Utilizing Attitude Information for Efficient Multi-User Millimeter-Wave Communications

Mingrui Li, Xiaowei Qin, Yunfei Chen, Weidong Wang, and Li Chen

***Abstract*—Sensing has played an important role in the 5G wireless network. It is well recognized that the energy efficiency of mobile devices can be further improved with the assistance of sensing. As an important sensing resource, attitude information is highly related to the angle of arrivals (AoAs) information when user equipments (UEs) rotate. It provides great potential to improve transceiver designs for multi-user millimeter-wave (mmWave) green communications. In this paper, we propose an attitude information aided block diagonalization receiver design algorithm. We first adopt angle of departure aided block diagonalization (AoD-BD) to eliminate inter-user interference (IUI). Then we compensate the rotation for each UE by only updating the receiver with attitude information from motion sensors. Compared to the joint transceiver design performance, we have theoretically proved that the sum spectral efficiency performance penalty of keeping transmitter unchanged is generally negligible. Furthermore, we consider the effect of imperfect attitude information, and derive a robust receiver by modeling the measurement error as Gaussian. Finally, the simulation results are presented to verify the effectiveness of the proposed design. It is noted that no feedback is required between the base station (BS) and UE during the rotation, which substantially reduces the system overhead.**

***Index Terms*—Attitude information, block diagonalization, millimeter-wave (mmWave), transceiver design, user equipment rotation**

I. Introduction

With the emergence of many applications in 5G wireless network, the high-quality wireless connectivity requires to be guaranteed. Meanwhile, in millimeter-wave (mmWave) systems, the massive increase in the number of antennas and users makes the problem much more challenging due to the high energy consumption [1]. To alleviate the problem, sensing-assisted communication is regarded as a promising way to further improve the energy efficiency [2], [3]. A typical example is the spectrum sensing in the context of cognitive radio, where the secondary user detects the presence of the primary user over a frequency band of interest, and then utilizes the idle spectrum resources to transmit information [4].

Recently, sensing-assisted communication is studied to be used in the beamforming design for mmWave communications [5], [6]. One key issue is the acquisition of channel state information (CSI). CSI is crucial for beamforming and often obtained through channel training. As the overhead of channel training grows with the number of antennas, it is necessary to take some measures to reduce the high training cost.

Traditional methods of estimating CSI require the transmitter to emit the training signals [7], [8], which might cause much feedback between the transmitter and the receiver. The overhead of conventional least squares (LS) method is given by $\mathcal{O}(N_T)$, where N_T is the number of transmit antennas. Since the channels are sparse in practical mmWave systems, some works [9], [10] exploited compressed sensing (CS) techniques to help perform explicit channel estimation, whose pilot overhead is given by $\mathcal{O}(L\ln N_T)$, where L is the sparsity level. The work in [11] further showed how many CS measurements are needed to approach the performance with perfect channel knowledge.

Compared to the aforementioned methods, sensing-assisted communication provides great potential to further reduce the training overhead by utilizing out-of-band information [12]–[15]. Out-of-band information is mainly divided into three categories: position information, signals from low frequency systems and attitude information. The main advantage of these information is that they do not consume resources in the designed system and can be easily obtained from sensors or other communication systems. Position information could help limit the range of beam directions. The work in [16] utilized the position information to help determine the angle of departure (AoD) angular support. A location-aided beamforming strategy was proposed in [17] for vehicular communications, and the results showed that it could significantly speed up the initial access. Since mmWave cellular systems are likely to be deployed in parallel with sub-6G communication systems, the lower frequency signals could be exploited to help estimate parameters of mmWave channel based on spatial correlation translation [18].

Attitude information has high correlations with beam direction. Thus, it is often used to overcome the effects of beam misalignment, which might severely deteriorate the link quality [19]. The work in [20] proposed an algorithm that used position-attitude information to help align the beams and complete beam tracking, which outperformed the existing algorithms with only position information. In [21], the authors revealed a new beam alignment mechanism based on IEEE 802.11ad that can dramatically reduce the system overhead. A beam tracking algorithm was developed to compensate the misalignment caused by the change of device behaviors in [22]. Attitude information has also been used to help reconstruct the channel matrix when user equipment (UE)

This research was supported by National Key R&D Program of China (Grant No. 2018YFA0701603). *(Corresponding author: Xiaowei Qin.)*

Mingrui Li, Xiaowei Qin, Weidong Wang and Li Chen are with CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China, Hefei 230027, Anhui, China (e-mail: lmingr@mail.ustc.edu.cn; qinxw@ustc.edu.cn; wdwang@ustc.edu.cn; chenli87@ustc.edu.cn).

Yunfei Chen is with the Department of Engineering, University of Durham, Durham, UK, DH1 3LE (e-mail: Yunfei.Chen@durham.ac.uk).

rotates. The work in [23] exploited attitude information to help the transmitter design in an uplink digital beamforming system.

All the the existing attitude information aided (AIA) transceiver designs discussed above are only proposed for single-user scenarios. How to utilize attitude information to reduce the system overhead in a multi-user system is still an open problem. As the training overhead grows with the number of users, it is of great significance to apply AIA methods for multi-user scenarios. Compared to single-user scenarios, AIA multi-user transceiver designs will be more challenging. First, the inter-user interference (IUI) makes the performance corresponding to different UEs coupled together. Specifically, when we compensate the rotation for one UE, other UEs need to be re-calibrated too. Second, since the attitude information is measured at UEs, joint transceiver design might generate too much feedback overhead and cause severe time delay. The most efficient method is to compensate the rotation by only updating the receiver. Third, compared to the joint transceiver design, it is unknown whether keeping the transmitter unchanged could help the system achieve the near-optimal spectral efficiency. A detailed discussion on performance loss is in need.

In this article, we study AIA transceiver designs for multi-user mmWave communications. First, we adopt AoD aided Block diagonalization (AoD-BD) method to eliminate IUI among different UEs. This is to ensure that one UE rotation will not affect other UEs. Then, the problem is formulated as a receiver design problem for each UE by keeping the transmitter unchanged during the rotation. A closed-form solution with the assistance of attitude information is derived by maximizing the sum spectral efficiency. Thus, an AIA-BD receiver design algorithm is proposed. Further, compared to the optimal sum spectral efficiency without rotation, we prove the superiority of AIA-BD by analyzing its performance. In addition, by modeling the measurement error of attitude information as Gaussian, we derive a robust AIA-BD receiver design. The main contributions of the article can be summarized as follows.

1) *AIA-BD receiver design algorithm:* Considering a block-fading channel, since the rotation may cause severe degradation of the sum spectral efficiency, we propose a low-complexity AIA-BD receiver design algorithm to compensate this loss. In our proposed algorithm, the system only needs to establish the initial link and BS keeps the precoding matrix unchanged during the subsequent rotation. Each UE adjusts the combining matrix with the help of attitude information from motion sensors. Our proposed algorithm efficiently reduces the training overhead and no feedback is required during the rotation.
2) *Theoretical performance of the proposed design:* The performance gap between the proposed AIA-BD receiver design and the optimal sum spectral efficiency before rotation mainly contains two parts. The first part is caused by the rotation. We have proved that the rotation will not affect the achievable sum spectral efficiency obtained by joint transceiver design when the number of user antennas approaches infinity. The second part comes from keeping the transmitter unchanged. We derive an upper bound of performance loss between AIA-BD and optimal joint transceiver design after UEs rotation. The results show that such performance loss is generally trivial in mmWave systems.
3) *Robust AIA-BD receiver with measurement error:* Due to hardware limitation, the precise attitude information is difficult to obtain. Thus, we take the measurement error of attitude information into consideration and reformulate the receiver design problem. By modeling the measurement error as Gaussian, we derive a robust AIA-BD receiver based on the Taylor expansion. The simulation results show that the proposed receiver could obtain a good performance even with a relative large measurement error.

The remainder of the paper is organized as follows. Section II describes the mmWave network and how it works under AoD-BD method. In section III, we present our AIA-BD method for the multi-user scenario. Section IV discusses the robustness of attitude information aided receiver design. Finally, simulation results are presented in Section V, while Section VI concludes the paper.

Notations: Bold uppercase $\mathbf{A}$ denotes a matrix and bold lowercase $\mathbf{a}$ denotes a vector. $\mathbf{A}^T$,$\mathbf{A}^H$, and $\mathbf{A}^{-1}$ denotes the transpose, conjugate transpose, and the inverse of $\mathbf{A}$ respectively. $\|\mathbf{A}\|_F$ is the Frobenius norm of $\mathbf{A}$. $tr(\mathbf{A})$ is its trace and $|\mathbf{A}|$ is its determinant; $\mathrm{diag}(a_1,\ldots,a_N)$ represents a block diagonal matrix whose diagonal entries are given by $\{a_1,\ldots,a_N\}$; $\mathbb{E}[\cdot]$ is the expectation operator; $\mathcal{CN}(\mu,\mathbf{\Sigma})$ denotes a Guassian random variable with mean μ and the variance $\mathbf{\Sigma}$.

II. System Model

A. Transmission Model

Consider a multi-user mmWave system that consists of one base station (BS) with N_t antennas and K UEs. Each UE is equipped with N_r antennas. We focus on the multi-user digital beamforming case in which the BS communicates with every UE via N_s streams, with $N_s \leq \min(N_t, N_r)$. Therefore, the total number of streams is KN_s. We consider a narrowband block-fading channel model which is widely used in the studies on beamforming [24], [25]. Then the received signal observed by the j-th UE can be written as

$$\mathbf{y}_j = \mathbf{W}_j^H \mathbf{H}_j \mathbf{F}\mathbf{x} + \mathbf{W}_j^H \mathbf{n}_j, \tag{1}$$

where $\mathbf{W}_j \in \mathbb{C}^{N_r \times N_s}$ is the combining matrix of the j-th UE and normalized to satisfy $\|\mathbf{W}_j\|_F^2 = N_s$, $\mathbf{F} \in \mathbb{C}^{N_t \times KN_s}$ is the transmit precoding matrix and normalized to satisfy $\|\mathbf{F}\|_F^2 = KN_s$. $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_K^T]^T \in \mathbb{C}^{KN_s}$ is the transmit signal vector in which the j-th UE's symbols are represented by $\mathbf{x}_j \in \mathbb{C}^{N_s}$, with $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \rho \mathbf{I}_{KN_s}$, where $\rho = P/KN_s$, and P is the average total transmitted power. Finally, $\mathbf{n}_j \sim \mathcal{CN}(\mathbf{0}, \sigma_j^2 \mathbf{I}) \in \mathbb{C}^{N_r}$ is the Gaussian noise corrupting the received signal.

Since mmWave channels are expected to have limited scattering characteristic, we adopt a narrowband geometric channel model between the BS and each UE [25]. The channel

$\mathbf{H}_j$ is assumed to be a sum of the contributions of L_j scattering clusters. Each cluster consists of one propagation path. Therefore, the discrete-time narrowband channel $\mathbf{H}_j$ can be written as

$$\mathbf{H}_j = \sqrt{\frac{N_t N_t}{L_j}} \sum_{i=1}^{L_j} \alpha_i^{(j)} \mathbf{a}_r(\theta_i^{(j)}) \mathbf{a}_t^H(\phi_i^{(j)}), \tag{2}$$

$\alpha_i^{(j)}$ is the complex gain of i-th scattering cluster. We assume that $L_j = L, 1 \leq j \leq K$. The variables $(\theta_i^{(j)}, \phi_i^{(j)}) \in [0, 2\pi]$ represent AoA and AoD of BS and UE. They are assumed to be randomly distributed with a constant angular spread (standard deviation) of σ_θ^2 and σ_ϕ^2 respectively. $\mathbf{a}_t(\phi)$ and $\mathbf{a}_r(\theta)$ are the antenna array response vectors of the uniform linear array (ULA) at the BS and UE, respectively, which can be written as

$$\mathbf{a}_t(\phi) = \frac{1}{\sqrt{N_t}} [1, e^{j\frac{2\pi}{\lambda} d \cos\phi}, \ldots, e^{j\frac{2\pi}{\lambda} d(N_t-1)\cos\phi}]^T, \tag{3}$$

$$\mathbf{a}_r(\theta) = \frac{1}{\sqrt{N_r}} [1, e^{j\frac{2\pi}{\lambda} d \cos\theta}, \ldots, e^{j\frac{2\pi}{\lambda} d(N_r-1)\cos\theta}]^T, \tag{4}$$

where λ is the signal wavelength, and d is the distance between antenna elements. Thus, equation (2) can be rewritten as

$$\mathbf{H}_j = \mathbf{A}_r^{(j)} \mathbf{\Sigma}^{(j)} \mathbf{A}_t^{(j)^H} \tag{5}$$

where

$$\mathbf{A}_r^{(j)} = [\mathbf{a}_r(\theta_1^{(j)}), \mathbf{a}_r(\theta_2^{(j)}), \ldots, \mathbf{a}_r(\theta_L^{(j)})], \tag{6}$$

$$\mathbf{A}_t^{(j)} = [\mathbf{a}_t(\phi_1^{(j)}), \mathbf{a}_t(\phi_2^{(j)}), \ldots, \mathbf{a}_t(\phi_L^{(j)})], \tag{7}$$

$$\mathbf{\Sigma}^{(j)} = \sqrt{\frac{N_t N_t}{L}} \mathrm{diag}(\alpha_1^{(j)}, \alpha_2^{(j)}, \ldots, \alpha_L^{(j)}). \tag{8}$$

Note the results presented in this paper can be extended for other antenna array structures as well. Assuming equal power allocation across all the user streams by the BS, the sum spectral efficiency is given by [26] and [27]

$$R(\mathbf{F}, \mathbf{W}) = \sum_{j=1}^{K} \log_2 \left(\left| \mathbf{I} + \rho \mathbf{R}_j^{-1} \mathbf{W}_j^H \mathbf{H}_j \mathbf{F}_j \mathbf{F}_j^H \mathbf{H}_j^H \mathbf{W}_j \right| \right), \tag{9}$$

where $\mathbf{R}_j^{-1} = \rho \sum_{i=1, i\neq j}^{K} \mathbf{W}_j^H \mathbf{H}_j \mathbf{F}_i \mathbf{F}_i^H \mathbf{H}_j^H \mathbf{W}_j + \sigma_j^2 \mathbf{W}_j^H \mathbf{W}_j$ is the interference from other UEs plus the noise covariance matrix.

B. Block Diagonalization Algorithm with Subspace Projection Based AoD

In this section, we adopt BD methods to eliminate IUI. Conventional BD algorithms [27], [28] require the BS to have full CSI, which might lead to a lower efficiency when the wireless channel changes rapidly. Since mmWave channel shows the spatial-domain sparsity and geometric properties, it is sufficient for achieving BD of the composite multi-user channel with only AoD message associated with every UE [29], [30]. Thus, we explain how BD with subspace projection based AoD (SP-AoD-BD) works as follows.

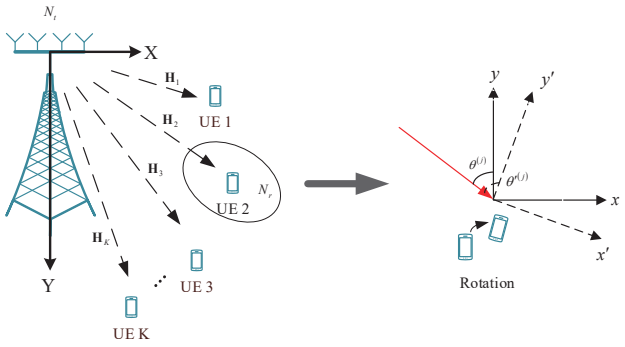


Fig. 1. X-Y represents the global coordinate system of BS while x-y represents the local coordinate system of UE. The dotted line represents the axis after rotation. ULA is implemented along x axis and the first element of antenna array is at the origin point. The red arrow indicates the direction of the propagation path.

Let $\mathbf{H}_{tot} = \left[\mathbf{H}_1^T, \mathbf{H}_2^T, \ldots, \mathbf{H}_K^T\right]^T$ denote the composite multi-user channel. From equation (5), it can be written as

$$\mathbf{H}_{tot} = \begin{bmatrix} \mathbf{A}_r^{(1)} \mathbf{\Sigma}^{(1)} & & \\ & \ddots & \\ & & \mathbf{A}_r^{(K)} \mathbf{\Sigma}^{(K)} \end{bmatrix} \begin{bmatrix} \mathbf{A}_t^{(1)^H} \\ \vdots \\ \mathbf{A}_t^{(K)^H} \end{bmatrix}, \tag{10}$$

where $[\mathbf{A}_t^{(1)}, \mathbf{A}_t^{(2)}, \ldots, \mathbf{A}_t^{(K)}]^H$ is known to the BS. For simplicity, assume the rank of the channel matrix of each UE is $r = L$. Let $\tilde{\mathbf{A}}_t^{(j)} = [\mathbf{A}_t^{(1)}, \ldots, \mathbf{A}_t^{(j-1)}, \mathbf{A}_t^{(j+1)}, \ldots, \mathbf{A}_t^{(K)}]^H = \tilde{\mathbf{U}}^{(j)} \tilde{\mathbf{\Lambda}}^{(j)} \tilde{\mathbf{V}}^{(j)^H}$, and we could opt for

$$\mathbf{F}_{1\_null}^{(j)} = \tilde{\mathbf{V}}^{(j)}(:, (K-1)L + 1 : N_t), \tag{11}$$

which corresponds to the orthogonal complement of the interference subspace [30]. Arbitrarily choosing L columns in (11) as the first precoding matrix to eliminate IUI might result in very poor performance, since the choice of basis may potentially be orthogonal to the user-signal. Therefore, a set of basis vectors that align with signal space needs to be carefully chosen, which can be handled with subspace projection.

Let $\mathbf{A}_t^{(j)^H} = \bar{\mathbf{U}}^{(j)} \bar{\mathbf{\Lambda}}^{(j)} \bar{\mathbf{V}}^{(j)^H}$, and we opt for $\mathbf{F}_{1\_sig}^{(j)} = \bar{\mathbf{V}}^{(j)}(:, 1 : L)$ which corresponds to the signal space. Let $\mathbf{P}_{1\_null}^{(j)}$ and $\mathbf{P}_{1\_sig}^{(j)}$ represent the projection matrices associated with $\mathbf{F}_{1\_null}^{(j)}$ and $\mathbf{F}_{1\_sig}^{(j)}$ respectively. Then we could obtain the first precoding matrix by performing singular value decomposition (SVD) on the matrix $\mathbf{P}_{1\_sig}^{(j)} \mathbf{P}_{1\_null}^{(j)}$, which can be written as

$$\mathbf{F}_1^{(j)} = \mathbf{V}_1^{(j)}(:, 1 : L) \in \mathbb{C}^{N_t \times L}, \tag{12}$$

where $\mathbf{V}_1^{(j)}$ is the right singular matrix of $\mathbf{P}_{1\_sig}^{(j)} \mathbf{P}_{1\_null}^{(j)}$. We can find that the first precoding matrix not only satisfies $\mathbf{H}_i \mathbf{F}_1^{(j)} = \mathrm{O}_{N_r \times L}, 1 \leq i \neq j \leq K$, but also align with the j-th user-signal space $\mathbf{F}_{1\_sig}^{(j)}$. Then, the total effective channel matrix will present a block structure as

$$\mathbf{H}_{tot}^{(eff)} = \mathbf{H}_{tot} \mathbf{F}_{tot1} = \begin{bmatrix} \mathbf{H}_1 \mathbf{F}_1^{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{H}_K \mathbf{F}_1^{(K)} \end{bmatrix}. \tag{13}$$

This allows SVD to be determined individually to maximize the spectral efficiency for each UE when equal power is allocated across streams with high SNR [31]. Define the effective channel matrix of j-th UE and its ordered SVD as

$$\mathbf{H}_j^{(eff)} = \mathbf{A}_r^{(j)}\mathbf{\Sigma}^{(j)}\mathbf{A}_t^{(j)^H}\mathbf{F}_1^{(j)} = \mathbf{U}_2^{(j)}\mathbf{\Lambda}_2^{(j)}\mathbf{V}_2^{(j)^H}, \quad (14)$$

where $\mathbf{U}_2^{(j)}$ and $\mathbf{V}_2^{(j)}$ are unitary matrices and $\mathbf{\Lambda}_2^{(j)}$ contains the singular values arranged in decreasing order. Furthermore, we define the following partitions

$$\begin{aligned}\mathbf{U}_2^{(j)} &= \left[\mathbf{U}_a^{(j)}\mathbf{U}_b^{(j)}\right],\\ \mathbf{\Lambda}_2^{(j)} &= \begin{bmatrix}\mathbf{\Lambda}_a^{(j)} & \\ & \mathbf{\Lambda}_b^{(j)}\end{bmatrix},\\ \mathbf{V}_2^{(j)} &= \left[\mathbf{V}_a^{(j)}\mathbf{V}_b^{(j)}\right],\end{aligned} \quad (15)$$

where $\mathbf{U}_a^{(j)} \in \mathbb{C}^{N_r\times N_s}$, $\mathbf{V}_a^{(j)} \in \mathbb{C}^{L\times N_s}$ and $\mathbf{\Lambda}_a^{(j)} \in \mathbb{C}^{N_s\times N_s}(N_s \le L)$.

Then we obtain the combining matrix and the second precoding matrix, which eliminates the inter-stream interference. They can be written as

$$\mathbf{W}_j = \mathbf{U}_a^{(j)}, \quad (16)$$

$$\mathbf{F}_2^{(j)} = \mathbf{V}_a^{(j)}. \quad (17)$$

Thus, the final precoding matrix of the j-th UE is given by

$$\mathbf{F}_j = \mathbf{F}(:;(j-1)N_s : jN_s) = \mathbf{F}_1^{(j)}\mathbf{F}_2^{(j)}. \quad (18)$$

With each UE's combining matrix chosen as (16) and precoding matrix chosen as (18), the sum spectral efficiency in (9) becomes

$$R = \sum_{j=1}^{K} R_j = \sum_{j=1}^{K}\log_2\left(\left|\mathbf{I} + \frac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^{(j)^2}\right|\right). \quad (19)$$

C. Rotation Models of UEs

We will analyze how the rotations affect the mmWave channel and transceiver designs in this section. Consider two adjacent channel estimations within a block. The block-length is related to the channel coherence time, which is generally small in mmWave systems. Thus, it is assumed that there is no large-scale movement of UE or environment variation. As to the small-scale displacement and rotations of UEs, the link quality is mainly affected by the latter one since the small-scale displacement of UEs can be ignored compared to the distance from the BS [19], [32]. Therefore, we can focus on the effects of UEs' rotations on the mmWave channel.

1) Before the rotation: We take the j-th UE as an example. The channel between BS and UE is given by (5) and the transceiver designs are given by (16) and (18).

2) After the rotation: As shown in Fig. 1, since the block-fading model applies to a channel in which several adjacent symbols are affected by the same fading value. It is believed that the surrounding environment is basically unchanged during the two adjacent estimations [32]–[34]. Therefore, the distribution of propagation paths in the global coordinate system at BS remains unchanged. The AoDs at BS are not affected by the rotation of UE. Moreover, the distance of UE from the BS does not change. This means the path loss coefficient is also a constant [23]. The rotation only changes the AoAs in the local coordinate system at UE [19]. The AoAs at UE change from $\theta_i^{(j)}$ to $\theta_i^{(j)} + \Delta\theta^{(j)}$. Here we use the superscript $'$ to denote the parameters after the rotation. Then, we can reconstruct the channel matrix as

$$\mathbf{H'}_j = \mathbf{A'}_r^{(j)}\mathbf{\Sigma}^{(j)}\mathbf{A}_t^{(j)^H} = \mathbf{C}_j\mathbf{H}_j, \quad (20)$$

where $\mathbf{A'}_r^{(j)} = [\mathbf{a}_r(\theta_1^{(j)} + \Delta\theta^{(j)}), \ldots, \mathbf{a}_r(\theta_L^{(j)} + \Delta\theta^{(j)})]$ is the array response matrix after the rotation and $\mathbf{C}_j = \mathbf{A'}_r^{(j)}(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}$ describes the relationship between the channel matrix before and after the rotation.

Remark 1 (Zero feedback transceiver design): Since attitude information is measured at UEs during the rotation, updating CSI at BS will generate too much overhead for feedback. Note that AoD message is sufficient for achieving BD structure. Thus, we consider keeping the precoding unchanged and leaving UEs to compensate the rotation with the observed AoA variations from motion sensors. This approach will be discussed further in Section III.

III. AIA-BD Receiver Design

In this section, we propose the AIA-BD algorithm that redesign the receiver by using attitude information. The proposed algorithm leverages the structure of block diagonalization to decouple different UEs and provides a near-optimal sum spectral efficiency after UEs rotation.

A. AIA-BD Receiver Design Algorithm

If the overall precoding matrix is kept to be the same as before the rotation $\mathbf{F'}_j = \mathbf{F}_j = \mathbf{F_1}^{(j)}\mathbf{F_2}^{(j)}, 1 \le j \le K$, the sum spectral efficiency will be a function of the combining matrix. We first simplify it by substituting the precoders (18) into (9). Note that some superscripts of singular matrices will be dropped in the derivation for simplicity.

Substituting the precoders (18) into (9) and using (14) and (20), one has

$$\begin{aligned}&R'_j(\mathbf{W'}_j)\\ &= \log_2\left(\left|\mathbf{I}_{N_s} + \rho(\sigma_j^2\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}\right.\right.\\ &\quad\left.\left.\times\,\mathbf{W'}_j^H\mathbf{H'}_j\mathbf{F}_j\mathbf{F}_j{}^H\mathbf{H'}_j^H\mathbf{W'}\right|\right)\\ &= \log_2\left(\left|\mathbf{I}_{N_s} + \rho(\sigma_j^2\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}\ \mathbf{W'}_j^H\mathbf{A'}_r^{(j)}\mathbf{\Sigma}^{(j)}\right.\right.\\ &\quad\left.\left.\times\,\mathbf{A}_t^{(j)^H}\mathbf{F}_1^{(j)}\mathbf{V}_a^{(j)}\mathbf{V}_a^{(j)^H}\mathbf{F}_1^{(j)^H}\mathbf{A}_t^{(j)}\mathbf{\Sigma}^{(j)^H}\mathbf{A'}_r^{(j)^H}\mathbf{W'}_j\right|\right)\\ &= \log_2\left(\left|\mathbf{I}_{N_s} + \rho(\sigma_j^2\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_2^{(j)}\mathbf{\Lambda}_2^{(j)}\right.\right.\\ &\quad\left.\left.\times\,\mathbf{V}_2^{(j)^H}\mathbf{V}_a^{(j)}\mathbf{V}_a^{(j)^H}\mathbf{V}_2^{(j)}\mathbf{\Lambda}_2^{(j)^H}\mathbf{U}_2^{(j)^H}\mathbf{C}_j^H\mathbf{W'}_j\right|\right)\\ &= \log_2\left(\left|\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}(\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}\right.\right.\\ &\quad\left.\left.\times\,\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a^{(j)}\mathbf{\Lambda}_a^{(j)}\mathbf{\Lambda}_a^{(j)^H}\mathbf{U}_a^{(j)^H}\mathbf{C}_j^H\mathbf{W'}_j\right|\right)\\ &= \log_2\left(\left|\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\right.\right.\\ &\quad\left.\left.\times\,(\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a\right|\right).\end{aligned} \quad (21)$$

It is noted that $\mathbf{W'}_j(\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}\mathbf{W'}_j^H$ is an orthogonal projective operator which has Hermitian and idempotent properties. Hence we could perform Schmidt Orthogonalization on $\mathbf{W'}_j$ as

$$\begin{aligned}&\mathbf{W'}_j\mathbf{P}_j\Big((\mathbf{W'}_j\mathbf{P}_j)^H\mathbf{W'}_j\mathbf{P}_j\Big)^{-1}(\mathbf{W'}_j\mathbf{P}_j)^H\\&=\mathbf{W'}_j\mathbf{P}_j\left(\mathbf{P}_j^{-1}(\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}(\mathbf{P}_j^H)^{-1}\right)\mathbf{P}_j^H\mathbf{W'}_j^H\\&=\mathbf{W'}_j(\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}\mathbf{W'}_j^H,\end{aligned}\tag{22}$$

where $\mathbf{P}_j \in \mathbb{C}^{N_s\times N_s}$ is an invertible matrix that consists of elementary transforms of Schmidt Orthogonalization. Therefore we can simplify $(\mathbf{W'}_j^H\mathbf{W'}_j)^{-1}$ to the identity matrix $\mathbf{I}_{Ns}$.

Although we impose the constraints $\mathbf{W'}_j^H\mathbf{W'}_j = \mathbf{I}_{Ns}$ on the combining matrix, the orthogonal basis transform has no effect on the column space of the matrix and hence, the constraints do not increase the complexity of the problem. Thus, the sum spectral efficiency in (9) becomes

$$\begin{aligned}\sum_{j=1}^{K} R'_j(\mathbf{W'}_j) = \quad &\sum_{j=1}^{K}\log_2\Big(\Big|\mathbf{I}_{N_s}+\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\mathbf{U}_a^H\mathbf{C}_j^H\\&\times\ \mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a\Big|\Big),\end{aligned}\tag{23}$$

where $\mathbf{W'}_j^H\mathbf{W'}_j = \mathbf{I}_{Ns}$ for $1 \le j \le K$.

The above object function is still non-concave according to the composition rules. Directly maximizing the sum spectral efficiency in (23) is intractable. The following proposition can be used to obtain an equivalent problem.

Proposition 1 (Approximate problem of maximizing the sum spectral efficiency): Maximizing the sum spectral efficiency in (23) can be approximated to K orthogonal Procrustes problems

$$\begin{aligned}\mathbf{P1}: &\min_{\mathbf{W'}_j}\ \|\mathbf{W'}_j-\mathbf{C}_j\mathbf{U}_a\|_F^2, 1\le j\le K\\&\mathbf{s.t.}\mathbf{W'}_j^H\mathbf{W'}_j=\mathbf{I}_{Ns}.\end{aligned}\tag{24}$$

Proof: This is given in the appendix A ■

From Proposition 1, we can easily obtain the solution by using Lagrange Multiplier method [35]. We first simplify the objective function

$$\begin{aligned}&\|\mathbf{W'}_j-\mathbf{C}_j\mathbf{U}_a\|_F^2 = tr\left((\mathbf{W'}_j-\mathbf{C}_j\mathbf{U}_a)^H(\mathbf{W'}_j-\mathbf{C}_j\mathbf{U}_a)\right)\\&=tr(\mathbf{W'}_j^H\mathbf{W'}_j)-2tr(\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a)+tr(\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{C}_j\mathbf{U}_a)\\&=-2tr(\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a)+N_s+tr(\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{C}_j\mathbf{U}_a).\end{aligned}\tag{25}$$

Since N_s and $tr(\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{C}_j\mathbf{U}_a)$ are constant, the problem will become

$$\begin{aligned}&\max_{\mathbf{W'}_j} tr(\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a)\\&\mathbf{s.t.}\mathbf{W'}_j^H\mathbf{W'}_j=\mathbf{I}_{Ns}.\end{aligned}\tag{26}$$

The Lagrange function of (26) is

$$L(\mathbf{W'}_j,\mathbf{\Lambda}_\lambda) = tr(\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a)-tr(\mathbf{\Lambda}_\lambda(\mathbf{W'}_j^H\mathbf{W'}_j-\mathbf{I}_{Ns}))\tag{27}$$

where $\mathbf{\Lambda}_\lambda \in \mathbb{C}^{N_s\times N_s}$ is the Lagrange multiplier matrix. The derivative of $L(\mathbf{W'}_j,\mathbf{\Lambda}_\lambda)$ can be calculated $\partial L(\mathbf{W'}_j,\mathbf{\Lambda}_\lambda)/\partial\mathbf{W'}_j = \mathbf{C}_j\mathbf{U}_a-\mathbf{W'}_j(\mathbf{\Lambda}_\lambda+\mathbf{\Lambda}_\lambda^H)$. Thus we can find that the maximum value is obtained when $\mathbf{W'}_j^H$ satisfies $\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a = (\mathbf{\Lambda}_\lambda+\mathbf{\Lambda}_\lambda^H)$. Assuming the singular value decomposition $\mathbf{C}_j\mathbf{U}_a = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$, we have

$$tr(\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a) = tr(\mathbf{W'}_j^H\mathbf{U}\mathbf{\Lambda}\mathbf{V}^H) = tr(\mathbf{V}^H\mathbf{W'}_j^H\mathbf{U}\mathbf{\Lambda}).\tag{28}$$

Therefore, the optimal solution of P1 is

$$\mathbf{W'}_j = \mathbf{U}(:,1:N_s)\mathbf{V}^H.\tag{29}$$

However, the solution still contains one SVD operation which could be further simplified. From the perspective of SVD decomposition, the essence of the receiver's compensation is to align the column space with the wireless channel column space after rotation. We note that the receiver design in (29) has the same column space as $\mathbf{C}_j\mathbf{U}_a$. We can choose $\mathbf{W'}_j = \alpha\mathbf{C}_j\mathbf{U}_a$ as the final combining matrix solution, where α makes $\|\mathbf{W'}_j\|_F^2 = N_s$. Notice $\mathbf{U}_a$ is just the receiver before rotation and $\mathbf{C}_j$ can be calculated with attitude information.

To conclude, with the transmitter unchanged during the rotation, we have the sub-optimal receiver

$$\mathbf{W'}_j = \alpha\mathbf{C}_j\mathbf{U}_a,\tag{30}$$

and the whole procedure is given in Alogrithm 1. In the first stage, the link configuration is completed and we obtain the initial precoding matrix and combining matrix with the help of channel training. The initial AoD message used for block diagonalization at BS can be achieved through uplink channel sounding as detailed in [11], [36]. The initial AoA knowledge can be obtained with the help of channel estimation algorithm, such as MUSIC algorithm [37]. In the second stage, UEs compensate for the rotation with attitude information during each channel estimation period. The max channel estimation times in one block is indicated by the variable $max\_estimation$, which can be calculated with channel coherent time and estimation interval.

Remark 2 (Other antenna array structures): The proposed AIA-BD receiver design algorithm is also applicable to other antenna arrays. The different antenna array structures only result in different ways we reconstruct the AoAs at receivers. For example, considering a three-dimension beamforming, UE rotation will change the elevation angles and azimuth angles of arrivals in the local coordinate system. However, in the global coordinate system at BS, the absolute directions of propagation paths would never change. By utilizing this characteristic, the AoAs can be reconstructed through coordinate transformation with the help of attitude information. Furthermore, the entire derivations for the receiver design do not involve the expressions of antenna array response vectors. thus the proposed algorithm can be easily extended to the other antenna array structures.

Remark 3 (Complexity analysis): Since the initial precoding matrix and combining matrix for each UE are computed by using SP-AoD-BD algorithm. Thus, the complexity of our proposed algorithm is the same as SP-AoD-BD algorithm in stage 1. In the adaptive compensation phase, with the transmitter unchanged, attitude information is utilized to help redesign the receiver for each UE. This significantly reduces the feedback overhead compared to the SP-AoD-BD joint transceiver design algorithm.

Algorithm 1 AIA-BD Receiver Design Algorithm

Require: $k=0$, $max\_estimation$, $\Delta\theta^{(j)}$, K
Stage1(Link Configuration):
1) For $j=1,2,\dots K$: Compute $\mathbf{F}_1^{(j)}$ with $\mathbf{A}_t$ in (11)
2) Determine $\mathbf{A}_r$ and $\mathbf{\Sigma}$ by downlink channel training and construct the effective channel matrix $\mathbf{H}_j^{(eff)}$
3) For $j=1,2,\dots K$: Compute $\mathbf{F}_2^{(j)}$ in (15) and $\mathbf{W}_j$ in (16) using SVD decomposition of $\mathbf{H}_j^{(eff)}$
Stage2(Adaptive Compensation):
Keep $\mathbf{F'}_j=\mathbf{F}_1^{(j)}\mathbf{F}_2^{(j)}$, let $\mathbf{W}_{j_{tmp}}=\mathbf{W}_j$
while $k<max\_estimation$ **do**
  for $j=1,2,\dots K$ **do**
    $\mathbf{A'}_r^{(j)}=[\mathbf{a}_r(\theta_1^{(j)}+\Delta\theta^{(j)}),\dots,\mathbf{a}_r(\theta_L^{(j)}+\Delta\theta^{(j)})]$
    $\mathbf{W'}_j=\mathbf{A'}_r^{(j)}(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}\mathbf{W}_{j_{tmp}}$
    $\mathbf{W}_{j_{tmp}}=\mathbf{W'}_j$
    $\mathbf{W'}_j\leftarrow\frac{\mathbf{W'}_j}{\|\mathbf{W'}_j\|}N_s$
  end for
  $k=k+1$
end while

Moreover, we compare the computational complexity of proposed AIA-BD receiver design to SP-AoD-BD algorithm. We consider how many times multiplication in the process of computing the combining matrix for each UE. Note the dimension of matrix $(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}$ is $L\times L$. By using Gaussian elimination, the complexity of calculating the inverse is $\frac{1}{2}L^3+\frac{3}{2}L$. The dimension of matrices $\mathbf{A'}_r^{(j)}$, $\mathbf{A}_r^{(j)^H}$ and $\mathbf{W}_{j_{tmp}}$ are $N_r\times L$, $L\times N_r$ and $N_r\times N_s$ respectively. Therefore, the whole complexity of computing $\mathbf{W'}_j=\mathbf{A'}_r^{(j)}(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}\mathbf{W}_{j_{tmp}}$ in the proposed algorithm can be written as

$$C1=\frac{1}{2}L^3+\frac{3}{2}L^2+2N_r^2L+N_r^2N_s. \tag{31}$$

To perform SP-AoD-BD algorithm, firstly, the effective channel matrix of each UE needs to be recalculated. The computational complexity of matrix multiplication of $\mathbf{H'}_j^{(eff)}=\mathbf{A'}_r^{(j)}\mathbf{\Sigma}^{(j)}\mathbf{A}_t^{(j)^H}\mathbf{F}_1^{(j)}$ is $N_rL^2+2N_rN_tL$. Then, SVD is operated to obtain the optimal combining matrix, and the complexity is given as $N_rL^2+L^2N_r$. Thus, the whole complexity of SP-AoD-BD algorithm can be written as

$$C2=2N_rL^2+2N_rN_tL+L^2N_r. \tag{32}$$

Without loss of generality, it is assumed that $N_r\le N_t$. It is noted that in the mmWave MIMO system, N_t is very large while L and N_s is quite small. It is obvious that the complexity of our proposed algorithm is not affected by the number of antennas at BS.

Furthermore, we have provided the curves of how the complexity scales with L, N_r and N_t in Fig. 2. It shows that the computational complexity of our proposed algorithm is much lower than SP-AoD-BD joint transceiver design algorithm in most cases.

B. Performance Analysis

The main idea of Algorithm 1 is to keep the transmitter the same as that before the rotation while updating the receiver with attitude information. This is a sub-optimal scheme because a joint transceiver design is the optimal one. So we evaluate the feasibility and rationality of our proposed algorithm in this section. We will prove that the sum spectral efficiency achieved by Algorithm 1 can reach R in equation (19).

The performance loss between Algorithm 1 and the sum spectral efficiency in (19) comes from two parts. The first part is the gap between R' and R which is caused by the rotation. Note R' and R are both achieved through joint transceiver designs. Denote the loss between them as ΔR_1. The second part is caused by the method of keeping the precoding matrix unchanged during the rotation. We denote the loss between $\sum_{j=1}^{K}R'_j(\mathbf{W'}_j^*)|_{\mathbf{F'}_j=\mathbf{F}_j}$ and R' as ΔR_2.

We first focus on ΔR_1. As the AoDs at BS are immune to the rotation, the first precoding matrix is always the same. We only have to focus on the procedure of maximizing the spectral efficiency for each UE which depends on SVD decomposition of the effective channel matrix $\mathbf{H'}_j^{(eff)}$. The following proposition shows that the achievable sum spectral efficiency is the same as that before the rotation for the j-th UE when N_r is very large.

Proposition 2 (ΔR_1 approaches zero when N_r is large): When the number of antennas at UE approaches infinity, the achievable sum spectral efficiency approaches that before rotation. It can be described as

$$\lim_{N_r\to\infty}\Delta R_1=\lim_{N_r\to\infty}(R-R')=0. \tag{33}$$

Proof: This is given in the appendix B ■

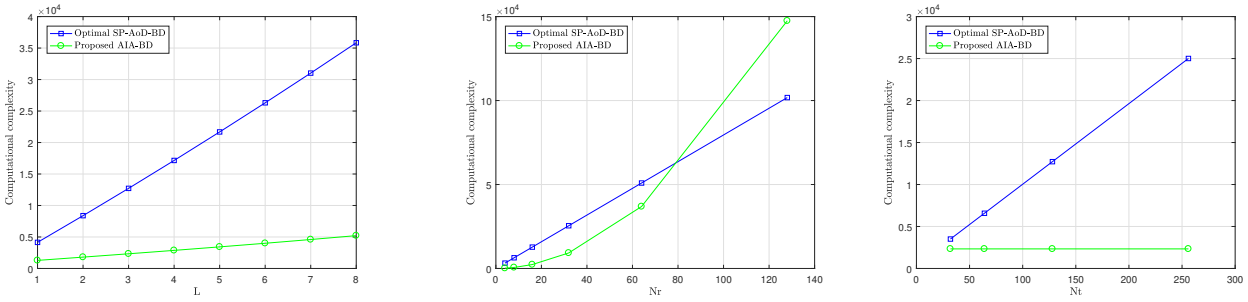
From Propostion 2, we have proved that the achievable sum spectral efficiency are almost the same with and without rotation. Furthermore, we evaluate the conclusion with simulation results in Fig 3. It can be seen that ΔR_1 is very small with different numbers of antennas in mmWave system. To explore the accurate performance gap ΔR_2 between AIA-BD and optimal joint transceiver design after the rotation, we have the following proposition

Proposition 3 (Upper bound of performance loss ΔR_2):

$$\begin{aligned}\Delta R_2&=\sum_{j=1}^{K}R'_j-\sum_{j=1}^{K}R'_j\left(\mathbf{W'}_j^*\right)\Big|_{\mathbf{F'}_j=\mathbf{F}_j}\\&\le\sum_{j=1}^{K}(L-N_s)\log_2\left(1+\frac{\rho}{\sigma_j^2}\left\|\mathbf{\Lambda}_b^{(j)}\right\|_2^2\|\mathbf{C}_j\|_2^2\right).\end{aligned} \tag{34}$$

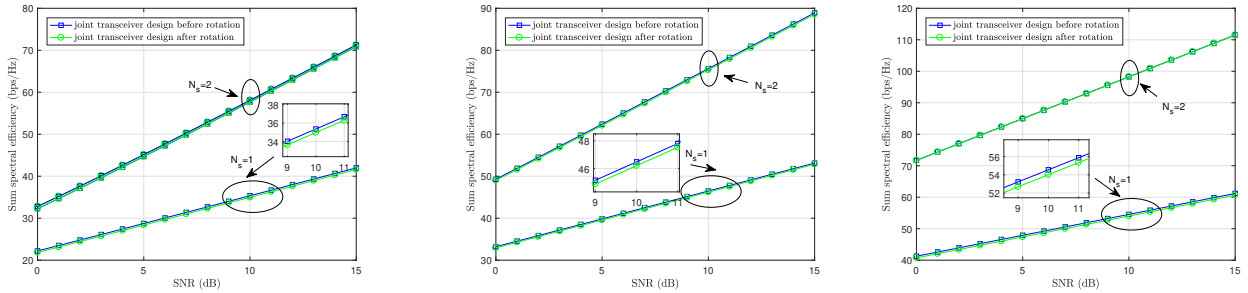
Proof: This is given in the appendix C ■

Remark 4 (Scenarios without performance loss): 1) In the case of only one user, the first precoding matrix of BD no longer exists. The whole procedure will be the same as the SVD algorithm [23]. Since SVD exploits the freedom of wireless channels sufficiently, the number of data stream N_s can reach the rank of channel matrix. In addition, the SVD algorithm is inclined to use larger singular values which means the element of $\mathbf{\Lambda}_b^{(j)}$ is rather small. Thus, the performance loss



(a) L ($N_t = 128, N_r = 16$) (b) N_r ($N_t = 128, L = 3$) (c) N_t ($N_r = 16, L = 3$)

Fig. 2. The computational complexity of our proposed AIA-BD algorithm is much lower than SP-AoD-BD joint transceiver design algorithm in most cases.



(a) N_t = 64, N_r = 8 (b) N_t = 128, N_r = 16 (c) N_t = 128, N_r = 32

Fig. 3. The gap ΔR_1 with different numbers of the BS and UE antennas. It shows the gap between between R' and R is generally negligible.

is negligible. 2) When the number of data streams of each UE is equal to the rank of channel matrix, the performance loss will be zero ($L = N_s$)

Since the proposed algorithm can reach the optimal sum spectral efficiency R' after the rotation, AIA-BD design can reach R in mmWave massive MIMO systems from proposition 2.

IV. Robust AIA-BD Receiver Design

Note that, in the practical system, the UEs may not have the accurate attitude information from motion sensors. This means UEs do not have perfect knowledge of the AoA variations caused by the rotation. Thus, the compensation method at receiver discussed in Section III may not achieve a good performance. So we reformulate the problem to provide an efficient solution that can handle this case.

A. Error Model

Assume that the AoDs at BS and path loss are perfect. The measurement error occurs when AoA variations are estimated [38]. Let $\Delta\theta^{(j)}$ denote the actual AoA variation and $\psi^{(j)}$ denote the observed AoA variation of the j-th UE. Thus we have

$$\Delta\theta^{(j)} = \psi^{(j)} - n_\theta^{(j)}, \tag{35}$$

where $n_\theta^{(j)}$ represents the measurement error and follows a Gaussian distribution $n_\theta^{(j)} \sim N(0, \sigma_{\theta_j}^2)$. The observed AoA variations and the estimated channel matrix from $\psi^{(j)}$ after rotation are given as

$$\mathbf{A'}_{r_{ob}}^{(j)} = [\mathbf{a}_r(\theta_1^{(j)} + \psi^{(j)} - n_\theta^{(j)}), \ldots, \mathbf{a}_r(\theta_L^{(j)} + \psi^{(j)} - n_\theta^{(j)})], \tag{36}$$

$$\mathbf{H'}_{j_{ob}} = \mathbf{A'}_{r_{ob}}^{(j)} \mathbf{\Sigma}^{(j)} {\mathbf{A}_t^{(j)}}^H = \mathbf{C}_{j_{ob}} \mathbf{H}_j, \tag{37}$$

where $\mathbf{C}_{j_{ob}} = \mathbf{A'}_{r_{ob}}^{(j)} ({\mathbf{A}_r^{(j)}}^H \mathbf{A}_r^{(j)})^{-1} {\mathbf{A}_r^{(j)}}^H$. Since the precoding matrix in the proposed AIA algorithm is not related to AoA, we can still keep the transmitter unchanged as in Section III. The difference is that we need to consider the statistical property of spectral efficiency for each UE because of the random variable $n_\theta^{(j)}$.

B. Robust Design of Receiver

In the rotation compensation phase, maximizing the sum spectral efficiency with measurement error can be approximated to solve the following receiver design problem.

Proposition 4 (Approximate problem of maximizing the sum spectral efficiency with measurement error): The robust receiver can be obtained by solving the following problem

$$\begin{aligned} \mathbf{P2}: &\min_{\mathbf{W'}_j} \mathrm{E}\left[\left\| \mathbf{W'}_j - \mathbf{C}_{j_{ob}} \mathbf{U}_a \right\|_F^2 \right], 1 \le j \le K \\ &\mathbf{s.t.} \mathbf{W'}_j^H \mathbf{W'}_j = \mathbf{I}_{Ns}. \end{aligned} \tag{38}$$

Proof: The approximate problem can be proved similar to Appendix A. ■

The closed-form solution can be obtained by using the Taylor expansion. From (30), we know the format of solution

which includes the random variable $n_\theta^{(j)}$ could be expressed as

$$\mathbf{W'}_j^* = \alpha \mathrm{E}\left[\mathbf{C}_{j_{ob}}\right]\mathbf{U}_a. \tag{39}$$

By using the definition of $\mathbf{C}_{j_{ob}}$ in (37), $\mathrm{E}\left[\mathbf{C}_{j_{ob}}\right]\mathbf{U}_a$ can be further simplified as

$$\begin{aligned}\mathrm{E}\left[\mathbf{C}_{j_{ob}}\right]\mathbf{U}_a &= \mathrm{E}\left[\mathbf{A'}_{r_{ob}}^{(j)}(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}\right]\mathbf{U}_a \\ &= \mathrm{E}\left[\mathbf{A'}_{r_{ob}}^{(j)}\right](\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}\mathbf{U}_a.\end{aligned} \tag{40}$$

The expectation of element in matrix $\mathbf{A'}_{r_{ob}}^{(j)}$ is

$$\mathrm{E}\left[\mathbf{A'}_{r_{ob}}^{(j)}(s,p)\right] = \mathrm{E}\left[\frac{1}{\sqrt{N_r}}e^{j\frac{2\pi}{\lambda}d(s-1)\cos\left(\theta_p^{(j)}+\psi^{(j)}-n_\theta^{(j)}\right)}\right], \tag{41}$$

and using the second order Taylor expansion to approximate $\mathbf{A'}_{r_{ob}}^{(j)}(s,p)$ when $n_\theta^{(j)}$ is small, one has

$$\mathbf{A'}_{r_{ob}}^{(j)}(s,p) = \mathbf{A'}_{r_0}^{(j)}(s,p) + n_\theta^{(j)}\mathbf{A'}_{r_1}^{(j)}(s,p) + {n_\theta^{(j)}}^2\mathbf{A'}_{r_2}^{(j)}(s,p), \tag{42}$$

where

$$\mathbf{A'}_{r_0}^{(j)}(s,p) = \frac{1}{\sqrt{N_r}}e^{j\frac{2\pi}{\lambda}d(s-1)\cos\left(\theta_p^{(j)}+\psi^{(j)}\right)}, \tag{43}$$

$$\begin{aligned}\mathbf{A'}_{r_1}^{(j)}(s,p) =& \frac{1}{\sqrt{N_r}}j\frac{2\pi}{\lambda}d(s-1)\sin\left(\theta_p^{(j)}+\psi^{(j)}\right) \\ &\times e^{j\frac{2\pi}{\lambda}d(s-1)\cos\left(\theta_p^{(j)}+\psi^{(j)}\right)},\end{aligned} \tag{44}$$

$$\begin{aligned}\mathbf{A'}_{r_2}^{(j)}(s,p) =& -\frac{1}{\sqrt{N_r}}\frac{1}{2}\left(\frac{4\pi^2}{\lambda^2}d^2(s-1)^2\sin^2\left(\theta_p^{(j)}+\psi^{(j)}\right)\right. \\ &\left. + j\frac{2\pi}{\lambda}d(s-1)\cos\left(\theta_p^{(j)}+\psi^{(j)}\right)\right) \\ &\times e^{j\frac{2\pi}{\lambda}d(s-1)\cos\left(\theta_p^{(j)}+\psi^{(j)}\right)}.\end{aligned} \tag{45}$$

By using (39), the final solution can be obtained as

$$\begin{aligned}\mathbf{W'}_j^* =& \alpha\mathrm{E}\left[\mathbf{A'}_{r_0}^{(j)}(s,p) + n_\theta^{(j)}\mathbf{A'}_{r_1}^{(j)}(s,p) + {n_\theta^{(j)}}^2\mathbf{A'}_{r_2}^{(j)}(s,p)\right] \\ &\times(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}\mathbf{U}_a \\ =& \alpha\mathbf{A'}_{r_0}^{(j)}(s,p)(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}\mathbf{U}_a \\ &+ \alpha\sigma_{\theta_j}^2\mathbf{A'}_{r_2}^{(j)}(s,p)(\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)})^{-1}\mathbf{A}_r^{(j)^H}\mathbf{U}_a \\ =& \alpha\left(\mathbf{C}_{j_0}+\sigma_{\theta_j}^2\mathbf{C}_{j_2}\right)\mathbf{U}_a.\end{aligned} \tag{46}$$

where $\mathbf{C}_{j_0}$ and $\mathbf{C}_{j_2}$ are the matrices which only depend on the observed AoA variation $\psi^{(j)}$. In (46), the first part $\alpha\mathbf{C}_{j_0}\mathbf{U}_a$ is actually the solution that calculated over the observed values as in (30).

It is noted that, in the derivation of (46), we utilize the Taylor expansion. It will lead to approximation errors. The dominant one between measurement error and approximation error will determine the final receiver. Let $\sigma_{threshold}^2$ denote the variance when the measurement error and approximation error have the same effects on the spectral efficiency. We have the following rules: when $\sigma_{\theta_j}^2 \leq \sigma_{threshold}^2$, we design the receiver with $\mathbf{W'}_j^* = \alpha\mathbf{C}_{j_0}\mathbf{U}_a$. Otherwise, we adopt $\mathbf{W'}_j^* = \alpha\left(\mathbf{C}_{j_0}+\sigma_{\theta_j}^2\mathbf{C}_{j_2}\right)\mathbf{U}_a$ to account for the error caused by the sensors.

TABLE I
MAIN PARAMETERS

Parameters	Values
Carrier frequency	28GHz
Number of clusters	3
Number of rays (Each cluster)	1
Max data streams (Each UE)	2
Antenna array structure	ULA
Antenna array element space	0.5λ
Angular spreads at BS	7°
Angular spreads at UE	7°

V. SIMULATION RESULTS AND DISCUSSION

In this section, we present the simulation results to demonstrate the performances of the proposed AIA algorithms. We assume the geometric channel model as discussed in Section II with $L_j = 3$ clusters. Each cluster consists of only one ray. The clusters are set as uniform distribution. In addition, it is assumed that $\alpha_i^{(j)} \sim \mathcal{CN}\left(0, \sigma_{\alpha,i}^2\right)$, where $\sigma_{\alpha,i}^2$ represents the average power of the i-th scattering cluster. The cluster powers are such that $\sum_{i=1}^{L_j}\sigma_{\alpha,i}^2 = \sqrt{N_t N_r / L_j}$. The angular spreads at both ends are equal and set as $\sigma_\theta^2 = \sigma_\phi^2 = 7^\circ$. In all our simulations, both BS and UEs are assumed to have their antennas arranged in ULA with element space $d = \lambda/2$. Furthermore, the channel estimation period is set as 500 symbols. Consider N_t symbols used as pilots without compressed sensing, then $500 - N_t$ symbols can be used for data transmission. The main parameters have also been shown in Table 1. The curves are obtained by averaging over one thousand channel realizations at each SNR point.

We first evaluate the performance of the proposed AIA-BD receiver design algorithm. Consider a multi-user mmWave system, where the BS has $N_t = 64$ and each UE has $N_r = 8$. For simplicity, we assume that each UE rotates around the y-axis from 0 degree to 15 degrees and receives at most two data streams. Fig. 4 reveals that the sum spectral efficiency will degrade significantly without compensation. In addition, we demonstrate the performance of AIA-BD algorithm by using SP-AoD-BD algorithm with CSI after the rotation for comparison. It is shown that the proposed algorithm achieves almost the same performance as SP-AoD-BD even with a relatively large angle. In practice, the rotation angle is very small between two adjacent compensations. This is because the duration of two adjacent compensations in mmWave systems is very short, which is generally milliseconds. The rapidest rotation of UE is hundreds degrees per second in practical scenarios. Therefore, the rotation angle is always small and the performance loss caused by keeping the transmitter unchanged can be ignored.

Now Assume that there are four UEs in a mmWave system. They rotate 30°, 10°, 10° and 18°, respectively, during one rotation. The BS has $N_t = 128$ and each UE has $N_r = 16$. In Fig. 5, the performance of the proposed AIA-BD receiver design is very close to the SP-AoD-BD algorithm. Furthermore, we also compare the proposed design with the AIA BeamSteering method [22], [39]. In the AIA BeamSteering algorithm, the antennas at transceivers are aligned to the direction with the largest gain after rotation and BD is ex-

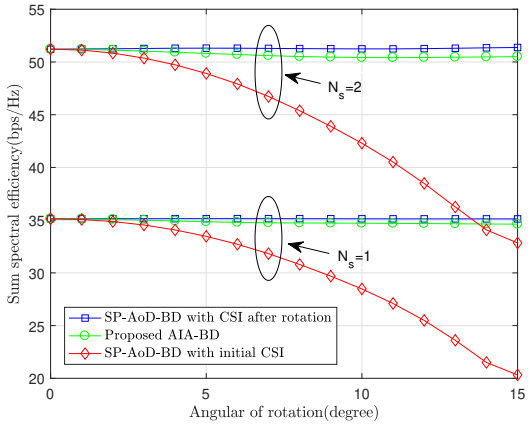


Fig. 4. The sum spectral efficiency versus the angle of rotation for different transmission schemes. Each UE works in a 64×8 mm-wave system and rotates around y-axis simultaneously. The SNR is set as 10 dB.

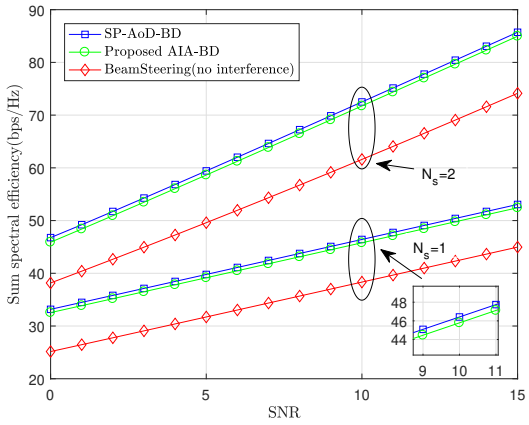


Fig. 5. The performances of various transceiver designs with each UE working in a 128×16 mm-wave system. The proposed AIA-BD algorithm performs better than AIA Beamsteering method, and achieves the sum spectral efficiency similar to SP-AoD-BD.

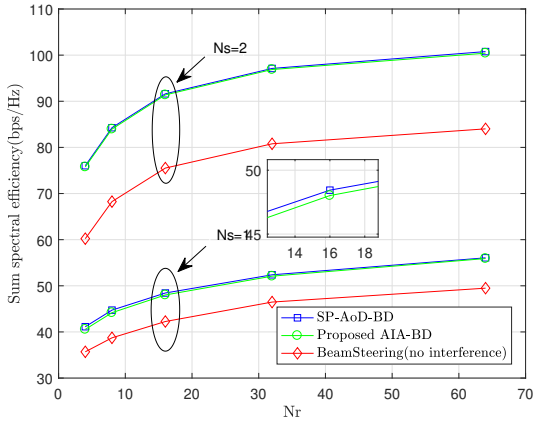
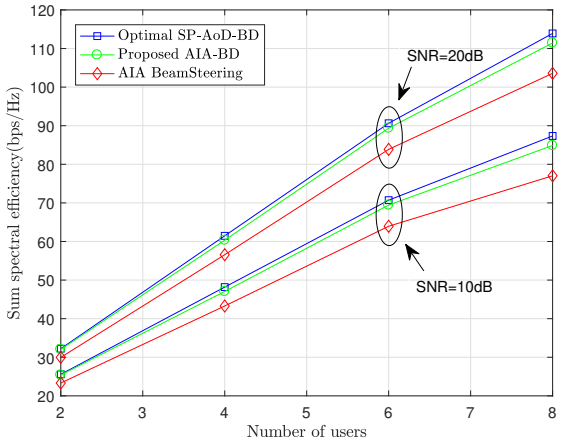
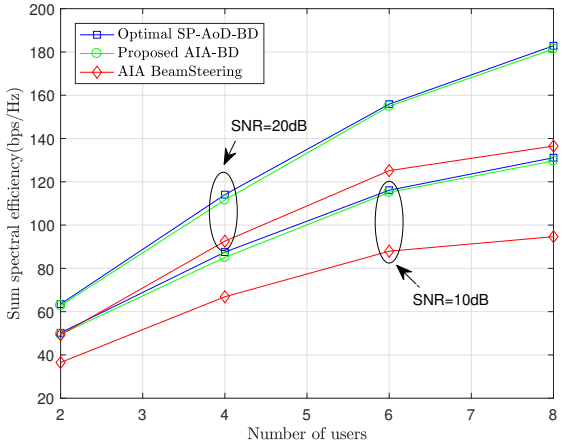


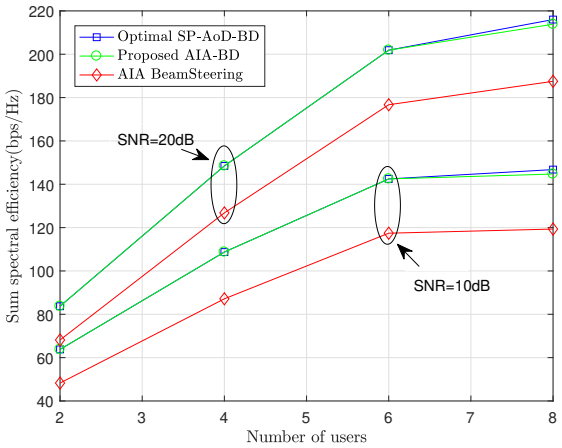
Fig. 6. The sum spectral efficiency achieved versus diffident number of antennas at UE. The SNR is set as 10dB and the BS is applied with 128 antennas. Other simulation parameters are set as in Fig. 4. This result evaluates that the proposed AIA-BD algorithm can approximate the SP-AoD-BD with various scale antenna array.



(a) N_s = 1



(b) N_s = 2



(c) N_s = 3

Fig. 7. Comparison of the sum spectral efficiency as a function of the number of users K at SNR values of 10dB and 20dB. In a multi-user scenario, BS has $N_t = 128$ and each UE is equipped with 16 antennas. Plots (a)-(c) correspond to the scenarios where the number of data streams of each UE range from 1 to 3.

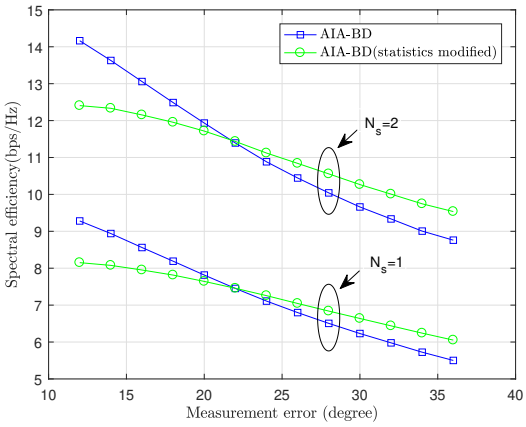


Fig. 8. The spectral efficiency achieved versus different measurement errors with statistics modified AIA-BD design and initial AIA-BD design. It is evident that the robust receiver with modification will perform better even with a relative large measurement error.

ploited to make sure that the precoding matrix of one UE is orthogonal to other UEs' channel. It is seen from Fig. 5 that the proposed AIA-BD receiver design attains a significantly better performance than the AIA BeamSteering method with practical antenna arrays. A gain of about 8 bps/Hz is observed when $N_s = 1$. With the increase of data streams, the sum spectral efficiency obtains a gain of 10 bps/Hz.

To show the performances of mmWave systems for different antenna arrays, we adjust the number of antennas at UE. Consider BS has $N_t = 128$ and transmits at most two data streams to each UE. Other settings are the same as Fig 4. Fig. 6 shows that the sum spectral efficiency achieved by AIA-BD algorithm has a great improvement over the AIA BeamSteering method. Meanwhile, when the number of antennas at UE becomes large, the increase of spectral efficiency is very small. This is because the channel training overhead has significantly increased.

Fig. 7 shows the sum spectral efficiency as a function of the number of users at SNR values of 10dB and 20dB. The BS has $N_t = 128$ and each UE has $N_r = 16$. Other settings are the same as Fig.4. One sees that the proposed AIA-BD algorithm can be well applied in multi-user scenarios. With the number of data streams increasing, the performance gap between the proposed design and joint transceiver design gradually approaches zero.

Fig. 8 illustrates the performance of proposed AIA-BD with robust design which utilizes the statistical information of measurement error σ_{θ_j}. We take UE1 as an example, as other UEs have the same results. Considering UE1 working in a 64×16 mmWave system. Other settings are the same as Fig. 4. The measurement error variance data can be obtained from the performance indicators of gyroscopes, such as fixed angle deviation stability and random angle deviation stability. We can find that with the measurement error increasing, the AIA-BD robust receiver will gradually give a better performance.

The above simulation results have showed that the proposed AIA-BD receiver algorithm can efficiently compensate the performance loss caused by UE's rotation and can approach the sum spectral efficiency obtained by joint transceiver design. The robust AIA-BD receiver could also achieve a better performance even with a relative large measurement error. Moreover, considering the implementation of the proposed design in practical mmWave systems, the number of user antennas is limited. Based on the complexity analysis and simulation results, each UE should be equipped with 8 or 16 antennas when BS has $N_t = 64$ or $N_t = 128$. To achieve higher sum spectral efficiency, the number of data streams for each UE should be as close as possible to L_j.

VI. Conclusion

In this article, we have proposed a low-complexity AIA-BD receiver design algorithm for multi-user mmWave communications. In our design, we have first decoupled UEs based on SP-AoD-BD. Then, the performance degradation caused by UEs' rotations is compensated only at the receivers with the attitude information from motion sensors. Compared with the joint transceiver design, we have theoretically proved that the proposed AIA-BD receiver design algorithm is capable of achieving the near-optimal performance. Moreover, in AIA-BD algorithm, no feedback is required between BS and UEs during the rotation to substantially reduce the training overhead. Simulation results have demonstrated the proposed algorithm could efficiently avoid the performance degradation and approach the performance obtained by joint transceiver design. Future work in this area includes the application for more complex scenarios. Attitude change and movement of UEs should be simultaneously considered in the transceiver design algorithm. Additionally, the angular velocity of rotation could be utilized for accurate beam management.

Appendix A
Proof of Proposition 1

Since we have decoupled the performance corresponding to different UE with AoD-BD method. The problem of maximizing the sum spectral efficiency is equivalent to maximize the spectral efficiency for each UE.

We take the j-th UE as an example in the following derivation. Note that its spectral efficiency in (23) becomes R_j in (19) when $\mathbf{W'}_j$ satisfies $\mathbf{C}_j^H\mathbf{W'}_j = \mathbf{U}_a$. Here we make the assumption that the mmWave system parameters are such that $\mathbf{C}_j^H\mathbf{W'}$ can be made sufficiently "close" to $\mathbf{U}_a$. This assumption is equivalent to that $R'_j(\mathbf{W'}_j)$ can generally reach R_j and we have proved it in proposition 2 and proposition 3. On the other hand, such assumption is similar to the high-resolution approximation used in codebooks and can be generally realized in mmWave systems [40]. Mathematically, this "closeness" can be defined by following two equivalent approximations.

1) The eigenvalues of the matrix $\mathbf{I}_{Ns} - \mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a$ are small. In the case of mmWave system, this can be equivalently stated as $\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j \approx \mathbf{I}_{Ns}$.
2) The singular values of $\mathbf{U}_b^H\mathbf{C}_j^H\mathbf{W'}_j$ are small or $\mathbf{U}_b^H\mathbf{C}_j^H\mathbf{W'}_j \approx \mathbf{0}$.

The above approximations allow us to prove that the spectral efficiency in (23) is very close to the objective function in (24).

$$\begin{aligned}
&R'_j(\mathbf{W'}_j)\\
&=\log_2\left(\left|\mathbf{I}_{N_s}+\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a\right|\right)\\
&\overset{(a)}{=}R_j+\log_2\left(\left|\mathbf{I}_{N_s}-\left(\mathbf{I}_{N_s}+\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\right)^{-1}\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\right.\right.\\
&\quad\left.\left.\times\left(\mathbf{I}_{N_s}-\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a\right)\right|\right)\\
&\overset{(b)}{\approx}R_j-tr\left(\left|\left(\mathbf{I}_{N_s}+\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\right)^{-1}\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\right.\right.\\
&\quad\left.\left.\times\left(\mathbf{I}_{N_s}-\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a\right)\right|\right)\\
&\overset{(c)}{\approx}R_j-tr\left(\mathbf{I}_{N_s}-\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a\right)\\
&=R_j-\left(N_s-\left\|\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\right\|_F^2\right)\\
&\overset{(d)}{\approx}R_j-\left(\left\|\mathbf{W}_j-\mathbf{C}_j\mathbf{U}_a\right\|_F^2\right),
\end{aligned}\tag{47}$$

where (a) is obtained by defining the matrices $\mathbf{A}=\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a$, $\mathbf{B}=\frac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2$ and noting that $(\mathbf{I}+\mathbf{BA})=(\mathbf{I}+\mathbf{B})\left(\mathbf{I}-(\mathbf{I}+\mathbf{B})^{-1}\mathbf{B}(\mathbf{I}-\mathbf{A})\right)$. The simplification in (b) follows from the approximation that the eigenvalues of the matrix $\mathbf{X}=\left(\mathbf{I}_{N_s}+\frac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\right)^{-1}\frac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^2\left(\mathbf{I}_{N_s}-\mathbf{U}_a^H\mathbf{C}_j^H\mathbf{W'}_j\mathbf{W'}_j^H\mathbf{C}_j\mathbf{U}_a\right)$ are small and use the following approximation $\log_2\left(|\mathbf{I}_{N_s}-\mathbf{X}|\right)\approx\log_2\left(1-tr\left(\mathbf{X}\right)\right)\approx-tr\left(\mathbf{X}\right)$. (c) follows from high SNR approximation and finally we use Euclidean distance to replace chordal distance in (d) because of the manifold's locally Euclidean property [41].

Thus, maximizing the sum spectral efficiency in (23) can be approximated to minimizing $\left\|\mathbf{W}_j-\mathbf{C}_j\mathbf{U}_a\right\|_F^2$ for each UE. Therefore, we can write the approximate problem as in (24).

Appendix B
Proof of proposition 2

Define the ordered SVD of j-th UE's effective channel matrix after rotation as

$$\mathbf{H'}_j^{(eff)}=\mathbf{A'}_r^{(j)}\mathbf{\Sigma}^{(j)}\mathbf{A}_t^{(j)^H}\mathbf{F}_1^{(j)}=\mathbf{U'}_2^{(j)}\mathbf{\Lambda'}_2^{(j)}\mathbf{V'}_2^{(j)^H},\tag{48}$$

where $\mathbf{U'}_2^{(j)}\in\mathbb{C}^{N_r\times L}$, $\mathbf{V'}_2^{(j)}\in\mathbb{C}^{N_r\times L}$ are unitary matrices, $\mathbf{\Lambda'}_2^{(j)}\in\mathbb{C}^{N_r\times L}$ consists of all the singular values. Furthermore, we can define the following partitions

$$\begin{aligned}
\mathbf{U'}_2^{(j)}&=\left[\mathbf{U'}_a^{(j)}\mathbf{U'}_b^{(j)}\right],\\
\mathbf{\Lambda'}_2^{(j)}&=\begin{bmatrix}\mathbf{\Lambda'}_a^{(j)} & \\ & \mathbf{\Lambda'}_b^{(j)}\end{bmatrix},\\
\mathbf{V'}_2^{(j)}&=\left[\mathbf{V'}_a^{(j)}\mathbf{V'}_b^{(j)}\right].
\end{aligned}\tag{49}$$

where $\mathbf{U'}_a^{(j)}\in\mathbb{C}^{N_r\times N_s}$, $\mathbf{V'}_a^{(j)}\in\mathbb{C}^{L\times N_s}$, $\mathbf{\Lambda'}_a^{(j)}\in\mathbb{C}^{N_s\times N_s}$. Let us focus on the matrix $\mathbf{H'}_j^{(eff)^H}\mathbf{H'}_j^{(eff)}$ and $\mathbf{H}_j^{(eff)^H}\mathbf{H}_j^{(eff)}$

Using (14),(15),(48) and (49), we have

$$\begin{aligned}
&\mathbf{H'}_j^{(eff)^H}\mathbf{H'}_j^{(eff)}\\
&=\mathbf{F}_1^{(j)^H}\mathbf{A}_t^{(j)}\mathbf{\Sigma}^{(j)^H}\mathbf{A'}_r^{(j)^H}\mathbf{A'}_r^{(j)}\mathbf{\Sigma}^{(j)}\mathbf{A}_t^{(j)^H}\mathbf{F}_1^{(j)}\\
&=\mathbf{V'}_2^{(j)}\mathbf{\Lambda'}_2^{(j)^2}\mathbf{V'}_2^{(j)^H},
\end{aligned}\tag{50}$$

$$\begin{aligned}
&\mathbf{H}_j^{(eff)^H}\mathbf{H}_j^{(eff)}\\
&=\mathbf{F}_1^{(j)^H}\mathbf{A}_t^{(j)}\mathbf{\Sigma}^{(j)^H}\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)}\mathbf{\Sigma}^{(j)}\mathbf{A}_t^{(j)^H}\mathbf{F}_1^{(j)}\\
&=\mathbf{V}_2^{(j)}\mathbf{\Lambda}_2^{(j)^2}\mathbf{V}_2^{(j)^H}.
\end{aligned}\tag{51}$$

By using the massive MIMO array in mmWave systems, the beams formulated will be sufficiently narrow, such that the inter-beam interference can be omitted. The array response matrix of ULA in the poor scattering channel is regarded to be orthogonal. Then, we could obtain $\lim_{N_r\to\infty}\mathbf{A'}_r^{(j)^H}\mathbf{A'}_r^{(j)}=\lim_{N_r\to\infty}\mathbf{A}_r^{(j)^H}\mathbf{A}_r^{(j)}=\mathbf{I}_{N_r}$ [42]. Thus we have

$$\begin{aligned}
\mathbf{H'}_j^{(eff)^H}\mathbf{H'}_j^{(eff)}\mathbf{v}_i&=\mathbf{H}_j^{(eff)^H}\mathbf{H}_j^{(eff)}\mathbf{v}_i\\
&=\mathbf{F}_1^{(j)^H}\mathbf{A}_t^{(j)}\mathbf{\Sigma}^{(j)^H}\mathbf{\Sigma}^{(j)}\mathbf{A}_t^{(j)^H}\mathbf{F}_1^{(j)}\mathbf{v}_i\\
&=\lambda_i^2\mathbf{v}_i,
\end{aligned}\tag{52}$$

where λ_i is the ith non-zero singular value of $\mathbf{\Lambda}_2^{(j)}$, $\mathbf{v}_i$ is the ith column of $\mathbf{V}_2^{(j)}$. According to the relationship between eigenvalue decomposition and singular value decomposition, we can write

$$\lim_{N_r\to\infty}\mathbf{\Lambda'}_2^{(j)}=\mathbf{\Lambda}_2^{(j)}.\tag{53}$$

We generalize this result to all the UEs and could obtain

$$\begin{aligned}
\lim_{N_r\to\infty}\Delta R_1&=\lim_{N_r\to\infty}\left(R-R'\right)\\
&=\lim_{N_r\to\infty}\left(\sum_{j=1}^{K}\log_2\left(\left|\mathbf{I}+\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda}_a^{(j)^2}\right|\right)\right)\\
&\quad-\lim_{N_r\to\infty}\left(\sum_{j=1}^{K}\log_2\left(\left|\mathbf{I}+\tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda'}_a^{(j)^2}\right|\right)\right)\\
&=0.
\end{aligned}$$

Appendix C
Proof of proposition 3

Note $\sum_{j=1}^{K}R'_j\left(\mathbf{W'}_j^*\right)\Big|_{\mathbf{F'}_j=\mathbf{F}_j}$ is the sum spectral efficiency attained by keeping transmitter unchanged. Here we consider $\mathbf{W'}_j=\mathbf{U'}_a^{(j)}$ in $\sum_{j=1}^{K}R'_j\left(\mathbf{W'}_j^*\right)\Big|_{\mathbf{F'}_j=\mathbf{F}_j}$ which is generally not the optimal solution when the attitude has changed, because the optimal one should have the same column space as the receiver in (24). Thus, we have

$$\sum_{j=1}^{K}R_j\left(\mathbf{W'}_j\right)\Big|_{\mathbf{W'}_j=\mathbf{U'}_a^{(j)},\mathbf{F'}_j=\mathbf{F}_j}\leq\sum_{j=1}^{K}R_j\left(\mathbf{W'}_j^*\right)\Big|_{\mathbf{F'}_j=\mathbf{F}_j},$$

and the performance loss satisfies the following inequality

$$\Delta R_2 \leq \sum_{j=1}^{K} R'_j - \sum_{j=1}^{K} R'_j\left(\mathbf{W'}_j\right)\Big|_{\mathbf{W'}_j=\mathbf{U'}_a^{(j)},\mathbf{F'}_j=\mathbf{F}_j}. \quad (54)$$

We extend the second term on the right side in (54)

$$\begin{aligned}
&\sum_{j=1}^{K} R'_j\left(\mathbf{W'}_j\right)\Big|_{\mathbf{W'}_j=\mathbf{U'}_a^{(j)},\mathbf{F'}_j=\mathbf{F}_j} \\
&= \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{U'}_a^{(j)^H}\mathbf{H'}_j\mathbf{F}_j\mathbf{F}_j{}^H\mathbf{H'}_j^H\mathbf{U'}_a^{(j)}\right|\right) \\
&= \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{U'}_a^{(j)^H}\mathbf{H'}_j^{(eff)}\mathbf{V}_a^{(j)}\right.\right. \\
&\quad \left.\left.\times\, \mathbf{V}_a^{(j)^H}\mathbf{H'}_j^{(eff)^H}\mathbf{U'}_a^{(j)}\right|\right) \\
&= \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{V}_a^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_a^{(j)}\right|\right).
\end{aligned} \quad (55)$$

We first use the partitions defined in (15) and define the partition of the matrix $\mathbf{V}_2^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_2^{(j)}$

$$\mathbf{V}_2^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_2^{(j)} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}, \quad (56)$$

where

$$\begin{aligned}
\mathbf{P}_{11} &= \mathbf{V}_a^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_a^{(j)}, \\
\mathbf{P}_{12} &= \mathbf{V}_a^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_b^{(j)}, \\
\mathbf{P}_{21} &= \mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_a^{(j)}, \\
\mathbf{P}_{22} &= \mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_b^{(j)}.
\end{aligned}$$

From the Schur complement identity for matrix determinants, the sum spectral efficiency in (55) can be further simplified

$$\begin{aligned}
&\sum_{j=1}^{K} R'_j\left(\mathbf{W'}_j\right) \\
&= \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_L + \tfrac{\rho}{\sigma_j^2}\mathbf{V}_2^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_2^{(j)}\right|\right) \\
&\quad - \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_{L-N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{P}_{22} - \left(\tfrac{\rho}{\sigma_j^2}\right)^2\mathbf{P}_{21}\right.\right. \\
&\quad \left.\left.\times\, \left(\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{P}_{11}\right)^{-1}\mathbf{P}_{12}\right|\right).
\end{aligned} \quad (57)$$

Notice the first term in (57)

$$\begin{aligned}
\mathbf{S}_1 &= \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_L + \tfrac{\rho}{\sigma_j^2}\mathbf{V}_2^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_2^{(j)}\right|\right) \\
&= \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_L + \tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda'}_a^{(j)^2}\mathbf{V'}_a^{(j)^H}\mathbf{V}_2^{(j)}\mathbf{V}_2^{(j)^H}\mathbf{V'}_a^{(j)}\right|\right) \\
&\overset{(a)}{=} \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{\Lambda'}_a^{(j)^2}\right|\right)
\end{aligned} \quad (58)$$

is exactly $\sum_{j=1}^{K} R'_j$, where (a) follows that $\mathbf{V}_2^{(j)}$, $\mathbf{V'}_2^{(j)}$ are unitary matrices.

Therefore, we only have to simplify the second term in (57)

$$\begin{aligned}
\mathbf{S}_2 &\overset{(a)}{=} \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_{L-N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)}\right.\right. \\
&\quad \times \left(\mathbf{I}_L - \tfrac{\rho}{\sigma_j^2}\mathbf{X}\left(\mathbf{I}_{N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\right) \\
&\quad \left.\left.\times\, \mathbf{\Lambda'}_a^{(j)}\mathbf{V'}_a^{(j)^H}\mathbf{V}_b^{(j)}\right|\right) \\
&\overset{(b)}{=} \sum_{j=1}^{K} \log_2\left(\left|\mathbf{I}_{L-N_s} + \tfrac{\rho}{\sigma_j^2}\mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)}\right.\right. \\
&\quad \left.\left.\times\, \left(\mathbf{I}_L + \tfrac{\rho}{\sigma_j^2}\mathbf{X}\mathbf{X}^H\right)^{-1}\mathbf{\Lambda'}_a^{(j)}\mathbf{V'}_a^{(j)^H}\mathbf{V}_b^{(j)}\right|\right) \\
&\overset{(c)}{=} \sum_{j=1}^{K}\sum_{i=1}^{L-N_s} \log_2\left(1 + \lambda_i\left\{\tfrac{\rho}{\sigma_j^2}\mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)}\right.\right. \\
&\quad \left.\left.\times\, \left(\mathbf{I}_L + \tfrac{\rho}{\sigma_j^2}XX^H\right)^{-1}\mathbf{\Lambda'}_a^{(j)}\mathbf{V'}_a^{(j)^H}\mathbf{V}_b^{(j)}\right\}\right) \\
&\overset{(d)}{\leq} \sum_{j=1}^{K} (L-N_s)\log_2\left(1 + \left\|\tfrac{\rho}{\sigma_j^2}\mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)}\right.\right. \\
&\quad \left.\left.\times\, \left(\mathbf{I}_L + \tfrac{\rho}{\sigma_j^2}XX^H\right)^{-1}\mathbf{\Lambda'}_a^{(j)}\mathbf{V'}_a^{(j)^H}\mathbf{V}_b^{(j)}\right\|_2\right) \\
&\overset{(e)}{\leq} \sum_{j=1}^{K} (L-N_s)\log_2\left(1 + \tfrac{\rho}{\sigma_j^2}\left\|\mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)}\right\|_2^2\right) \\
&\overset{(f)}{=} \sum_{j=1}^{K} (L-N_s)\log_2\left(1 + \tfrac{\rho}{\sigma_j^2}\left\|\mathbf{\Lambda}_b^{(j)}\mathbf{U}_b^{(j)^H}\mathbf{C}_j^H\mathbf{U'}_a^{(j)}\right\|_2^2\right) \\
&\overset{(g)}{\leq} \sum_{j=1}^{K} (L-N_s)\log_2\left(1 + \tfrac{\rho}{\sigma_j^2}\left\|\mathbf{\Lambda}_b^{(j)}\right\|_2^2\|\mathbf{C}_j\|_2^2\right),
\end{aligned} \quad (59)$$

where (a) is obtained by defining $\mathbf{P}_{11} = \mathbf{X}^H\mathbf{X}$, (b) follows the inverse matrix equation $(\mathbf{I}+\mathbf{BA})^{-1} = \mathbf{I} - \mathbf{B}(\mathbf{I}+\mathbf{AB})^{-1}\mathbf{A}$, let $\mathbf{B} = \frac{\rho}{\sigma_j^2}\mathbf{X}$, $\mathbf{A} = \mathbf{X}^H$, the $\lambda_i\{\mathbf{A}\}$ in (c) represents the ith singular value of matrix $\mathbf{A}$, (d) follows the definition of 2-norm, (e) and (g) follows the submultiplicative property of 2-norm. Finally, (f) follows the SVD of channel matrix $\mathbf{V}_b^{(j)^H}\mathbf{V'}_a^{(j)}\mathbf{\Lambda'}_a^{(j)} = \mathbf{V}_b^{(j)^H}\mathbf{H}_j^{(eff)^H}\mathbf{C}_j^H\mathbf{U'}_a^{(j)} = \mathbf{\Lambda}_b^{(j)}\mathbf{U}_b^{(j)^H}\mathbf{C}_j^H\mathbf{U'}_a^{(j)}$.

Therefore, we have

$$\Delta R_2 \leq \sum_{j=1}^{K} (L-N_s)\log_2\left(1 + \frac{\rho}{\sigma_j^2}\left\|\mathbf{\Lambda}_b^{(j)}\right\|_2^2\|\mathbf{C}_j\|_2^2\right).$$

REFERENCES

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

[2] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6g and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[3] W. Saad, M. Bennis, and M. Chen, "A vision of 6g wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, 2020.

[4] D. Cohen, S. Tsiper, and Y. C. Eldar, "Analog-to-digital cognitive radio: Sampling, detection, and hardware," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 137–166, 2018.

[5] A. Ali, N. Gonzalez-Prelcic, R. W. Heath, and A. Ghosh, "Leveraging sensing at the infrastructure for mmwave communication," *IEEE Commun Mag*, vol. 58, no. 7, pp. 84–89, 2020.

[6] M. Hashemi, C. E. Koksal, and N. B. Shroff, "Out-of-band millimeter wave beamforming and communications to achieve low latency and high energy efficiency in 5g systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 875–888, 2018.

[7] T. Marzetta, "Blast training: Estimating channel characteristics for high capacity space-time wireless," 01 2000.

[8] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[9] A. Akbarpour-Kasgari and M. Ardebilipour, "Massive mimo-ofdm channel estimation via distributed compressed sensing," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 376–379, 2019.

[10] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, 2016.

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

[12] N. Gonzalez-Prelcic, A. Ali, V. Va, and R. W. Heath, "Millimeter-wave communication with out-of-band information," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 140–146, 2017.

[13] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, 2018.

[14] F. Maschietti, D. Gesbert, P. de Kerret, and H. Wymeersch, "Robust location-aided beam alignment in millimeter wave massive mimo," in *IEEE Glob. Commun. Conf., GLOBECOM - Proc.*, 2017, pp. 1–6.

[15] Q. Zhang, X. Zhang, and C. Yang, "Camera sensing assisted fast mmwave beam tracking for connected automated vehicles," *IEEE Internet Things J.*, pp. 1–1, 2022.

[16] G. Lee and R. W. Heath, "Fast link configuration for mmwave multiuser MIMO downlink using spatial aod angular supports," in *IEEE Veh. Technol. Conf.*, 2017, pp. 1–2.

[17] N. Garcia, H. Wymeersch, E. G. Ström, and D. Slock, "Location-aided mm-wave channel estimation for vehicular communication," in *IEEE Workshop Signal Process. Adv. Wireless Commun. SPAWC*, 2016, pp. 1–5.

[18] A. Ali, N. González-Prelcic, and R. W. Heath, "Estimating millimeter wave channels using out-of-band measurements," in *Inf. Theory Appl. Workshop, ITA*, 2016, pp. 1–6.

[19] W. Kim, J. Bae, and S.-J. Lee, "Gain degradation effect due to beam misalignment on mmwave beamforming for 5g cellular communication," in *Int. Conf. ICT Converg.: Innov. Towar. IoT, 5G, Smart Media Era, ICTC*, 2015, pp. 1252–1256.

[20] J. Zhang, W. Xu, H. Gao, M. Pan, Z. Feng, and Z. Han, "Position-attitude prediction based beam tracking for uav mmwave communications," in *IEEE Int. Conf. Commun.*, 2019, pp. 1–7.

[21] J. Bao, D. Sun, and H. Li, "Motion sensor aided beam tracking in mobile devices of millimeter-wave communications," in *IEEE Int. Conf. Commun.*, 2018, pp. 1–7.

[22] D.-S. Shim, C.-K. Yang, J. H. Kim, J. P. Han, and Y. S. Cho, "Application of motion sensors for beam-tracking of mobile stations in mmwave communication systems," *Ah S Sens*, vol. 14, no. 10, pp. 19 622–19 638, 2014. [Online]. Available: https://www.mdpi.com/1424-8220/14/10/19622

[23] S. Zhou, L. Chen, and W. Wang, "Attitude information aided digital beamforming in millimeter-wave mimo systems," *IEEE Syst. J.*, pp. 1–12, 2021.

[24] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.

[25] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J Sel Topics Signal Process*, vol. 8, no. 5, pp. 831–846, 2014.

[26] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of mimo channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, 2003.

[27] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser mimo systems," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201–211, 2016.

[28] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, 2004.

[29] M. Alaaeldin, E. Alsusa, K. G. Seddik, and W. Mesbah, "Aod-adaptive channel feedback for fdd massive mimo systems with multiple-antenna users," *IEEE Access*, vol. 10, pp. 4431–4447, 2022.

[30] R. Rajashekar and L. Hanzo, "Iterative matrix decomposition aided block diagonalization for mm-wave multiuser mimo systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1372–1384, 2017.

[31] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[32] L. Simić, N. Perpinias, and M. Petrova, "60 ghz outdoor urban measurement study of the feasibility of multi-gbps mm-wave cellular networks," in *Proc. IEEE Conf. Comput. Commun. Workshops*, 2016, pp. 554–559.

[33] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1273–1289, 2001.

[34] V. Lau, Y. Liu, and T.-A. Chen, "On the design of mimo block-fading channels with feedback-link capacity constraint," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 62–70, 2004.

[35] A. Antoniou and W.-S. Lu, *Fundamentals of Constrained Optimization*. New York, NY: Springer US, 2021, pp. 285–338. [Online]. Available: https://doi.org/10.1007/978-1-0716-0843-2_10

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

[37] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans Antennas Propag*, vol. 34, no. 3, pp. 276–280, 1986.

[38] L. Jiang and H. Jafarkhani, "Multi-user analog beamforming in millimeter wave mimo systems based on path angle information," *IEEE Trans Wireless Commun*, vol. 18, no. 1, pp. 608–619, 2019.

[39] Y. Ahn, T. Kim, and C. Lee, "A beam steering based hybrid precoding for mu-mimo mmwave systems," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2726–2729, 2017.

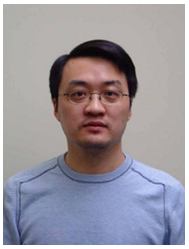
[40] J. Roh and B. Rao, "Design and analysis of mimo spatial multiplexing systems with quantized feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2874–2886, 2006.

[41] J. M. Lee, *Smooth Manifolds*. New York, NY: Springer New York, 2003, pp. 1–29. [Online]. Available: https://doi.org/10.1007/978-0-387-21752-9_1

[42] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7704–7719, 2020.

Mingrui Li received the B.E. degree in communications engineering from Xidian University, Xian, China, in 2020. He is currently pursuing the M.E. degree with the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China. His research interests include wireless communications and sensing-assisted communication.

Xiaowei Qin received the B.S. and Ph.D. degrees from the Department of Electrical Engineering and Information Science, University of Science and Technology of China (USTC), Hefei, China, in 2000 and 2008, respectively, where he has been a Member of Staff with the Key Laboratory of Wireless-Optical Communications, Chinese Academy of Sciences, since 2014. His research interests include optimization theory, service modeling in future heterogeneous networks, and wireless artificial intelligence in mobile communication networks.



Yunfei Chen (Senior Member, IEEE) received his B.E. and M.E. degrees in electronics engineering from Shanghai Jiaotong University, Shanghai, P.R.China, in 1998 and 2001, respectively. He received his Ph.D. degree from the University of Alberta in 2006. He is currently working as an Associate Professor at the University of Warwick, U.K. His research interests include wireless communications, cognitive radios, wireless relaying and energy harvesting.

Weidong Wang received the B.S. degree from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1989, and the M.S. degree from the University of Science and Technology of China, Hefei, China, in 1993. He is currently a Full Professor with the Department of Electronic Engineering and Information Systems, University of Science and Technology of China. His research interests include wireless communication, microwave and millimeter, and radar technology. He is a member of the Committee of Optoelectronic Technology, Chinese Society of Astronautics.

Li Chen received the B.E. in electrical and information engineering from Harbin Institute of Technology, Harbin, China, in 2009 and the Ph.D. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2014. He is currently an Associate Professor with the Department of Electronic Engineering and Information Science, University of Science and Technology of China. His research interests include integrated communication and computation and integrated communication and sensing.