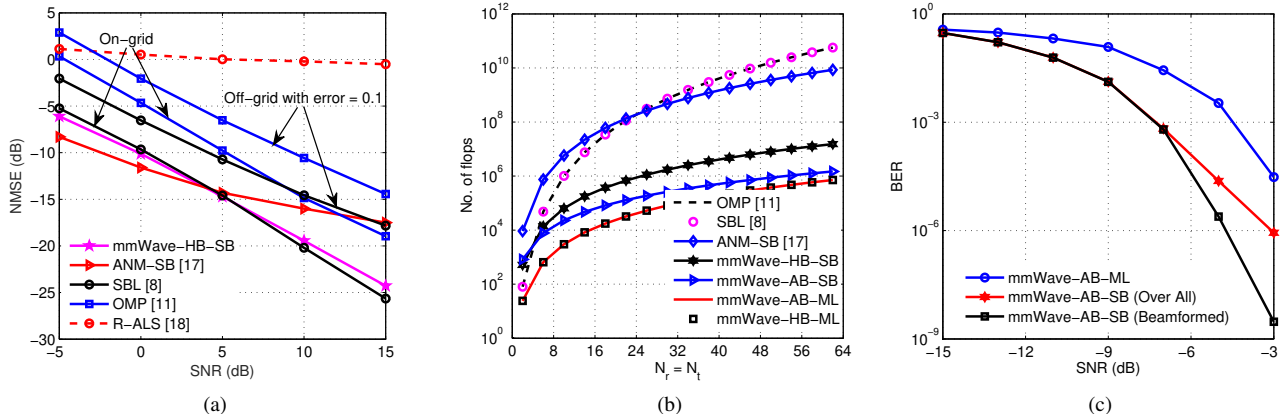


Fig. 7: Comparison between mmWave-HB-SB and the existing schemes: (a) NMSE versus SNR with  $N_t = N_r = 32$ ,  $N_B = N_{RF} = 2$ ,  $N_{WB} = 60$ ,  $N_{cl} = 8$ ,  $N_{ray,i} = 1$  and grid-size  $G = 32$ ; and (b) Complexity comparison with  $N_{D1} = 200$ ,  $N_{WB} = 60$ ,  $G = N_r = N_t$ ,  $L = N_{cl}N_{ray,i} = 4 \times 6 = 24$ . (c) Overall BER of the mmWave-AB-ML and mmWave-AB-SB schemes with  $N_t = M_P = 8$ ,  $N_r = 32$ ,  $N_{D1} = 200$ ,  $N_{D2} = 10^6$ , and  $N_{ray,i} = N_{cl} = 4$ .

proposed mmWave-AB-SB techniques can be seen to improve upon increasing  $M_P$ .

### B. mmWave-HB MIMO Channel Estimation

Fig. 6(a) demonstrates the performance gain of the proposed mmWave-HB-SB scheme over the training-based mmWave-HB-ML scheme as well as the analytical results derived in Section-III-D for both the systems. The MSE of the mmWave-HB-SB scheme exhibits 3 dB, 6 dB and 9 dB performance gain over its ML counterparts, for  $32 \times 32$ ,  $64 \times 32$  and  $128 \times 32$  mmWave-HB MIMO systems, respectively, in accordance with the formula  $10 \log_{10} \frac{2N_r}{N_t}$  derived in (54). It can also be observed that the proposed mmWave-HB-SB scheme approaches the MSE given by the complex-constrained CRLB derived in Section-III-C.

Fig. 6(b) compares the NMSE versus SNR performance of the proposed mmWave-HB-SB scheme both with perfect and imperfect knowledge of the decorrelating matrix  $\mathbf{D}$ . It can be seen from Fig. 6(b) that the estimation performance of the proposed scheme degrades with a reduction in the number of spatially white symbols. This can be attributed to the increased estimation error arising in the estimation of  $\mathbf{D}$  from equation (37). As shown in (38), the matrix  $\mathbf{D}$  is estimated using the sample correlation matrix  $\hat{\mathbf{R}}_Y$ , which approaches the true correlation matrix  $\mathbf{R}_Y$  as the number of spatially-white data blocks  $N_{WB}$  increases. Consequently, the estimate  $\hat{\mathbf{D}}$  approaches the true decorrelating matrix  $\mathbf{D}$ . The same trend is reflected in Fig. 6(b), wherein the NMSE of the proposed mmWave-HB-SB scheme progressively approaches the scenario of perfect  $\mathbf{D}$ , as the number of spatially-white data block  $N_{WB}$  increases.

Fig. 6(c) shows the NMSE versus number pilot beams  $N_B$  performance of the mmWave-HB-ML and the proposed mmWave-HB-SB schemes. Similar to Fig. 5(c) for the mmWave-AB MIMO systems, it can be seen that the proposed mmWave-HB-SB scheme requires a significantly lower number of pilot beams  $N_B$  than the conventional ML

scheme for an identical NMSE performance. For example, for  $\text{SNR} = 5$  dB, the proposed mmWave-HB-SB scheme requires only  $N_B = 4$  pilot beams to achieve an NMSE of  $-15$  dB, whereas the conventional mmWave-HB-ML scheme requires  $N_B = 10$  pilot beams to achieve an identical NMSE performance.

Fig. 7(a) compares the NMSE performance of the proposed mmWave-HB-SB and the state-of-the-art existing sparse techniques such as SBL [8] and OMP [11], also the atomic-norm-minimization (ANM) [17]- and R-ALS [18]-based semi-blind channel estimation techniques for mmWave-HB MIMO systems. It can be observed that not only does the mmWave-HB-SB significantly outperform the OMP and SBL schemes, but is also robust to practical aberrations such as grid mismatch, where the AoDs/ AoAs do not coincide with the dictionary points. It can also be observed that at low SNR, the NMSE of the proposed scheme is close to that of the ANM-SB scheme, while at high SNR, the former significantly outperform the latter. Furthermore, it can be readily observed that the NMSE of the proposed mmWave-HB-SB scheme is significantly lower than that of the R-ALS [18] scheme. This is due to the fact that i) the problem formulation in [18] based on a low-rank matrix decomposition is non-convex in nature, which is subsequently solved using the R-ALS scheme by decomposing it into two alternating convex sub-problems; ii) the low-rank matrix decomposition therein relies on the assumption  $N_t \ll N_r$ , which need not to satisfy in a typical hybrid MIMO system; and iii) unlike the channel estimation model formulated in Eq. (26)-(30), the scheme in [18] does not excite all the angular modes of the mmWave MIMO channel. Thus, the proposed scheme is ideally suited for mmWave MIMO channel estimation in practical settings.

Fig. 7(b) shows the complexity comparison of the proposed and several existing state-of-the-art schemes. As described in Section-IV, both the mmWave-AB-ML and mmWave-HB-ML schemes can be seen to have an identical computational cost, which is slightly lower than that of the proposed mmWave-

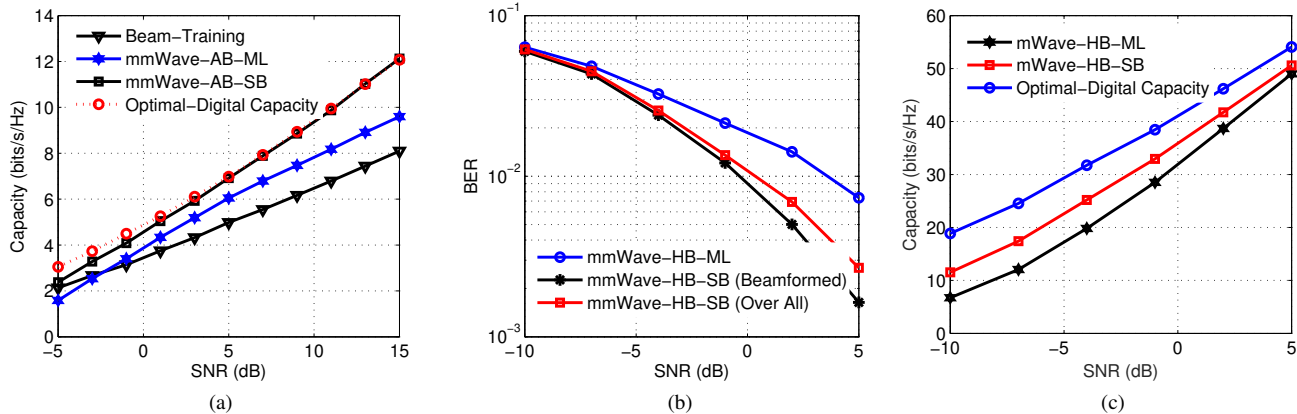


Fig. 8: (a) Overall capacity of the mmWave-AB-ML and mmWave-AB-SB schemes with  $N_t = M_P = N_r = 8$ ,  $N_{D1} = 200$ ,  $N_{D2} = 10^3$ , and  $N_{ray,i} = N_{cl} = 4$ . (b) Overall BER of the mmWave-HB-ML and mmWave-HB-SB schemes with  $N_t = N_r = 32$ ,  $N_B = 8$ ,  $N_{RF} = 8$ ,  $N_{WB} = 100$ ,  $N_{cl} = N_{ray,i} = 4$  and  $N_{PD} = 1000$ . (c) Overall capacity of the mmWave-HB-ML and mmWave-HB-SB schemes with  $N_t = N_r = 32$ ,  $N_B = 8$ ,  $N_{RF} = 8$ ,  $N_{WB} = 200$ ,  $N_{cl} = N_{ray,i} = 4$  and  $N_{PD} = 2000$ .

AB-SB technique. Furthermore, it can also be observed that the complexity of the mmWave-HB-SB scheme is only moderately higher than that of the mmWave-HB-ML scheme, while it is significantly lower in comparison to the existing SBL [8], OMP [11] and ANM-SB [17] schemes. Thus, the proposed semi-blind channel estimation schemes yield the best performance complexity trade-off in practical systems.

### C. Overall Performance Comparison of the ML and the proposed SB Techniques

Fig. 7(c) shows the BER of both the mmWave-AB-ML and of the proposed mmWave-AB-SB techniques. It can be observed that the overall BER of our proposed mmWave-AB-SB technique, which includes the performance of both the spatially-white data as well as the beamformed data symbols, is superior to that of the mmWave-AB-ML technique. Observe that in order to obtain an overall superior BER for the mmWave-AB-SB technique, it is desirable that  $\frac{N_{D2}}{N_{D1}} \gg 1$ , since the BER for beamformed symbols is lower due to the higher diversity gain attained. Fig. 8(a) compares the capacities of the mmWave-AB MIMO system, with channel estimation performed using the mmWave-AB-SB and mmWave-AB-ML techniques. To benchmark the performance, the capacity of a ‘genie’-transceiver, which employs the optimal digital beamformer and combiner vectors obtained using perfect knowledge of the mmWave MIMO channel, has also been plotted. The capacity of the mmWave-AB MIMO technique is also compared to that of a hypothetical beam-training method, which estimates the dominant AoA and AoD pair within an angular uncertainty of 0.1 radian. The improved channel estimation accuracy of the mmWave-AB-SB technique is also reflected in the improved MIMO capacity, which is very close to that of the conventional fully digital MIMO beamformer/combiner with perfect CSI.

Fig. 8(b) and Fig. 8(c) demonstrate the resultant BER versus SNR and the capacity versus SNR performance of mmWave-HB systems with the channel estimation performed using the

proposed SB and the existing ML schemes. It can be seen that the overall performance of the proposed mmWave-HB-SB estimator is significantly better than that of its ML counterpart in terms of both the capacity as well as BER.

## VI. CONCLUSIONS

We have conceived the mmWave-AB-SB and mmWave-HB-SB schemes, for analog- and hybrid-beamforming based mmWave MIMO systems, respectively. A unique aspect of these schemes is that they exploit both the training as well as the statistical information available in the unknown data symbols for enhancing the estimation performance. The mmWave-AB-SB has the advantage that it does not necessitate the estimation of the complete channel matrix  $\mathbf{H}$  and directly estimates the beamforming vectors. A first order perturbation analysis based framework has also been presented for deriving the analytical MSE expression of the proposed mmWave-AB-SB scheme. The mmWave-HB-SB scheme developed subsequently exploits the decorrelating-unitary decomposition of the channel matrix  $\mathbf{H}$ . It is shown that while the former can be estimated exclusively using the unknown data symbols, the known training symbols are only used for the estimation of the latter. The lower number of free parameters in the unitary matrix to be estimated by the proposed technique leads to its significantly improved estimation performance, as shown by our explicit constrained CRLB analysis. Furthermore, its performance gain over the conventional technique shows a progressively improving trend with the number of RAs  $N_r$ . Our simulation results have demonstrated the improved NMSE, BER and capacity of the proposed schemes in comparison to the existing ones.

## APPENDIX A

Taking  $(\cdot)^H$  on both the sides of (11), we arrive at:

$$\check{\mathbf{y}}^H = \mathbf{X}^H \mathbf{F}_{RF}^H (\sigma_1 \mathbf{q}_1) + \mathbf{V}^H \mathbf{W}_{RF}^H \hat{\mathbf{s}}_{1,s}^H. \quad (56)$$

Furthermore, the LS problem in (12) can be recast as

$$\hat{\mathbf{q}}_{1,s} = \arg \min_{\|\mathbf{q}_1\|=1} \|\hat{\mathbf{y}}^H - \mathbf{X}^H \mathbf{F}_{\text{RF}}^H (\sigma_1 \mathbf{q}_1)\|^2. \quad (57)$$

The solution to (57) can be readily obtained as

$$\hat{\sigma}_{1,s} \hat{\mathbf{q}}_{1,s} = (\mathbf{F}_{\text{RF}} \mathbf{X} \mathbf{X}^H \mathbf{F}_{\text{RF}}^H)^{-1} \mathbf{F}_{\text{RF}} \mathbf{X} \hat{\mathbf{y}}^H = \frac{1}{P_p} \mathbf{F}_{\text{RF}} \mathbf{X} \hat{\mathbf{y}}^H. \quad (58)$$

Thus,  $\hat{\sigma}_{1,s} = \frac{1}{P_p} \|\mathbf{F}_{\text{RF}} \mathbf{X} \hat{\mathbf{y}}^H\|$ , since  $\|\hat{\mathbf{q}}_{1,s}\| = 1$ . Substituting  $\hat{\sigma}_{1,s}$  into (58), completes the proof.

## APPENDIX B

This Appendix presents some relevant results from matrix perturbation theory [24]. Let  $\mathbf{T} \in \mathbb{C}^{N \times N}$  be a Hermitian symmetric matrix, whose eigenvalues and the corresponding eigenvectors are denoted by  $\{\lambda_i\}_{i=1}^N$  and  $\{\mathbf{e}_i\}_{i=1}^N$ , respectively. Let the first-order perturbation of the matrix  $\mathbf{T}$ , denoted by  $\hat{\mathbf{T}}$ , be given as  $\hat{\mathbf{T}} = \mathbf{T} + \Delta \mathbf{T}$ , where  $\Delta \mathbf{T}$  is the error matrix. Let  $\{\hat{\mathbf{e}}_i\}_{i=1}^N$  denote the eigenvectors of the perturbed matrix  $\hat{\mathbf{T}}$ . Then, for small perturbations, the eigenvector  $\hat{\mathbf{e}}_i$  can be approximated in terms of the true eigenvectors  $\mathbf{e}_i$  as

$$\hat{\mathbf{e}}_i \approx \mathbf{e}_i + \sum_{j=1, j \neq i}^{R_T} \frac{\mathbf{e}_j^H \Delta \mathbf{T} \mathbf{e}_i}{\lambda_i - \lambda_j} \mathbf{e}_j, \quad (59)$$

where  $R_T = \text{rank}(\mathbf{T})$ . An important assumption in (59) is that the non-zero eigenvalues of the matrix  $\mathbf{T}$  are distinct, i.e.,  $\lambda_j \neq \lambda_k$  for  $j \neq k$ . For  $i = 1$ , the expression above can be written as

$$\hat{\mathbf{e}}_1 \approx \mathbf{e}_1 + \sum_{j=2}^{R_T} \frac{\mathbf{e}_j^H \Delta \mathbf{T} \mathbf{e}_1}{\lambda_1 - \lambda_j} \mathbf{e}_j. \quad (60)$$

The above expression can be compactly rewritten as  $\hat{\mathbf{e}}_1 \approx \mathbf{E} \tilde{\mathbf{e}}$ , where the matrix  $\mathbf{E} \in \mathbb{C}^{N \times R_T}$  and the vector  $\tilde{\mathbf{e}} \in \mathbb{C}^{R_T \times 1}$  are defined as

$$\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{R_T}] \quad \text{and} \quad \tilde{\mathbf{e}} = \left[ 1, \frac{\mathbf{e}_2^H \Delta \mathbf{T} \mathbf{e}_1}{\lambda_1 - \lambda_2}, \dots, \frac{\mathbf{e}_{R_T}^H \Delta \mathbf{T} \mathbf{e}_1}{\lambda_1 - \lambda_{R_T}} \right]^T. \quad (61)$$

Furthermore, for a unit-norm vector  $\tilde{\mathbf{e}}_1 = \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|}$  and  $\tilde{\mathbf{e}} = \frac{\tilde{\mathbf{e}}}{\|\tilde{\mathbf{e}}\|} = [1 + \Delta e_1, \Delta e_2, \dots, \Delta e_{R_T}]^T$ , an equivalent relationship is obtained as  $\tilde{\mathbf{e}}_1 = \mathbf{E} \tilde{\mathbf{e}}$ . It follows from [34] that the quantities  $\{\Delta e_i\}_{i=1}^{R_T}$  can be approximated as

$$\Delta e_1 \approx -\frac{1}{2} \sum_{j=2}^{R_T} |\Delta e_j|^2 \quad \text{and} \quad \Delta e_i \approx \frac{\mathbf{e}_i^H \Delta \mathbf{T} \mathbf{e}_1}{\lambda_1 - \lambda_i}, \quad i = 2, \dots, R_T. \quad (62)$$

## APPENDIX C

Ignoring the terms of the order  $\mathcal{E}^H \mathcal{E}$ , the MSE contaminating the estimate  $\hat{\mathbf{q}}_{1,s}$  can be evaluated using the expression in (20) as

$$\|\hat{\mathbf{q}}_{1,s} - \mathbf{q}_1\|^2 = \frac{\|\mathcal{E}^H \mathbf{s}_1\|^2}{\sigma_1^2} + \frac{|\mathbf{s}_1^H \mathcal{E} \mathbf{q}_1 + \mathbf{q}_1^H \mathcal{E}^H \mathbf{s}_1|^2}{4\sigma_1^2}. \quad (63)$$

Since we have  $\mathbf{V} = \mathbf{W}_{\text{RF}}^H \bar{\mathbf{V}}$ , the matrix  $\mathcal{E}$  in (18) can be recast as  $\mathcal{E} = \bar{\mathbf{V}} \mathbf{X}^H \mathbf{F}_{\text{RF}}^H / P_p$ . Furthermore, since each element of the noise  $\bar{\mathbf{V}}$  is i.i.d.  $\mathcal{CN}(0, \sigma_v^2)$ , it follows from the properties,  $\mathbf{X} \mathbf{X}^H = P_p \mathbf{I}_{M_P}$  and  $\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{RF}}^H = \mathbf{I}_{N_t}$ , that each element of the matrix  $\mathcal{E}$  is i.i.d. as  $\mathcal{CN}(0, \sigma_v^2 / P_p)$ . Let  $x = \mathbf{s}_1^H \mathcal{E} \mathbf{q}_1$  and  $y = \mathbf{q}_1^H \mathcal{E}^H \mathbf{s}_1$ . In view of the fact that both  $x$  and  $y$  represent linear transformations of the matrix  $\mathcal{E}$ , we obtain  $\mathbb{E}[xy^*] = 0$ , and  $\mathbb{E}[|x|^2] = \mathbb{E}[|y|^2] = \sigma_v^2 / P_p$  since  $\|\mathbf{q}_1\| = \|\mathbf{s}_1\| = 1$ . Exploiting the above properties, the MSE in  $\hat{\mathbf{q}}_{1,s}$  can be evaluated as

$$\mathbb{E}[\|\hat{\mathbf{q}}_{1,s} - \mathbf{q}_1\|^2] = \frac{\sigma_v^2}{2P_p \sigma_1^2} (2N_t - 1). \quad (64)$$

## APPENDIX D

Exploiting the relationship between  $\tilde{\mathbf{c}}$  and  $\mathbf{c}$ , it can be readily seen that

$$\begin{aligned} \tilde{c}_1 &= \frac{\sigma_1 (1 + \Delta c_1)}{\sqrt{\sigma_1^2 |1 + \Delta c_1|^2 + \sum_{i=2}^{N_r} \sigma_i^2 |\Delta c_i|^2}} \\ &\approx (1 + \Delta c_1) \left[ 1 - \frac{1}{2} \left( 2\Delta c_1 + \sum_{i=2}^{N_r} \frac{\sigma_i^2}{\sigma_1^2} |\Delta c_i|^2 \right) \right] \\ &\approx 1 - \frac{1}{2} \sum_{i=2}^{N_r} \frac{\sigma_i^2}{\sigma_1^2} |\Delta c_i|^2. \end{aligned} \quad (65)$$

Substituting the above expression into (22), yields the desired result in (23). The average value of the quantity  $|\Delta c_i|^2$  is derived next. Stacking the received vectors  $\mathbf{y}_i$  in (9), for  $1 \leq i \leq N_{D1}$ , as  $\mathbf{Y}_D = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{N_{D1}}] \in \mathbb{C}^{N_r \times N_{D1}}$ , one obtains  $\mathbf{Y}_D = \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_D + \mathbf{W}_{\text{RF}}^H \mathbf{V}_D$ , where  $\mathbf{F}_D = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_{N_{D1}}] \in \mathbb{C}^{N_t \times N_{D1}}$  is the matrix of spatially-white data vectors and  $\mathbf{V}_D = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N_{D1}}] \in \mathbb{C}^{N_r \times N_{D1}}$  is the corresponding noise matrix. Recall that the vector  $\mathbf{s}_1$  is estimated using the SVD of the covariance matrix  $\mathbf{R}_y$  of the vector  $\mathbf{W}_{\text{RF}} \mathbf{y}_i$ . Let the quantity  $\alpha_D$  be defined as  $\alpha_D = N_{D1} P_d / N_t$ . Scaling with the factor  $\alpha_D$ , the SVD ( $\hat{\mathbf{R}}_y$ ) can be rewritten as

$$\hat{\mathbf{S}} \hat{\mathbf{T}}^2 \hat{\mathbf{S}}^H = \mathbf{H} \mathbf{H}^H + \mathbf{E}_D, \quad (66)$$

where the matrix  $\mathbf{E}_D$  is defined as

$$\mathbf{E}_D = \mathbf{H} \mathbf{E}_f \mathbf{H}^H + \mathbf{H} \mathbf{E}_{f_v} + \mathbf{E}_{f_v}^H \mathbf{H}^H + \mathbf{E}_v. \quad (67)$$

In the above expression, we have  $\mathbf{E}_f = (\mathbf{F}_D \mathbf{F}_D^H - \alpha_D \mathbf{I}_{N_t}) / \alpha_D$ ,  $\mathbf{E}_v = (\mathbf{V}_D \mathbf{V}_D^H - N_{D1} \mathbf{I}_{N_r}) / \alpha_D$  and  $\mathbf{E}_{f_v} = \mathbf{F}_D \mathbf{V}_D^H / \alpha_D$ . Note that, since the noise matrix  $\mathbf{V}_D$  and the spatially-white data matrix  $\mathbf{F}_D$  are mutually independent, the elements of the matrices  $\mathbf{E}_f$ ,  $\mathbf{E}_{f_v}$  and  $\mathbf{E}_v$  are pairwise uncorrelated. Thus, it follows



that  $\mathbb{E}[|\mathbf{E}_f(k, l)|^2] = (P_d/N_t)^2/(N_{D1}(P_D/N_t)^2) = 1/N_{D1}^2$ ,  $\mathbb{E}[|\mathbf{E}_v(k, l)|^2] = \sigma_v^4/(N_{D1}(P_D/N_t)^2) = N_{D1}\sigma_v^4/\alpha_D^2$  and  $\mathbb{E}[|\mathbf{E}_{fv}(k, l)|^2] = (\sigma_v^2 P_d/N_t)/(N_{D1}(P_D/N_t)^2) = \sigma_v^2/\alpha_D$ . Now, it follows from the result of first-order perturbation in (62) of Appendix B that  $\Delta c_i \approx \mathbf{s}_i^H \mathbf{E}_D \mathbf{s}_1 / (\sigma_1^2 - \sigma_i^2)$ . Thus, using (67), the quantity  $\mathbb{E}[|\Delta c_i|^2]$  can be expressed as

$$\mathbb{E}[|\Delta c_i|^2] \approx \left( \mathbb{E}[|\mathbf{s}_i^H \mathbf{H} \mathbf{E}_f \mathbf{H}^H \mathbf{s}_1|^2] + \mathbb{E}[|\mathbf{s}_i^H \mathbf{E}_{fv} \mathbf{H}^H \mathbf{s}_1|^2] \right. \\ \left. + \mathbb{E}[|\mathbf{s}_i^H \mathbf{H} \mathbf{E}_{fv} \mathbf{s}_1|^2] + \mathbb{E}[|\mathbf{s}_i^H \mathbf{E}_v \mathbf{s}_1|^2] \right) \frac{1}{(\sigma_1^2 - \sigma_i^2)^2}. \quad (68)$$

Exploiting  $\text{SVD}(\mathbf{H}) = \mathbf{S}\mathbf{\Gamma}\mathbf{Q}^H$  and the aforementioned properties, the different terms in the above expression can be evaluated as  $\mathbb{E}[|\mathbf{s}_i^H \mathbf{H} \mathbf{E}_f \mathbf{H}^H \mathbf{s}_1|^2] = \sigma_1^2 \sigma_i^2 / N_{D1}^2$ ,  $\mathbb{E}[|\mathbf{s}_i^H \mathbf{E}_{fv} \mathbf{H}^H \mathbf{s}_1|^2] = \sigma_v^2 \sigma_1^2 / \alpha_D$ ,  $\mathbb{E}[|\mathbf{s}_i^H \mathbf{H} \mathbf{E}_{fv} \mathbf{s}_1|^2] = \sigma_v^2 \sigma_i^2 / \alpha_D$  and  $\mathbb{E}[|\mathbf{s}_i^H \mathbf{E}_v \mathbf{s}_1|^2] = N_{D1} \sigma_v^4 / \alpha_D^2$ . Upon substituting these results into (68), we get

$$\mathbb{E}[|\Delta c_i|^2] \approx \left( \frac{\sigma_1^2 \sigma_i^2}{N_{D1}^2} + \frac{N_t \sigma_v^2 (\sigma_1^2 + \sigma_i^2)}{N_{D1} P_d} + \frac{N_t^2 \sigma_v^4}{N_{D1} P_d^2} \right) \\ \times \frac{1}{(\sigma_1^2 - \sigma_i^2)^2}. \quad (69)$$

#### APPENDIX E

Exploiting the properties of complex derivatives from [22], the matrix  $\mathbf{J}(\boldsymbol{\rho})$  can be constructed by differentiating  $\mathbf{f}(\boldsymbol{\rho})$  in (47). Furthermore, the matrix  $\mathbf{B} \in \mathbb{C}^{2N_t \times N_t}$  is derived using the property that it forms an orthonormal basis for the null-space of the matrix  $\mathbf{J}(\boldsymbol{\rho})$ . Both the matrices  $\mathbf{J}(\boldsymbol{\rho})$  and  $\mathbf{B}$  are expressed below.

$$\mathbf{J}(\boldsymbol{\rho}) = \begin{bmatrix} \mathbf{u}_1^H & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{u}_1^T & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{u}_1^H & \mathbf{0} & \cdots & \mathbf{u}_2^T & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{u}_2^H & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{u}_1^T & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{u}_2^H & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{u}_2^T & \mathbf{0} & \cdots \\ \mathbf{u}_3^H & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{u}_1^T & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{u}_1^H & \cdots & \mathbf{u}_3^T & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad \text{and}$$

$$\mathbf{B} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{u}_1 & \mathbf{0} & \mathbf{u}_2 & \mathbf{0} & \mathbf{u}_3 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{u}_1 & \mathbf{0} & \mathbf{u}_2 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{u}_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ -\mathbf{u}_1^* & -\mathbf{u}_2^* & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{u}_3^* & \cdots \\ \mathbf{0} & \mathbf{0} & -\mathbf{u}_1^* & \mathbf{u}_2^* & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{u}_1^* & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

#### APPENDIX F

##### CONSTRAINED CRLB FOR THE MMWAVE-HB-SB SCHEME

Upon employing (46), the matrix  $\mathbf{B}^H \mathbf{C}_\rho \mathbf{B}$  in (48) can be simplified to

$$\mathbf{B}^H \mathbf{C}_\rho \mathbf{B} = \frac{P_p N_B}{\sigma_v^2} \tilde{\mathbf{\Gamma}}. \quad (70)$$

Let the orthonormal basis matrix  $\mathbf{B}$  be partitioned as  $\mathbf{B} = [\mathbf{B}_1^T, \mathbf{B}_2^T]^T$ . Correspondingly, the matrix  $\text{CRLB}(\boldsymbol{\rho})$  in (48) can be partitioned as

$$\text{CRLB}(\boldsymbol{\rho}) = \begin{bmatrix} [\text{CRLB}(\boldsymbol{\rho})]_{1,1} & [\text{CRLB}(\boldsymbol{\rho})]_{1,2} \\ [\text{CRLB}(\boldsymbol{\rho})]_{2,1} & [\text{CRLB}(\boldsymbol{\rho})]_{2,2} \end{bmatrix}, \quad (71)$$

where each block is of size  $N_t^2 \times N_t^2$ . From (46) and (48), it is clear that the constrained CRLB for  $\text{vec}(\mathbf{U})$  can be obtained as  $\text{CRLB}(\boldsymbol{\rho})_{1,1} = (\sigma_v^2 / (P_p N_B)) \mathbf{B}_1 \tilde{\mathbf{\Gamma}}^{-1} \mathbf{B}_1^H$ . Since  $\mathbf{H} = \mathbf{D}\mathbf{U}^H$ , the parameter vector obeys  $\mathbf{h} = \text{vec}(\mathbf{H}^T) = [\mathbf{D} \otimes \mathbf{I}_{N_t}] \text{vec}(\mathbf{U}^*)$ . The error covariance of the mmWave MIMO channel matrix estimate  $\hat{\mathbf{H}}$  is bounded as

$$\mathbb{E}[(\hat{\mathbf{h}} - \mathbf{h})(\hat{\mathbf{h}} - \mathbf{h})^H] \geq [\mathbf{D} \otimes \mathbf{I}_{N_t}] ([\text{CRLB}(\boldsymbol{\rho})]_{1,1})^* [\mathbf{D} \otimes \mathbf{I}_{N_t}]^H \\ = \frac{\sigma_v^2}{P_p N_B} [\mathbf{S}\mathbf{\Gamma} \otimes \mathbf{I}_{N_t}] \mathbf{B}_1^* \tilde{\mathbf{\Gamma}}^{-1} \mathbf{B}_1^T [\mathbf{S}\mathbf{\Gamma} \otimes \mathbf{I}_{N_t}]^H = \mathbf{C}_H.$$

#### REFERENCES

- [1] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, 2017.
- [2] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [3] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, pp. 436–453, 2016.
- [4] K. Venugopal and R. W. Heath, "Millimeter wave networked wearables in dense indoor environments," *IEEE Access*, vol. 4, pp. 1205–1221, 2016.
- [5] M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Edge computing meets millimeter-wave enabled VR: Paving the way to cutting the cord," in *IEEE Wireless Communications and Networking Conference*. IEEE, 2018, pp. 1–6.
- [6] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, p. 4091, 2005.
- [7] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.
- [8] S. Srivastava, A. Mishra, A. Rajorija, A. K. Jagannatham, and G. Ascheid, "Quasi-static and time-selective channel estimation for block-sparse millimeter wave hybrid MIMO systems: Sparse Bayesian (SBL) based approaches," *IEEE Transactions on Signal Processing*, vol. 67, no. 5, pp. 1251–1266, 2019.
- [9] T. Nitsche, C. Cordeiro, A. B. Flores, E. W. Knightly, E. Perahia, and J. C. Widmer, "IEEE 802.11 ad: directional 60 GHz communication for multi-gigabit-per-second Wi-Fi," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 132–141, 2014.
- [10] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.
- [11] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370–2386, 2016.
- [12] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays," *IEEE Transactions on Signal Processing*, vol. 67, pp. 4809–4824, 2019.
- [13] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*. Pearson Education, 2014.
- [14] Z. Guo, X. Wang, and W. Heng, "Millimeter-wave channel estimation based on 2-D beamspace MUSIC method," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5384–5394, 2017.

- [15] A. Liao, Z. Gao, Y. Wu, H. Wang, and M. Alouini, "2d unitary ESPRIT based super-resolution channel estimation for millimeter-wave massive MIMO with hybrid precoding," *IEEE Access*, vol. 5, pp. 24 747–24 757, 2017.
- [16] W. Ma and C. Qi, "ESPRIT-based channel estimation for frequency-selective millimeter wave massive MIMO system," in *IEEE International Conference on Communications, ICC, Shanghai, China, May 20-24, 2019*. IEEE, 2019, pp. 1–6.
- [17] H. Chu, L. Zheng, and X. Wang, "Semi-blind millimeter-wave channel estimation using atomic norm minimization," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2535–2538, 2018.
- [18] S. Liang, X. Wang, and L. Ping, "Semi-blind detection in hybrid massive MIMO systems via low-rank matrix completion," *IEEE Trans. Wireless Communications*, vol. 18, no. 11, pp. 5242–5254, 2019.
- [19] A. Mezghani and A. L. Swindlehurst, "Blind estimation of sparse broadband massive MIMO channels with ideal and one-bit adcs," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2972–2983, 2018.
- [20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [21] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.
- [22] S. M. Kay, "Fundamentals of statistical signal processing, Volume i: Estimation theory," *PTR Prentice-Hall, Englewood Cliffs, NJ*, 1993.
- [23] C. Rusu, R. Mendez-Rial, N. González-Prelcic, and R. W. Heath, "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Tran. on Wireless Commun.*, vol. 15, no. 12, pp. 8380–8393, 2016.
- [24] J. H. Wilkinson, *The algebraic eigenvalue problem*. Oxford Clarendon, 1965, vol. 662.
- [25] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, p. 15, 2008.
- [26] R. Vershynin, "How close is the sample covariance matrix to the actual covariance matrix?" *Journal of Theoretical Probability*, vol. 25, no. 3, pp. 655–686, 2012.
- [27] A. K. Jagannatham and B. D. Rao, "Cramér-Rao lower bound for constrained complex parameters," *IEEE Signal Processing Letters*, vol. 11, no. 11, pp. 875–878, 2004.
- [28] A. van den Bos, "A Cramér-Rao lower bound for complex parameters," *IEEE Trans. Signal Processing*, vol. 42, no. 10, p. 2859, 1994.
- [29] P. Stoica and B. C. Ng, "On the Cramér-Rao bound under parametric constraints," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 177–179, 1998.
- [30] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Processing*, vol. 54, no. 9, pp. 3658–3663, 2006.
- [31] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU Press, 2012, vol. 3.
- [32] P. Singh, S. Srivastava, A. K. Jagannatham, and L. Hanzo, "Technical report: Second-order statistics-based semi-blind techniques for channel estimation in millimeter wave mimo systems with analog and hybrid beamforming." IIT Kanpur, Tech. Rep., 2020. [Online]. Available: [http://www.iitk.ac.in/mwn/documents/MWNLab\\_TR\\_mmWave\\_SB\\_MIMO.pdf](http://www.iitk.ac.in/mwn/documents/MWNLab_TR_mmWave_SB_MIMO.pdf).
- [33] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3959–3973, 2018.
- [34] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 34, no. 2, pp. 331–341, 1986.