Ray-of-Arrival Passing for Indirect Beam Training in Cooperative Millimeter Wave MIMO Networks

Matthew Kokshoorn, He Chen, Yonghui Li, and Branka Vucetic

Abstract

This paper is concerned with the channel estimation problem in multi-cell millimeter wave (mmWave) wireless systems. We develop a novel Ray-of-Arrival Passing for In-Direct (RAPID) framework, in which a network consisting of multiple BS are able to work cooperatively to estimate jointly the UE channels. To achieve this aim, we consider the spatial geometry of the mmWave environment and transform conventional angular domain beamforming concepts into the more euclidean, Ray-based domain. Leveraging this model, we then consider the conditional probabilities that pilot signals are received in each direction, given that the deployment of each BS is known to the network. Simulation results show that RAPID is able to improve the average estimation of the network and significantly increase the rate of poorer quality links. Furthermore, we also show that, when a coverage rate threshold is considered, RAPID is able to improve greatly the probability that multiple link options will be available to a user at any given time.

I. INTRODUCTION

In order to meet the unprecedented throughput demands of next-generation communications systems, mobile networks are expected to become significantly denser in urban areas [1]–[3]. By reducing the typical cell size, the number of devices that each base station (BS) needs to support can be decreased; however, the smaller inter-cell spacing can lead to increased interference between cells [4]. Supporting densification, the mmWave frequency range (30 GHz-300 GHz) is an appealing spectrum band, due to its much higher atmospheric losses. This propagation characteristic naturally attenuates inter-cell interference and thus can permit carrier frequencies to be reused in cells in closer proximity to one another [5]–[7]. Furthermore, due to its vastly underutilized spectrum, the mmWave frequency range also offers a substantial increase in bandwidth compared to the over-congested microwave spectrum used in existing wireless systems [8], [9]. Although frequency reuse may see a
benefit from these propagation losses, the signals also experience significant reflection and penetration losses that together make wireless communication in the mmWave band very challenging [10]. Overcoming these losses is a critical issue that must be resolved, in order for mmWave networks to meet reliability expectations of emerging technologies such as vehicular communications and the industrial Internet-of-Things (IoT) [11], [12].

The most widely accepted means of overcoming mmWave propagation losses is to employ large multiple-input multiple-output (MIMO) antenna arrays with directional beamforming [8], [13], [14]. Beamforming is also advantageous in further reducing interference in small cell networks, as a coordinated approach in this regard can minimize unwanted signals [4], [15]. Moreover, as spacing between antenna array elements is typically proportional to the carrier wavelength, large mmWave antenna arrays can be implemented while occupying a much smaller form factor. However, the much larger bandwidth expectation in an mmWave system also indicates that conventional digital MIMO architectures may have an unrealistic power consumption, due to the large number of high-rate Analog-to-Digital Converters (ADCs) and digital-to-analog converters (DACs).

Even for completely digital systems, estimating large MIMO channel matrices can be a challenging problem. At mmWave frequencies, this becomes more difficult, due to both hardware constraints and the necessity to use beamforming to overcome propagation losses during initial access and channel estimation. On the other hand, due to these propagation losses, measurements have shown that the mmWave channel is sparse in the spatial domain [8]. Leveraging this phenomenon, each channel can be decomposed into its underlying physical parameters, including an angle-of-arrival (AoA), an angle-of-departure (AoD), and a path coefficient, for a small number of propagation paths. Conversely, microwave frequency MIMO channels exhibit rich scattering environments, which leads to channel models that characterize the superposition of many paths. To capture mmWave sparsity, an alternative MIMO matrix representation, known as the “virtual channel matrix,” can be formed from the set of channel gains between each beamforming direction [16], [17]. Thanks to its sparsity, it is often advantageous to estimate this matrix directly by using compressed sensing (CS) techniques to reduce the number of required measurements [18], [19].

An intuitive benchmark approach to estimating the virtual channel matrix involves exhaustively searching the channel for all possible propagation paths. This may be achieved by transmitting and receiving pilot signals while sequentially adopting combinations of beamforming vectors between the transceiver to search for any paths. Improving on this point, hierarchical codebook-based estimation strategies have been proposed, in order to reduce
significantly estimation overheads by applying a divide-and-conquer type of beam search [18]–[23]. However, as progressive beam refinement converges toward a single path in each estimation process, these approaches have an overhead that is proportional to the number of users and paths. For this reason, approaches that do not adapt toward a specific user, such as Exhaustive Search (ES) and random directional beamforming (RDB), are more appropriate, as they can support the estimation of multiple user channels in the same period [24]–[26]. Furthermore, in order to realize a robust and reliable network, it may also be important for each user to maintain multiple link options to the network [27]–[29]. In particular, this multi-connectivity would be a crucial form of redundancy to support beam switching when one path direction suddenly becomes blocked. Fortunately, in the ultra-dense, user-centric cells expected in next-generation mobile systems, it follows that each user equipment (UE) will typically be in the coverage of large numbers of BS in a given communication period. Furthermore, due to the close relationship between the environment and the physical parameters that make up the mmWave channel, we are able to infer path directions from the network’s spatial geometry.

For a sparse mmWave channel within dense multi-BS deployments, channels between each UE and BS can exhibit high spatial correlation. As these channels can be expressed directly as a function of the physical environment and array orientations, channel decompositions for multiple BS may have propagation paths that correspond to a common scatterer, or direct line-of-sight (LOS) paths to the same UE. Leveraging the same duality between localization and estimation, joint strategies have been proposed for microwave systems in [30]–[32]. Extending these concepts to mmWave systems, [33] proposed a user cooperation estimation strategy which also able to support non-line-of-sight (NLOS) estimation. Although these strategies show a significant improvement over conventional estimations, they often require a precise time difference of arrival (TDoA) or phase information, and they neglect many of the hardware constraints such as hybrid beamforming and quantized phase shifters. Furthermore, as uniform linear arrays (ULAs) are typically considered in mmWave systems, the array orientation is a commonly overlooked variable. In particular, the use of ULAs results in an angle ambiguity problem, where “forward” beamforming pilots/measurements are indistinguishable from the rearward directions (See [24] and Figure 4 therein).

Motivated by the strict hardware constraints in mmWave systems and the need to meet network reliability requirements, we aim to develop a joint channel estimation strategy that is able to utilize spatial dependencies among multiple BS, in order to assist the network’s estimation to each UE. Specifically, in the uplink, we consider a UE that broadcasts beam-
formed pilot signals which can be jointly received by multiple BS. We also show that, if each BS knows the relative position of its neighbors, this physical deployment information can be utilized to identify conditional geometric relationships that exist between virtual channel estimates. To facilitate this aim, we leverage ray tracing principals to transform the channel AoD/AoA measurements into a more Euclidean-focused ray-of-arrival (RoA) and ray-of-reparture (RoD). We then show that a given pair of RoA measurements, received from two different positions, can be considered to sample jointly a position in Euclidean space. Similarly, we leverage geometric dependencies among the RoD to infer conditional transmit directions. We refer to the developed scheme as “Ray-of-Arrival Passing for In-Direct (RAPID) beam training.”

In contrast to the existing work, by focusing on the virtual channel information, we are able to apply our approach to hardware constrained estimation. Furthermore, to provide generality and reduce computational redundancy, we also consider that each BS is only able to share entries from its already estimated channel. As this matrix is inherently sparse, this greatly reduces the bandwidth required to share information among the network. Furthermore, by considering the virtual channel estimate, RAPID is agnostic to how each independent estimation is carried out, and therefore can be implemented on top of existing channel estimation strategies. Results show that the proposed scheme can greatly increase the achievable rate between the transceiver, particularly for links that would have normally been quite poor. By considering a minimum rate requirement, we also show that RAPID is able to significantly increase the coverage probability for having a greater number of available links.

We summarize the main contributions of this as follows:

• We investigate a multi-cell user-centric mmWave communication system, in which a UE broadcasts pilot signals to a number of BSs. We generalize the concept of the AoD/AoA beamforming to a ray-based RoD/RoA estimation. We apply this model to the widely used ULA, and develop an estimator that is robust to the angle ambiguity problem. We also show that for a BS pair with a known relative displacement, many of their virtual channel entries are mutually dependent.

• We use the Ray-based model to develop a Bayesian estimator so that each BS may compute the probability of a path on its beamforming directions, given the channel estimates provided by the other network BSs. Results show a significant improvement for both the average achievable rate and network coverage when compared to conventional schemes.

• In order to reduce sharing overheads in bandwidth constrained networks, we exploit the
channel sparsity and propose to use limited information exchange among BS. To this end, we reduce the interchange to only the most dominant virtual channel entries, that exhibit a mutually dependent relationship for another BS. In addition to this, we also show that the only prior information required for RAPID is the relative position and orientation of each BS. In this sense, if the UE is also aware of this relative deployment, the proposed scheme can also be applied to the downlink. By adopting multiple access scheme for downlink pilots such as Code Division Multiple Access (CDMA), no sharing overhead would be required in this case.

II. SYSTEM MODEL

Consider a mmWave cellular network consisting of \( B \) BS—each equipped with an array of \( N_{BS} \) antenna. We adopt a user-centric deployment model, in which a UE is located at the origin of a two-dimensional coordinate system (i.e., \((x_u, y_u) = (0, 0)\)). We further assume that the UE is equipped with an array of \( N_{UE} \) antennas. Relative to the origin of this system, we consider that the deployment of the \( b \)th BS antenna array can be described by a 2D translation and a rotation, denoted by \( D_b = (x_b, y_b) \) and \( \Theta_b \), respectively. We consider both the BS and UE antenna arrays to have an orientation denoted by \( \Theta_b \in [-\pi, \pi] \) and \( \psi_u \in [-\pi, \pi] \), respectively. We consider this orientation to be defined as the counter-clockwise angle from the x-axis to the ULA. We further denote the relative displacement vector from the \( p \)th BS to the \( q \)th BS as \( \Delta_{p,q} = [\delta_{x_{p,q}}, \delta_{y_{p,q}}] = D_p - D_q \). In this paper, we use the term “local reference frame” to refer to angles relative to a particular antenna array. Conversely, angles in the “global reference frame” refer to absolute angles in the global 2D coordinate system. This distinction is important, as UE orientation cannot be known to the network a priori. An example deployment configuration is shown in Fig. [I]

To estimate the uplink channel matrix, we assume that each UE simultaneously broadcasts a sequence of beamformed pilot signals\(^1\). Similarly, all BS collect these signals by adopting a sequence of beamforming vectors. We consider that both the UE and each BS are equipped with a limited number of radio frequency (RF) chains, denoted by \( R_{BS} \) and \( R_{UE} \), respectively. Denote \( f_i \) as the \( N_{UE} \times 1 \) transmit beamforming vector adopted by the \( i \)th RF chain at the UE. Similarly, denote by \( w_j^{(b)} \), the \( N_{BS} \times 1 \) receiving beamforming vector adopted by the \( j \)th RF chain of the \( b \)th BS.

\(^1\)Orthogonality among multiple UE can be achieved by carrier-independent multiple access schemes such as CDMA or Time Division Multiple Access (TDMA).
Fig. 1 – An example deployment illustrating each parameter in our model. The UE can be seen at the origin, with antenna array orientation denoted by $\psi_u$. Similarly, each BS is deployed at $b_mD_b = (x_b, y_b)$ and with orientation $\Theta_b$. To make the illustration more explicit, we show $B = 3$ BS, indexed as $b_1 = 1$, $b_2 = 2$ and $b_3 = 3$. Relating to the Euclidean channel model, we label the angle-of-departure (AoD) in the UE local reference frame (i.e., relative to its own array orientation) $\phi_L^b$ and the corresponding global reference frame (i.e., relative to the global coordinate system) angle $\phi_G^b$. Similarly for each BS, we have the local reference frame angle-of-arrival (AoA) as $\theta_L^b$ and the global angle as $\theta_G^b$.

Following [18], we consider the beamforming vectors, at each link end, as being limited to networks of RF phase shifters. As such, all elements of $f_i$ and $w_j^{(b)}$ are constrained to have a constant modulus and unit norm, such that $||f_i|| = 1, \forall \ i = 1, \cdots, R_{UE}$, and $||w_j^{(b)}|| = 1, \forall \ j = 1, \cdots, R_{BS}, b = 1, \cdots, B$. We further assume that due to hardware constraints, each of the phase shifters (i.e., the entries of $f_i$ and $w_j^{(b)}$) is digitally controlled and takes on quantized values from the predetermined set

$$\left\{ \frac{1}{\sqrt{N}} \exp(jq_k) \right\}, \forall k = 1, \cdots, N, \quad (1)$$

where $q_k = \pi - 2\pi(k - 1)/N$ and $N \in \{ N_{UE}, N_{BS} \}$ is the number of antennas in the array.
That is, each UE (BS) phase shifter can only use one $N_{\text{UE}}$ ($N_{\text{BS}}$) uniformly spaced point around the unit circle, respectively, which can therefore be digitally controlled by $\lceil \log_2 N \rceil$ bits.

Let $F = [f_1, f_2, \ldots, f_{R_{\text{UE}}}]$ denote the $N_{\text{UE}} \times R_{\text{UE}}$ UE beamforming matrix, with columns representing the $R_{\text{UE}}$ RF beamforming vectors. The corresponding $N_{\text{UE}} \times 1$ UE transmit signal can be represented as

$$x = \sqrt{\frac{P}{R_{\text{UE}}}} F s, \quad (2)$$

where $P$ is the UE’s pilot transmit power and $s$ is the $R_{\text{BS}} \times 1$ vector of transmit pilot symbols corresponding to each beamforming vector with $E[ss^H] = I_{R_{\text{UE}}}$. We adopt the widely used block-fading channel model, such that the signal observed by the $b$th BS can be expressed as

$$r^{(b)} = H^{(b)} x + q^{(b)} = \sqrt{\frac{P}{R_{\text{UE}}}} H^{(b)} F s + q^{(b)}, \quad (3)$$

where $H^{(b)}$ denotes the $N_{\text{BS}} \times N_{\text{UE}}$ MIMO channel matrix between the UE and the $b$th BS, and $q^{(b)}$ is an $N_{\text{BS}} \times 1$ complex additive white Gaussian noise (AWGN) vector for the $u$th user, following distribution $\mathcal{CN}(0, N_0 I_{N_{\text{BS}}})$.

Each BS processes the received pilot signals with each of the $R_{\text{BS}}$ RF chains. By denoting $W^{(b)} = [w^{(b)}_1, w^{(b)}_2, \ldots, w^{(b)}_{R_{\text{BS}}}]$ as the $N_{\text{BS}} \times R_{\text{BS}}$ combining the matrix at the $b$th BS, we express the $R_{\text{BS}} \times 1$ vector of the $b$th BS received signals as

$$y^{(b)} = (W^{(b)})^H H^{(b)} x + n^{(b)} \quad (4)$$

where the noise term follows the distribution

$$n^{(b)} = (W^{(b)})^H q^{(b)} \sim \mathcal{CN}(0, N_0 (W^{(b)})^H W^{(b)}). \quad (5)$$

We follow [34] and adopt a two-dimensional (2D) sparse geometric channel model. We consider that only a single dominant path is present between the UE and each BS, leaving the extension to joint scatterer estimation as a future work. Using this model, each candidate uplink channel between the UE and the $b$th BS can be characterized in its local reference frame by an AoD, $\phi^L_b$, an AoA, $\theta^L_b$, and a path coefficient, namely $\alpha_b$. The corresponding MIMO channel between the UE and the $b$th BS can be expressed in terms of these physical
parameters as

$$H^{(b)} = \alpha_b \sqrt{N_{\text{UE}}N_{\text{BS}}} a_{\text{BS}}(\theta^{L}_{b})(a_{\text{UE}}(\phi^{L}_{b}))^{H}$$  \hspace{1cm} (6)$$

where \(a_{\text{BS}}(\theta^{L}_{b})\) and \(a_{\text{UE}}(\phi^{L}_{b})\) denote the BS and UE arrays’ spatial signatures, respectively. We adopt a flat block fading model and assume that the path coefficient remains unchanged through the entire channel estimation process. We assume that the value of the path coefficient follows the zero mean complex distribution \(\alpha_b \sim \mathcal{CN}(0, \sigma_b^2)\), where the expected power, \(\sigma_b^2\), is inversely proportional to the radial displacement between the BS and UE as \(\sigma_b^2 = r_b^{-\beta}\), and where \(r_b = ||D_b||_2 = \sqrt{x_b^2 + y_b^2}\) is the radial distance and \(\beta\) is the path loss exponent.

We consider that the BS and each UE are equipped with ULA. We can then write

$$a_{\text{BS}}(\phi^{L}_{b}) = u(\phi^{L}_{b}, N_{\text{UE}}) \quad \text{and} \quad a_{\text{UE}}(\theta^{L}_{b}) = u(\theta^{L}_{b}, N_{\text{BS}}),$$

respectively, whereby

$$u(\epsilon, N) \triangleq \frac{1}{\sqrt{N}}[1, e^{j2\pi d \cos(\epsilon)/\lambda}, \ldots, e^{j2\pi d(N-1)/\lambda \cos(\epsilon)}]^T.$$  \hspace{1cm} (7)$$

In (7), \(N \in \{N_{\text{UE}}, N_{\text{BS}}\}\) is the number of antenna elements in the array, \(\lambda\) denotes the signal wavelength, and \(d\) denotes the spacing between antenna elements. With half-wavelength spacing, the distance between antenna elements satisfies \(d = \lambda/2\).

To estimate channel information, beamforming vectors are selected from a predetermined set of candidate beamforming vectors at each link end. We denote the candidate beamforming matrices as \(F_c\) and \(W_c\), the columns of which comprise all candidate beamforming vectors at the UE and BS, respectively. For ease of practical implementation, we consider the candidate beams to be subject to quantized phase-shifting constraints, and therefore they represent the set of all possible beamforming vectors that may be used later for data communication. Following (1), this leads to \(N_{\text{UE}}\) orthogonal transmitting candidate beams and \(N_{\text{BS}}\) orthogonal receiving candidate beams. The \(N_{\text{BS}} \times N_{\text{UE}}\) matrix formed by the product of the MIMO channel and these two candidate beamforming matrices is commonly referred to as the “virtual channel matrix” \([18]\), given by

$$V^{(b)} = \frac{1}{\sqrt{N_{\text{UE}}N_{\text{BS}}}} (W_c)^H H^{(b)} F_c.$$  \hspace{1cm} (8)$$

We therefore aim to estimate this matrix so that beam pairs that result in strong channel gains can be identified for data communication. The key challenge here is determining how to design a sequence of beamforming vectors in such a way that the channel parameters can be quickly and accurately estimated, leaving more time for data communication and thus achieving a higher throughput.
To facilitate our proposed cooperative channel estimation scheme, we assume that all BS are able to maintain a reliable link between one another and are thus able to share mutually dependent information. Initially, we consider complete information sharing, however we later restrict this to the bandwidth constrained channel by only sharing dependent measurements of significant signal strength. In the following sections, we leverage mutual channel information, in order to develop a cooperative BS framework.

III. RAPID Beam Training

In this section, we first extend our channel model to consider the Euclidean deployment scenario. We then introduce the sequence of measurement beamforming vectors adopted in our proposed estimation scheme, and then we extend these into the 2D geometric model. In doing so, we propose a shift from the conventional single-link-oriented AoD and AoA model, to a more Euclidean-focused RoD and RoA model. Subsequent sections then develop a means of jointly computing the probability of each beamforming combination, given the mutually dependent information provided by cooperating BS.

A. Euclidean Space MIMO Channel

In order to develop our joint estimation strategy, we begin by incorporating the 2D deployment into the angular channel in (6). To this end, we denote the global AoD of the bth propagation path as $\phi_b^G$, i.e., the angle of the propagation from the global frame, irrespective of the UE orientation. Similarly, we denote the global AoA at the BS end as $\phi_b^G$. Recalling the array orientations $\psi_u$ and $\Theta_b$, we can relate global angles into the local reference frame AoD and AoA (i.e., the local beam steering directions) as $\phi_b^L = \phi_b^G - \psi_u$ and $\theta_b^L = \theta_b^G - \Theta_b$. Substituting these into (6) leads subsequently to a global description of the channel model

$$H(b) = \alpha_b \sqrt{N_{UE}N_{BS}} a_{BS}(\theta_b^G - \Theta_b)(a_{UE}(\phi_b^G - \psi_u))^H.$$  \hspace{1cm} (9)

By observing the geometric relationships in Fig. 1 we can rate the global AoD and AoA further by considering BS deployments with signed trigonometric relationships $\tan(\phi_b^G) = y_b/x_b$ and $\tan(\theta_b^G) - \pi = y_b/x_b$. Using the four-quadrant inverse tangent function, denoted by $\text{atan2}(a,b)$, the LOS dominant channel in (9) can be rewritten as

$$H(b) = \alpha_b \sqrt{N_{UE}N_{BS}} a_{BS}(\text{atan2}(-y_b, -x_b) - \Theta_b)(a_{UE}(\text{atan2}(y_b, x_b) - \psi_u))^H$$  \hspace{1cm} (10)

$$= \alpha_b \sqrt{N_{UE}N_{BS}} E_{BS}(-y_b, -x_b, \Theta_b)(E_{UE}(y_b, x_b, \psi_u))^H$$ \hspace{1cm} (11)
where \( E_{BS}(-y_b, -x_b, \Theta_b) = a_{BS}(\text{atan2}(-y_b, -x_b) - \Theta_b) \) and \( E_{UE}(y_b, x_b, \psi_u) = a_{UE}(\text{atan2}(y_b, x_b) - \psi_u) \) describe antenna spatial signatures in terms of the Euclidean deployment parameters for the BS and UE, respectively.

If each BS is deployed without knowing the orientation and relative position surrounding the BS, each independent estimation is limited to the angular parameters in (9), subject to the constraints of mmWave beamforming. Accurate estimation of these parameters permits accurate beam selection communication and consequently an accurate recovery of the fading coefficient from the pilot measurement. However, for a deployment where the orientation and relative position of surrounding BS is known, the \( b \)th BS can focus on the estimation of the parameters in (10) as \( \hat{x}_b, \hat{y}_b \) and the UE orientation \( \hat{\psi}_u \). By expressing these estimation parameters in terms of \( p \)th BS as \( [\hat{x}_b, \hat{y}_b] = [\hat{x}_p, \hat{y}_p] + \Delta_{b,p}, \forall b = 1, \cdots, B \), it is evident that each BS can then reconstruct not only its own channel, but also the channel of other BS in the network. Furthermore, as the UE can only exist in a single position, estimations among BS are mutually dependent. This relationship supports the spatial correlation in the Euclidean channel and gives motivation for cooperation among BS to achieve accurate joint channel estimation.

### B. Candidate Beamforming Measurements

In this paper, we follow [24] and adopt random directional beam steering at each link end. To achieve this aim, we elaborate on the UE candidate beamforming matrix in (8) as \( W_c = [w_c(1), \ldots, w_c(N_{UE})] \). Similarly, we define the BS candidate beamforming matrix as \( F_c = [f_c(1), \ldots, f_c(N_{BS})] \), following which, in each pilot transmission time slot, a unique pseudo-random candidate beamforming vector is adopted by each RF chain at the UE and similarly at each BS. In order for each BS to collect simultaneously and fairly pilot signals from all users, we consider the pseudo random selection of candidate beams as having equal probability\(^2\)\(^\text{f}^\text{o}\). As each random selection is assumed to derive from a pseudo random process, the entire selection sequence can be predicted by both the UE and each BS, so long as the UE maintains a synchronized random seed within the network.

By recalling the ULA response in (7), the resulting set of orthogonal candidate beams that satisfy the quantized phase-shifting constraints in (1) becomes [24]

\(^2\text{In [24], each candidate beam is assigned a non-uniform probability of selection, which is later adaptively re-weighted to improve performance. As this results in each receiver adapting its beams toward a single user, we do not consider this approach herein.}\)
\[ f_c(n_u) = a_{UE}(\bar{\phi}_{n_u}), \forall \bar{\phi}_{n_u} = \cos^{-1}\left(1 - \frac{2n_u}{N_{UE}}\right), \ n_u \in \mathcal{N}_U \]  

and

\[ w_c(n_b) = a_{BS}(\bar{\theta}_{n_b}), \forall \bar{\theta}_{n_b} = \cos^{-1}\left(1 - \frac{2n_b}{N_{BS}}\right), \ n_b \in \mathcal{N}_B \]

where the candidate beam steering indexes at the UE and BS are denoted by \( n_u \)th and \( n_b \)th, respectively, and belong to the sets \( \mathcal{N}_U = \{0, \ldots, N_{UE} - 1\} \) and \( \mathcal{N}_B = \{0, \ldots, N_{BS} - 1\} \). Due to the quantized phase-shifting constraints, each candidate beam steering vector is orthogonal to the others, and therefore together they satisfy \( F_c F^H_c = I_{N_{UE}} \) and \( W_c W^H_c = I_{N_{BS}} \). The example set of candidate beam patterns in Fig. 1 shows each BS with \( N_{BS} = 8 \) and the UE with \( N_{UE} = 16 \). In the same figure, it is also evident that the candidate beams on the range \([0, \pi]\) are repeated in the range \([0, -\pi]\) i.e., \( a_{BS}(\theta) = a_{BS}(-\theta) \) and \( a_{UE}(\phi) = a_{UE}(-\phi) \), due to the one-dimensional nature of ULA, which leads to the candidate beams’ indexes ambiguously describing angles from either range. We discuss this in greater detail in subsequent sub-sections.

By using a random sequence of candidate beams to transmit and receive each pilot symbol, as described in (12) and (13), the sequence of \( M \) measurements that are collected by the \( b \)th BS can be expressed by the \( R_{BS} \times 1 \) measurement vector by

\[
y^{(b)} = \sqrt{\frac{P}{R_{UE}}} \left[ (W_1^{(b)})^H H^{(b)} F_1 s_1 \right] + \left[ n_1^{(b)} \right] + \left[ n_2^{(b)} \right] + \cdots + \left[ n_m^{(b)} \right],
\]

where \( F_m \) and \( W_m^{(b)} \) are the matrices whose columns consist of the \( R_{UE} \) and \( R_{BS} \) randomly selected candidate beam steering vectors at the UE and BS, respectively. Due to orthogonality among the BS candidate beams, the noise elements in (5) now follow an i.i.d., AWGN distribution.

By rearranging (8) to get \( H^{(b)} = \sqrt{N_{UE}N_{BS}}W_c V^{(b)} F^H_c \), we can substitute this result into
and express the measurement vector in the common CS form \[35\] as follows

\[
y^{(b)} = A_g \begin{bmatrix} A_1^{(b)} \\ \vdots \\ A_m^{(b)} \end{bmatrix} \text{vec}(V^{(b)}) + \begin{bmatrix} n_1^{(b)} \\ \vdots \\ n_m^{(b)} \end{bmatrix}
\]

\[
= A_g A^{(b)} v^{(b)} + n^{(u,m)}.
\]

where \( A_m^{(b)} = (s_m^T F_m^T F_m^*) \otimes ((W_m^{(b)})^H W_c) \) is the \( R_{BS} \times N_{UE} N_{BS} \) sparse sensing matrix, \( v^{(b)} = \text{vec}(V^{(b)}) \) is the vectorized virtual channel matrix between the UE and the \( b \)-th BS, and \( A_g = \sqrt{P N_{UE} N_{BS}} / R_{UE} \) is a scalar measurement gain.

C. Independent Base Station Channel Estimation

Following the measurement sequence in the previous sub-sections, we assume that each BS independently estimates its own the virtual channel, \( \hat{v}^{(b)} \), based on measurements it has collected over the span of \( T_E \) time slots in \( y^{(b,m)} \). Considering the CS matrix \( A^{(b,m)} \), this sparse recovery problem can formulated as

\[
\hat{v}^{(b)} = \arg\min_v \left[ \|y^{(b)} - A_g A^{(b)} v\|_2^2 + \gamma \|v\|_1 \right].
\]

In this paper, we consider that each independent channel estimation is obtained using the Bernoulli Gaussian (BG) Generalized Approximate Message Passing (GAMP) approach described in \[36\], \[37\]. After obtaining this initial channel estimate, each BS can then convert the vectorized channel estimate back into its matrix form (i.e., \( \hat{V}^{(b)} \)). In the following subsections, we develop a framework that permits each BS to then share its mutually dependent indexes with the rest of the network so that the joint probability of each beam combination may be computed. Although we have adopted a BG GAMP-based estimator in this paper, in practice RAPID is not limited to any particular independent estimation/recovery technique; rather, any approximate solution to (17) may be considered as an input into our proposed algorithm.

D. Bipolar Candidate Ray Measurements

Following a similar process as the Euclidean channel formulation in \[10\], we now also seek to transform the candidate beamforming vectors into the Euclidean deployment model. To this end, we consider that each of the candidate beamforming vectors, conventionally considered to measure an angular AoD/AoA, instead corresponds to a ray-based RoD/RoA. We model
each ray to begin at the center of each ULA and extend with a radial distance denoted by \( r_b \) in the direction corresponding to each AoD/AoA. By adopting a bipolar parametric line model, we can describe the \((x,y)\) coordinate pairs that lie on the \( n_{b} \)th candidate RoA for the \( b \)th BS as

\[
P_{b} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r_b \cos(\tilde{\theta}_{n_b} - \Theta_b) \\ r_b \sin(\tilde{\theta}_{n_b} - \Theta_b) \end{bmatrix} + \begin{bmatrix} x_b \\ y_b \end{bmatrix}
\]

\[
= r_b \mathbf{R}(n_b) + \mathbf{D}_b^T, \quad \forall \ r_b > 0.
\]

(18)

(19)

By recalling the relationship between the candidate index and angle, \( \tilde{\theta}_{n_b} = \cos^{-1}(1 - \frac{2n_b}{N_{BS}}) \), we further elaborate \( \mathbf{R}(n_b) \) as

\[
\mathbf{R}(n_b) = \begin{bmatrix} L_x(n_b) \\ L_y(n_b) \end{bmatrix} = \begin{bmatrix} \cos(\cos^{-1}(1 - \frac{2n_b}{N_{BS}}) - \Theta_b) \\ \cos(\cos^{-1}(1 - \frac{2n_b}{N_{BS}}) - \Theta_b - \frac{\pi}{2}) \end{bmatrix}
\]

\[
= \begin{bmatrix} \pm \sqrt{1 - (1 - \frac{2n_b}{N_{BS}})^2} \sin(\Theta_b) + (1 - \frac{2n_b}{N_{BS}})\cos(\Theta_b) \\ \pm \sqrt{1 - (1 - \frac{2n_b}{N_{BS}})^2} \cos(\Theta_b) - (1 - \frac{2n_b}{N_{BS}})\sin(\Theta_b) \end{bmatrix}.
\]

(20)

(21)

where the simplification in (21) follows the trigonometric property \( \cos(\cos^{-1}(a) - b) = \pm \sqrt{(1 - a^2)}\sin(b) + a\cos(b) \).

At this point, it is important to consider the square root term in (21). In particular, it is notable that each candidate beam index corresponds to two indistinguishable ROAs in the Euclidean space, as indicated by the plus-minus sign. As previously eluded to, this is an inherent property that arises from the use of uniform linear arrays, due to the symmetric property \( \mathbf{a}_{BS}(+ \tilde{\theta}_{n_b}) = \mathbf{a}_{BS}(- \tilde{\theta}_{n_b}) \forall n_b \in N_B \). In the context of AoD/AoA estimation, this leads to an ambiguity problem, in that any given angle estimate could be one of two possibilities. For point-to-point systems, there is generally little benefit in resolving this ambiguity, as the transceiver will still be unable to direct its beam in only one of the directions.\(^3\) However, in order for estimated directions to be considered in a Euclidean deployment, angle ambiguity can be an important source of uncertainty. This directional ambiguity is illustrated in Fig. 2 where the UE is shown as being positioned on two different ROAs extending from the right-hand BS. Although a single BS cannot, by itself, determine which of the two RoA directions correspond to a propagation path, there must be one globally

\(^3\)In more complex multi-user systems, this information could, however, be utilized to coordinate the reduction of interference among users [15].
consistent solution among all BS. More generally, Fig. 2 also illustrates the RoA-based model.

To consider this angle ambiguity in our proposed approach, we replace the “unipolar” BS candidate beam indexes \( n_b \in \mathcal{N}_B \) with a super set of “bipolar” indexes \( \bar{n}_b \in \bar{\mathcal{N}}_B = \{-N_B, N_B\} = \{-N_{BS} + 1, \cdots, N_{BS} - 1\} \). With this bipolar definition, we define more rigorously each candidate beamforming vector in (13) as

\[
w_c(n_b) = a_{BS}(\pm \bar{\theta}_{n_b}), \quad \bar{n}_b \in \bar{\mathcal{N}}_B
\]

which leads to

\[
w_c(n_b) = w_c(|\bar{n}_b|) = a_{BS}(\bar{\theta}_{n_b}), \quad \forall \bar{\theta}_{n_b} = \text{sgn}(\bar{n}_b)\cos^{-1}\left(1 - \frac{2|\bar{n}_b|}{N_{BS}}\right), \quad \bar{n}_b \in \bar{\mathcal{N}}_B
\] (22)

Using this definition, we can also express more explicitly the ambiguity in (20) as

\[
R(\bar{n}_b) = \begin{bmatrix}
\text{sgn}(\bar{n}_b)\sqrt{1 - \left(1 - \frac{2|\bar{n}_b|}{N_{BS}}\right)^2}\sin(\Theta_b) + \left(1 - \frac{2|\bar{n}_b|}{N_{BS}}\right)\cos(\Theta_b) \\
\text{sgn}(\bar{n}_b)\sqrt{1 - \left(1 - \frac{2|\bar{n}_b|}{N_{BS}}\right)^2}\cos(\Theta_b) - \left(1 - \frac{2|\bar{n}_b|}{N_{BS}}\right)\sin(\Theta_b)
\end{bmatrix}
\] (23)

Similarly, for each UE candidate beamforming vector, we can elaborate on (12) as

\[
f_c(n_u) = f_c(|\bar{n}_u|) = a_{UE}(\bar{\phi}_{n_u}), \quad \forall \bar{\phi}_{n_u} = \text{sgn}(\bar{n}_u)\cos^{-1}\left(1 - \frac{2|\bar{n}_u|}{N_{UE}}\right), \quad \bar{n}_u \in \bar{\mathcal{N}}_U
\] (24)

where \( \bar{\mathcal{N}}_U = \{-\mathcal{N}_U, \mathcal{N}_U\} = \{-N_{UE} + 1, \cdots, N_{UE} - 1\} \).

Although this bipolar RoA model has limited benefit in a point-to-point system, it facilitates the sharing of information among transceiver arrays concurrently operating in the same Euclidean space. More specifically, when a pair of BS adopts a pair of intersecting candidate RoA beamforming vectors, the pair of measurements cannot be considered only to sample their independent angular directions, but also to sample the position at which the two ROAs intercept. This increases mutual information and therefore can be used to enhance joint estimation performance. We later use geometric reasoning to find the conditional relationship between RoD extending from the UE.

E. Identifying Mutually Dependent Rays

We now build on the previous model by jointly considering the RoA of another BS operating concurrently in the same space as the \( b \)th BS. Specifically, we consider the \( \bar{n}_p \)th candidate RoA extending from the \( p \)th BS, where \( p \neq b \). By recalling (18), we can describe the common intercept between the \( \bar{n}_p \)th RoA and the \( \bar{n}_b \)th RoA, if it exists, by

\[
P_b = P_p
\] (25)

\[
r_b R(\bar{n}_b) + D_b^T = r_p R(\bar{n}_p) + D_p^T
\] (26)
where the solution is only valid if both radial distances are in a positive range, namely $r_b > 0$ and $r_p > 0$. By recalling $\Delta_{p,b} = \Delta_p - D_b = [\delta_{xb,p}, \delta_{yb,p}]$, (25) can be rearranged to become

$$\Delta_{p,b}^T = [\mathbf{R}(\bar{n}_b) \mathbf{R}(\bar{n}_p)] \begin{bmatrix} r_b \\ r_p \end{bmatrix}$$

(27)

\[
\begin{bmatrix}
\delta_{xb,p} \\
\delta_{yb,p}
\end{bmatrix} = \begin{bmatrix}
R_x(\bar{n}_b) & -R_x(\bar{n}_p) \\
R_y(\bar{n}_b) & -R_y(\bar{n}_p)
\end{bmatrix} \begin{bmatrix}
r_b \\
r_p
\end{bmatrix}.
\]

(28)

By multiplying each side of (27) by the inverse of the $2 \times 2$ square matrix, we can solve for the pair of radial distances from each BS to the common intercept among their RoA. We therefore express this pair in vector form as

$$\begin{bmatrix}
r_b \\
r_p
\end{bmatrix} = \begin{bmatrix}
R_x(\bar{n}_b) & -R_x(\bar{n}_p) \\
R_y(\bar{n}_b) & -R_y(\bar{n}_p)
\end{bmatrix}^{-1} \begin{bmatrix}
\delta_{xb,p} \\
\delta_{yb,p}
\end{bmatrix}.$$ (29)

Leveraging the closed form expression for a $2 \times 2$ matrix inverse, we can then directly express the radial distances $r_b$ and $r_p$ as a function of the RoA index pair $\bar{n}_b$ and $\bar{n}_p$ as

$$r_b(\bar{n}_b, \bar{n}_p) = \frac{R_x(\bar{n}_p)\delta_{yb,p} - R_y(\bar{n}_p)\delta_{xb,p}}{R_x(\bar{n}_p)R_y(\bar{n}_b) - R_x(\bar{n}_b)R_y(\bar{n}_p)}, \exists r_b > 0$$

(30)

and

$$r_p(\bar{n}_b, \bar{n}_p) = \frac{R_x(\bar{n}_b)\delta_{yb,p} - R_y(\bar{n}_b)\delta_{xb,p}}{R_x(\bar{n}_p)R_y(\bar{n}_b) - R_x(\bar{n}_b)R_y(\bar{n}_p)}, \exists r_p > 0.$$ (31)

We can then express the set of indexes that result in intercepts among the $\bar{n}_b$th RoA at the $b$th BS and all RoA at the $p$th BS as

$$\mathcal{R}_{\bar{n}_b}^{(p)} = \left\{ \bar{n}_p \in \mathcal{N}_B \mid \exists r_b(\bar{n}_b, \bar{n}_p) > 0 \land r_p(\bar{n}_b, \bar{n}_p) > 0 \right\}.$$ (32)

Finally, by noting that the solution in (30)-(31) only depends on the relative displacement and orientation between each BS pair, we can substitute each radial distance back into $\mathbf{P}_b = r_b \mathbf{R}(\bar{n}_b) + D_b^T$ and $D_b^T$ to describe the corresponding set of intercept positions relative to the $b$th BS. We therefore express the set of positions corresponding to each intercept in (32) as

$$\mathcal{P}_{\bar{n}_b}^{(p)} = \left\{ (r_b(\bar{n}_b, \bar{n}_p) R_x(\bar{n}_b), r_b(\bar{n}_b, \bar{n}_p) R_y(\bar{n}_b)) \mid \bar{n}_p \in \mathcal{R}_{\bar{n}_b} \right\}.$$ (33)

where each $(\bar{x}_b, \bar{y}_b) \in \mathcal{P}_{\bar{n}_b}^{(p)}$ describes the displacement, from the $b$th BS, to an intercept
between the $\hat{n}_b$th RoA and the $\hat{n}_p$th RoA.

The quantization of points described by the set in \((33)\) is a result of quantized phase-shifting constraints and, subsequently, the finite set of candidate beamforming directions that form the virtual channel matrix. For this reason, the grid of points formed by all intercepts from every BS RoA can therefore be thought of as a virtual channel in the Euclidean space. In the following sections, we develop a joint estimation tool to leverage the mutual information collected by the BS channel measurements. Intuitively, RAPID permits estimated information from one BS to assist in another.

\subsection*{F. RAPID Beam Probabilities}

From \((17)\), recall that, after the UE has finished transmitting its sequence of pilot symbols, each BS is able to make an independent estimate of the up-link channel. In this sub-section, we consider that the $b$th BS has its own estimate, denoted by $\hat{V}^{(b)}$, and also has access to the estimated virtual channel from the $p$th BS, denoted as $\hat{V}^{(p)}$. Leveraging the model developed in the previous sub-sections, we now aim to utilize the mutual dependency among virtual channel entries to find the joint probability of each direction. In doing so, we therefore collectively increase overall network performance.

We begin by considering a single entry of the $b$th BS’s virtual channel $V^{(b)}_{n_b,n_u}$, which here denotes the estimated path gain between the $n_b$th BS candidate beamforming vector $n_u$th and the UE candidate beamforming vector. From \((12)-(13)\), we can therefore elaborate this particular entry as

$$V^{(b)}_{n_b,n_u} = \frac{1}{\sqrt{N_{UE}N_{BS}}} (w_c(n_b))^H f^{(b)} c(n_u). \tag{34}$$

By recalling \((6)\), we can then consider the conditional probability density function (PDF) of the channel in \((34)\), given the channels AoD and AoA that are perfectly aligned with the beamforming vectors $f_c(n_u)$ and $w_c(n_b)$ i.e., $\theta^L_b \in \{-\bar{\theta}_{n_b}, \bar{\theta}_{n_b}\}$ and $\phi^L_b \in \{-\bar{\phi}_{n_u}, \bar{\phi}_{n_u}\}$. From \((6)\), we can substitute these angles into \((34)\) to yield

$$f(V^{(b)}_{n_b,n_u}|n_b,n_u) = f(\alpha_b(w_c(n_b))^H a_{BS}(\pm \bar{\theta}_{n_b}) (a_{UE}(\pm \bar{\phi}_{n_u}))^H f_c(n_u)|n_b,n_u)$$

$$= f(\alpha_b|n_b,n_u). \tag{35}$$

By itself, this conditional destiny function is the same as that used in independent estimation, with the PDF of $\alpha_b$ being limited to $f(\alpha_b|r_b) \sim \mathcal{CN}(0, r_b^{-\beta})$ for some unknown radial distance,
Fig. 2 – An example deployment model ($B = 2$) showing the set of bipolar ray-based rays-of-arrival (RoAs) extending from each BS as dashed lines. In this example, the true user equipment (UE) position is shown at the origin, with bold-solid beam patterns depicting the ray-of-reparture (RoD) directions expected to have strong gains. The scenario illustrated here is conditional upon the true $n_u$ and $n_b$. As such, the other UE positions and pairs of orientations show the conditional rotation of the correct directions. It is evident that the origin UE position aligns both $n_b$ and the conditionally dependent $n_p$. Therefore, both $V_r^{(b)}_{[n_b],[n_u]}$ and $V_r^{(p)}_{[n_p],[n_u]}$ would be expected to be strong. In each other position, either one or all of the directions do not align, and thus they correspond to expected weak codependent pairs for each virtual channel. In particular, the geometry of the lower right-hand UE position does result in the correct alignment of each AoD and will therefore index the correct columns of both $V^{(b)}$ and $V^{(p)}$. However, as the RoA to the $p$th BS is misaligned, the resulting joint probability will be low. Furthermore, the expected channel gain for this distant position will be significantly lower than the one observed from the true UE position at the origin.

$r_b$. However, by utilizing the models developed in the previous sub-sections, we are now able to consider all available information and extend each conditional $n_b$ and $n_u$ into the global deployment, namely by considering the implication of each conditional for all other BS.

By recalling array orientations and the resulting relationship between the global and local angles as $\phi^L_b = \phi^G_b - \psi_u$, $\forall \ b = 1, \ldots, B$, and $\theta^L_b = \theta^G_b - \Theta_b$, $\forall \ b = 1, \ldots, B$, we can deduce
the following conditionally related angle sets.

**Theorem 1** RAPID Theorem of Mutually Dependent Beam Angles: If the channel between the UE and the $b$th BS has a LOS propagation path such that $\theta_L^b \in \{-\bar{\theta}_{n_b}, \bar{\theta}_{n_b}\}$ and $\phi_L^b \in \{-\bar{\phi}_{n_u}, \bar{\phi}_{n_u}\}$ with a corresponding virtual channel density function, as described by (35), then for the $p$th BS to have jointly a LOS path to the same UE, it follows that the $p$th BS must also have

$$f \left( \frac{\left( w_c(|\tilde{n}_p|) \right) H^{(p)} \alpha_{UE}(\tilde{\theta}_p, \tilde{n}_p, n_u)}{\sqrt{N_{UE} N_{BS}}} \right)_{n_b, n_u} = f(\alpha_p|n_b, n_u),$$

where the conditional AoD to the $p$th BS, $\phi_L^p(\tilde{n}_b, \tilde{n}_p, n_u)$, satisfies the relationship

$$\phi_L^p(\tilde{n}_b, \tilde{n}_p, n_u) = \text{atan2}(r_b(\tilde{n}_b, \tilde{n}_p) R_y(\tilde{n}_b) - \delta_{y_{p,b}}, r_b(\tilde{n}_b, \tilde{n}_p) R_x(\tilde{n}_b) - \delta_{x_{p,b}}) + \bar{\theta}_{\tilde{n}_b} + \bar{\phi}_{n_u} + \Theta_b. \quad (37)$$

**Lemma 1.1 Conditional Line Dependency:** In order for the $b$th BS to receive pilot signals with the $n_b$th candidate beam, the propagation source (i.e., the UE) must be positioned at some point along the RoA line, indexed by $n_b = |\tilde{n}_b|, \forall \tilde{n}_b \in \{-n_b, n_b\}.$

**Lemma 1.2 Mutually Observable Positions:** For the $p$th BS to have jointly received pilot signals from the same propagation source, from (32) it follows that there must exist some RoA index that satisfies $\tilde{n}_p \in R_{n_b}^{(p)} \exists R_{\tilde{n}_b}^{(p)} \forall \tilde{n}_b \in \{-n_b, n_b\}. That is, there must be some RoA line extending from the $p$th BS that intercepts with the $b$th BS’s RoA, indexed by $\tilde{n}_b \in \{-n_b, n_b\}.$ Furthermore, this intercept has a relative displacement of $(\tilde{x}_b, \tilde{y}_b) \in P_{\tilde{n}_b}^{(p)}$ from the $b$th BS.

**Lemma 1.3 Mutual Orientation Dependency:** In order for this common propagation source to qualify as being a line of sight path from the UE, the orientation of the UE must direct the conditionally considered $n_u$th candidate beam toward the $b$th BS. In other words, it must satisfy the relationship $\phi_b^G = \pi - \theta_b^G$, which leads to the condition that the UE orientation must be one of $\psi_u = \pi - \theta_b^L - \Theta_b - \phi_b^L, \forall \theta_b^L \in \{-\bar{\theta}_{n_b}, \bar{\theta}_{n_b}\}, \phi_b^L \in \{-\bar{\phi}_{n_u}, \bar{\phi}_{n_u}\}.$
**Lemma 1.4** Conditional RoD: For the UE to have jointly a LOS path to the $p$th BS (i.e., from $(\tilde{x}_b, \tilde{y}_b)$ to $(\delta_{x_{p,b}}, \delta_{y_{p,b}})$, relative to the $b$th BS), it must have a global AoD that satisfies $\phi_p^G(\tilde{n}_b, \tilde{n}_p, n_u) = \phi_p^G - \psi_u$ and therefore as elaborated in (37).

With orientation $\psi_u$ from Lemma 1.3, the local AoD can then be expressed as $\phi_p^L(\tilde{n}_b, \tilde{n}_p, n_u) = \phi_p^G - \psi_u$ and therefore as elaborated in (37).

**Lemma 1.5** Conditional Radial Displacement and Path PDF: If the $b$th BS and $p$th BS have jointly LOS paths to a UE positioned at the intercepting point between the RoA pair $\tilde{n}_b$ and $\tilde{n}_p$, then, from (30, 31), the UE radial distance is given by $r_b(\tilde{n}_b, \tilde{n}_p)$ for the $b$th BS and $r_p(\tilde{n}_b, \tilde{n}_p)$ for the $p$th BS. Next, the conditional likelihood function for a $b$th BSs path coefficient can be expressed as

$$f(\hat{\alpha}_b | r_b(\tilde{n}_b, \tilde{n}_p)) = \frac{\exp\left(\frac{-|\hat{\alpha}_b|^2}{\pi(r_p(\tilde{n}_b, \tilde{n}_p)^{-\beta} + \text{Var}[\hat{\alpha}_b])}\right)}{1 + \left(f(\hat{\alpha}_b | r_b(\tilde{n}_b, \tilde{n}_p)) + f(\hat{\alpha}_b | -r_b(\tilde{n}_b, \tilde{n}_p))\right)}$$

and the path coefficient of $b$th BS as

$$f(\hat{\alpha}_p | r_p(\tilde{n}_b, \tilde{n}_p)) = \frac{\exp\left(\frac{-|\hat{\alpha}_p|^2}{\pi(r_p(\tilde{n}_b, \tilde{n}_p)^{-\beta} + \text{Var}[\hat{\alpha}_p])}\right)}{1 + \left(f(\hat{\alpha}_p | r_p(\tilde{n}_b, \tilde{n}_p)) + f(\hat{\alpha}_p | -r_p(\tilde{n}_b, \tilde{n}_p))\right)}$$

**Corollary 1.1:** From (37) we can use Bayes’ rule to obtain the conditional probability of $r_b(\tilde{n}_b, \tilde{n}_p)$ as

$$P_r(r_b(\tilde{n}_b, \tilde{n}_p) | \hat{\alpha}_b) = \frac{f(\hat{\alpha}_b | r_b(\tilde{n}_b, \tilde{n}_p))}{f(\hat{\alpha}_b | r_b(\tilde{n}_b, \tilde{n}_p)) + f(\hat{\alpha}_b | -r_b(\tilde{n}_b, \tilde{n}_p))}$$

and similarly for $P_r(r_p(\tilde{n}_b, \tilde{n}_p) | \hat{\alpha}_p)$.

By considering that the conditional occurrence of events $r_b(\tilde{n}_b, \tilde{n}_p)$ and $r_b(\tilde{n}_b, \tilde{n}_p)$ is, by definition, completely dependent, we can express the probability of their intersection as

$$P_r(\tilde{n}_b, \tilde{n}_p | \hat{\alpha}_b, \hat{\alpha}_p) = P_r(r_b(\tilde{n}_b, \tilde{n}_p) \cap r_p(\tilde{n}_b, \tilde{n}_p) | \hat{\alpha}_b, \hat{\alpha}_p)$$

$$= P_r(r_b(\tilde{n}_b, \tilde{n}_p) | \hat{\alpha}_b)P_r(r_p(\tilde{n}_b, \tilde{n}_p) | \hat{\alpha}_p)$$

and the union among all mutually exclusive solutions $\tilde{n}_p \in R_{\tilde{n}_b}^{(p)}$ that are jointly conditioned...
with \( \hat{n}_b \) as

\[
Pr(\hat{n}_b|\hat{\alpha}_b, \hat{\alpha}_p) = Pr\left( \bigcup_{\tilde{n}_b \in R_{\hat{n}_b}^{(p)}} \hat{n}_b, \tilde{n}_p|\hat{\alpha}_b, \hat{\alpha}_p \right) \frac{1}{|R_{\hat{n}_b}^{(p)}|} \sum_{\tilde{n}_p \in R_{\hat{n}_b}^{(p)}} Pr(\hat{n}_b, \tilde{n}_p|\hat{\alpha}_b, \hat{\alpha}_p)
\]  

(45)

Finally, by substituting the conditional path estimates \( \hat{\alpha}_b = \hat{V}_{n_b,n_u}^{(b)} \) and \( \hat{\alpha}_p = (w_c(\vert \hat{n}_p \vert))^H W_c \hat{V}^{(p)} F^H_{\alpha} a_{UE}(\phi^L_p(\hat{n}_b, \hat{n}_p, n_u)) \) and by approximating the estimation variance as being dominated by the AWGN noise components, i.e., \( \text{Var}[\hat{\alpha}_p] = \text{Var}[\hat{\alpha}_b] = N_0 \), we can consider this probability across each of the equiprobable ranges \( |n_b| = \hat{n}_b \in \{-n_b, n_b\}, |n_u| = \hat{n}_u \in \{-n_u, n_u\} \) to obtain

\[
Pr(n_b, n_u|\hat{V}^{(b)}, \hat{V}^{(p)}) = \frac{1}{4} \sum_{\hat{n}_b \in \{-n_b, n_b\}} \sum_{\hat{n}_u \in \{-n_u, n_u\}} Pr(\hat{n}_b|\hat{\alpha}_b, \hat{\alpha}_p)
\]

(46)

\[
= \sum_{\hat{n}_b \in \{-n_b, n_b\}} \sum_{\hat{n}_u \in \{-n_u, n_u\}} \frac{1}{4|R_{\hat{n}_b}^{(p)}|} Pr(r_p(\hat{n}_b, \hat{n}_p)\big| \hat{V}_{n_b,n_u}^{(b)}) \times
\]

\[
Pr\left( r_p(\hat{n}_b, \hat{n}_p)\big| (w_c(\vert \hat{n}_p \vert))^H W_c \hat{V}^{(p)} F^H_{\alpha} a_{UE}(\phi^L_p(\hat{n}_b, \hat{n}_p, n_u)) \right)
\]

(47)

where \( \phi^L_p(\hat{n}_b, \hat{n}_p, n_u) \) can be found from (37).

**Example 1.1:** An example that illustrates conditional geometry is shown in Fig. 2. Specifically, we show four conditional UE positions and orientations, given the RoA \( \hat{n}_b = 3 \) and \( \hat{n}_b = -3 \) from the right-hand BS. As is the case in our system model, the true UE position is shown in the center (i.e., \( \hat{n}_b = -3 \) in the example), with the correct RoD directions shown as the beamforming directions with bold-solid lines (i.e., \( \hat{n}_u \)). It follows that the correct directions will be expected to correspond to virtual channel entries that exhibit a strong path gain. The dashed and solid lines shown for the other conditional UE positions represent the two possible UE array orientations, each of which results in the considered correct beamforming vector (i.e., the expected strong estimate) directed back toward the right-hand BS at (20,10). Focusing on the UE in the center, it is notable that one of the two array orientations perfectly aligns the correct beamforming direction toward the left-hand BS (i.e., it corresponds to the expected strong measurement in both virtual channels). As such, the probability of the conditional virtual channel entry will be high.

In contrast to the correct estimate, we can consider the alternative conditional UE positions at coordinates (-20,-12.5) and (-18.5,12.5). At these positions, it is evident that neither of two conditional orientations direct the observed strong beamforming directions toward to the true BS positions. As such, these candidate positions will yield low probabilities, as the expected
measurements will not agree with the observed measurements. Focusing on the bottom-right UE position, we see that one of the two orientations does result in the alignment of the strong UE beamforming directions. However, as the resulting RoA from the left-hand BS is now incorrect, it will still correspond to a virtual channel entry with a weak gain and therefore result in low conditionally probability.

**Note 1.1:** It is worth noting that the geometric reasoning in Theorem 1 intentionally considers the scenario that applies to a joint LOS—as asserted through Lemmas (1.2)-(1.4). This set of conditions collectively considers geometric properties that would be consistent with a common line of sight path among two BS. In order to extend this framework to one that jointly estimates NLOS paths, the developed model could also consider common scatterers alongside the already considered LOS components. More specifically, for NLOS, it may also be considered that, for the intercept of two RoA pairs as a common NLOS propagation source (i.e., a scatterer), the RoD from the UE must be the same, or very similar, for both BS estimates. We have left this extension as future work.

**G. RAPID Summary**

Leveraging the expressions from the previous sub-sections, we now give a complete description our RAPID beam training algorithm. We propose that after each BS has collected UE pilot symbols for $T_E$ channel estimation time slots, they each carry out their own channel estimation, before exchanging their estimates with nearby BS. Initially, we assume that this information exchange is made possible by either a wired/wireless front-haul link between each BS. We then propose a bandwidth limited exchange later in this sub-section.

Following (46), after each BS has exchanged its initial set of virtual channel estimates, the $b$th BS can then compute its a priori virtual channel probabilities as

$$P_T(n_b, n_u) = \frac{1}{B-1} \sum_{p=1, p\neq b}^B P_T(n_b, n_u | \hat{V}^{(b)}, \hat{V}^{(p)}).$$

(48)

With this result, each BS can then select those UE/BS candidate beamforming pairs which have greatest probability of a path for data communication. The BS can then feed back the selected UE candidate beamforming indexes, requiring just $\log_2(N)$ per index, for use in the
following data communication\footnote{Alternatively, after the initial $T_E$ estimation time slots, the until-now transmitting UE could instead start receiving with its continued pseudo-random beamforming sequence. As the UE has been simultaneously associated with several BS, all of which know this sequence and now have an estimate of which UE beamforming directions are suitable for communication, this would only require, on average, $N_{IS}/(B \times R_{UE})$ time slots before a high path probability UE beam direction is adopted. This opening could then be used to feed back the UE side information and initiate communication.}

The achievable link rate between the UE and each BS can then be expressed as

$$R_{\text{opt}}^{(u)} = \log_2 | \mathbf{I} + \frac{P}{N_0} \mathbf{W}_d^H \hat{\mathbf{H}} \mathbf{F}_d \hat{\mathbf{H}}^H \mathbf{W}_d |. \quad (49)$$

where $\mathbf{W}_d^H$ and $\mathbf{F}_d$ are BS and UE beamforming matrices consisting of the candidate beamforming vectors selected for communication. The remainder of this section proposes bandwidth-constrained information sharing and the application of RAPID in downlink estimations.

1) Bandwidth Constrained Ray Passing: Due to the difficulties inherent in mmWave communication, and the cost of wired back-haul in dense networks, it is possible that any communication channels between BS links may be bandwidth-constrained. To reduce this overhead, we assume that each BS is only able to share a limited number of entries from a virtual channel estimate.

Fortunately, for any given BS pair, $\hat{\mathbf{V}}^{(b)}$ and $\hat{\mathbf{V}}^{(p)}$, the complete set of entries is not needed, as they do not all have statistical dependencies. More specifically, as the two BS exist in a 2D plane with RoA bounded by positive radial distances, only half of the total number of $|\mathbf{N}_B| = 2N_{BS}$ RoA directions from one BS have any directional component in relation to another BS. As such, the largest number of ROAs that can have a mutual intercept between two BS is $(|\mathbf{N}_B|/2)^2 = N_{BS}^2$. Mathematically, we can denote the entries of the $b$th BS’s virtual channel estimate that are passed to the $p$th BS as

$$\hat{\mathbf{V}}^{(b,p)}_{|\bar{\mathbf{n}}_b|,n_u} \leftarrow \hat{\mathbf{V}}^{(b)}_{|\bar{\mathbf{n}}_b|,n_u}, \forall n_u, \bar{\mathbf{n}}_b \in \mathcal{R}_{n_p}^{(p)}, \forall \bar{\mathbf{n}}_p \in \mathbf{N}$$. \quad (50)

Returning to bandwidth constrained sharing, in some BS deployments as little as half of these $N_{BS}^2$ RoA intercepts correspond to unique entries in the virtual channel matrix. For example, when $\Theta_b = \Theta_p = 0$ and both BS are positioned on the x-axis, the ROAs that are able to have intercepts are half in the positive AoA range and half in the negative AoA. Recalling the ULA beam ambiguity problem (i.e., $\mathbf{a}_{BS}(\theta) = \mathbf{a}_{BS}(-\theta)$), the positive and negative angle ranges correspond to the same entries in the virtual channel matrix, due to the absolute index $|\mathbf{N}_B|$. In this case, only $N_{BS}N_{UE}/2$ entries need to be shared. Conversely, in other cases, such as when $\Theta_b = \Theta_p = 0$ and both BS are positioned on the y-axis, the angular range with
a radial component toward the other BS is either all positive or all negative, and thus all $N_{\text{BS}}N_{\text{UE}}$ entries are statistically relevant, to some extent.

For very large MIMO systems, this may still lead to an undesirable sharing overhead. Fortunately, owing to sparsity of the mmWave channel, many of the estimated virtual channel entries are approximately zero and therefore can be neglected with little loss of performance. As such, we propose that, from the already reduced sets of virtual channel entries in (50), only the $N_d$ most dominant entries are shared. As the geometric representation of the mmWave channel is inherently sparse, this has little effect on the performance of the system, provided $N_d$ is still greater than the number of paths. Furthermore, this also decreases computational complexity, as only RoA pairs of significance need to be considered.

Using this approach, we denote the constrained matrix received by the $b$th BS from the $p$th BS as $\tilde{V}^{(p)}$, such that $||\tilde{V}^{(p)}||_0 = N_d$ as

$$\tilde{V}^{(p)}_{\tilde{n}_b,\tilde{n}_u} \Leftarrow \tilde{V}^{(p)(b)}_{\tilde{n}_p,\tilde{n}_u}, \forall \tilde{n}_u, \tilde{n}_p \in \mathcal{R}_{\tilde{n}_b}, \forall \tilde{n}_b \in \tilde{N}$$  \hspace{1cm} (51)

With this constrained information, we can rewrite (48) as

$$P_r(n_b, n_u) = \frac{1}{B-1} \sum_{p=1, p\neq b}^{B} P_r(n_b, n_u | \tilde{V}^{(b)}, \tilde{V}^{(p)}).$$  \hspace{1cm} (52)

We show the complete RAPID beam training approach in Algorithm 1.

2) **RAPID Downlink Beam Training:** Up to this point, we have introduced RAPID as a cooperative uplink channel estimation strategy; however, by considering that the only prior knowledge required to compute (48) is the relative positions of each network BS and their orientations, RAPID can also be implemented in downlink at the UE. To this end, the UE would only require this static network deployment information, along with a coarse estimate its network position, in order to reduce the number of considered BS. Then, similarly to the uplink description, each BS can broadcast beamformed pilot signals with orthogonal spreading codes while each UE collects the signals with a sequence of beamforming directions. After the $T_E$ estimation time slots, each UE can make a direct estimate of the downlink channel and implement RAPID with no communication overhead for the network.

In this scenario, although there is no sharing overhead, the computational burden that was originally in the up-link and distributed among several BS would now need to be carried out by a single UE. As the UE is expected to have less computational power and more stringent energy requirements, it may still be beneficial for the UE to only consider the $N_d$
Fig. 3 – results where the network consists of $B = 3$ base station (BS), each equipped with $N_{BS} = 32$ antennas and $R_{BS} = 8$ radio frequency (RF) chains and the user is equipped with $N_{UE} = 16$ antennas and $R_{UE} = 4$ RF chains. We assume the expected number of paths is $E[L] = 3$. (a) Shows the average maximum and mean achievable link rates after estimation, while (b) shows the minimum achievable link rate.
Fig. 4 – Numerical results where the network consists of $B = 3$ base station (BS), each equipped with $N_{BS} = 32$ antennas and $R_{BS} = 8$ radio frequency (RF) chains and the user is equipped with $N_{UE} = 16$ antennas and $R_{UE} = 4$ RF chains. We assume the expected number of paths is $E[L] = 3$. (c) shows the cumulative density function (CDF) of network coverage for $P = 0$ dBm and (d) $P = 10$ dBm.
Algorithm 1: Ray-of-Arrival Passing for In-Direct (RAPID) Beam Training.

1. **UE Input**: Each UE has a network-known candidate beamforming pilot sequence $\sqrt{P} \{ s_m \}_{m=1}^{T_E}$.

2. **BS Input**: The orientation, $\Theta_p$, and relative position of nearby BS $\Delta_p$, $\forall p = 1, \ldots, B$. Each BS has $N_d$, $N_0$, $\beta$, $A_b$, and knows $A_m(b) = (s_m^T P_m F_c^*) \otimes (W_m(b) W_c^*)$, $\forall m = 1, \ldots, T_E$.

3. **Initialization**: Each BS pre-computes the RoA intercepts $R_n$.

4. **Ray Passing and Indirect Estimation**:

   - For $m = 1, 2, \ldots, T_E$ do
     - UE Transmits beamformed pilots $x_m = \sqrt{P} F_m s_m$.
     - For $b = 1, 2, \ldots, B$ do
       - The $b$th BS receives with $W_m$ to obtain $y_m(b) = (W_m(b)^T H(b) x_m + q_m(b))$.

5. **Transmission and Independent Estimation**:

   - For $m = 1, 2, \ldots, T_E$ do
     - The $b$th BS as $W_m$ to obtain $y_m(b) = (W_m(b)^T H(b) x_m + q_m(b))$.

6. **Ray Passing and Direct Estimation**:

   - For $b = 1, 2, \ldots, B$ do
     - The $b$th BS uses $y_m(b) = [y_{b1}(b), \ldots, y_{bT_E}(b)]$ and $A_m(b) = [A_{m1}(b), \ldots, A_{mT_E}(b)]$ for independent sparse recovery:
       - $\hat{v}(b) = \arg\min_{v} \|y_m(b) - A_m(b) v\|^2 + \gamma \|v\|^2$.

7. **Output**: $Pr(n_b, n_u)$.

most dominant and dependent entries from each estimate. This downlink estimation strategy would also still permit the UE to adapt its beamforming directions during the estimation process, as proposed in [24].
IV. Numerical Results

We now provide some numerical results to evaluate the performance of our proposed scheme. We consider an mmWave system where a UE is equipped with $N_{UE} = 16$ antennas and $R_{UE} = 4$ RF chains. We assume this UE to be within the range of a network of $B$ BS, each equipped with $N_{BS} = 32$ antennas and $R_{BS} = 8$ RF chains. We adopt a user-centric deployment network model in which each BS is positioned within a 100m x 100m grid, with the UE at its center (i.e., a maximum of 50m away from the UE along the x- or y-axis). We consider each BS to follow a uniform random distribution within this space, while the orientation of each BS also follows a uniformly random distribution in the continuous range $[0, 2\pi]$. Similarly, we consider UE orientation to also follow a uniform random distribution in the range $[0, 2\pi]$. The resulting deployment dependent channels can therefore be found from (10), in which the path loss exponent is considered as $\beta = 4$ to represent severe mmWave propagation losses. We consider each receiver’s noise power to be $N_0 = 10^{-5}$, such that the propagation path signal-to-noise ratio (SNR) can be expressed as $\sigma_b^2 / N_0$, which leads to a minimum link SNR of $r_b^{-\beta} / N_0 = -24$ dB at $\max[r_b] = \sqrt{2} \times 50m$.

We apply RAPID to RDB with $T_E$ channel estimation time slots and an ES-based channel estimation with $T_{ES} = N_{UE}N_{BS}/R_{BS} = 64$ estimation time slots. To compare the performance of each scheme as a result of estimation, we show the average best-available link rate (i.e., the maximum achievable rate given all estimated channels) along with the average of the worst available link rate (i.e., the minimum achievable link rate given all estimated channels). For completeness, we also include the average of all available links. To demonstrate coverage probability and link redundancy, we also show the cumulative density function (CDF) of the UE with $N_{LO}$ link options, whereby we impose the requirement that a link must satisfy $R > R_{th}$ along with a coverage rate threshold $R_{th}$.

In Fig. 3 (a), we show both the maximum achievable link rate and the average achievable link rate for a network of $B = 3$ BS, with RDB using $T_E = 48$. In most cases, it is evident that RDB tends to outperform ES despite using fewer measurement timeslots, particularly at a high SNR, because RDB has greater measurement diversity due to pilots being transmitted with multiple beamforming directions in each time slot. This effectively allows the receiver to sample several entries of the virtual channel at once. Conversely, ES sequentially transmits a pilot with only one beamforming direction at a time and is included as a benchmark approach. By comparing the average link rate of the schemes in Fig. 3 (a), it is notable that those using RAPID show little advantage at a low SNR; however, as transmit power increases, both ES
Fig. 5 – Numerical results where the network consists of $B = 6$ base station (BS), each equipped with $N_{BS} = 32$ antennas and $R_{BS} = 8$ radio frequency (RF) chains and the user is equipped with $N_{UE} = 16$ antennas and $R_{UE} = 4$ RF chains. (a) Shows the average maximum and mean achievable link rates after estimation, while (b) shows the minimum achievable link rate.
Fig. 6 – Numerical results where the network consists of $B = 6$ base station (BS), each equipped with $N_{BS} = 32$ antennas and $R_{BS} = 8$ radio frequency (RF) chains and the user is equipped with $N_{UE} = 16$ antennas and $R_{UE} = 4$ RF chains. (c) shows the cumulative density function (CDF) of network coverage for $P = 0$ dBm and (d) $P = 10$ dBm.
and RDB are able to achieve an average link rate increase of around 1 bit/s/Hz. Interestingly, comparing this to maximum link rate performance, we find that the there is very little increase as SNR increases, because—in most cases—the BS that has the best channel in relation to the UE is the one that stands to gain the least from sharing its information with the other BS. Conversely, in Fig. 3 (b), we show the average of the minimum achievable link rate, which effectively represents the BS that has the worst channel in relation to the UE, which is therefore the BS that stands to gain the most from exchanging information. As the UE transmit power increases, we can see that the minimum link rate of the systems using RAPID increase significantly by up to around 2bps/Hz at a transmit power of 10 dBm.

Turning our attention to Figs. 4 (a) and (b), we show the CDFs for a number of achievable link options, for \( P = 0 \) dBm and \( P = 10 \) dBm, respectively. In both cases, we can see that RAPID is able to increase significantly the probability of having a larger number of available link options, particularly for lower-rate requirement thresholds. This is an inherent property of RAPID’s ability to improve significantly the weaker network links. This low rate threshold region also fits for mmWave systems, as throughput gains are expected to come from large bandwidths as opposed to complex modulation schemes. For much greater rate thresholds we see that the available link probabilities of all systems tend to converge.

In Fig. 5 and Fig. 6, we increase BS density to \( B = 6 \) for the same deployment area. We also reduce the number of RDB time slots to \( T_E = 32 \). In (a) and (b), we again show the minimum, mean, and maximum link rates. Again, we can see from (b) that the minimum link rate is able to increase by around 2 bits/s/Hz by applying RAPID despite the worst of the \( B = 6 \) channels being much worse than \( B = 3 \), as in Fig. 3. Looking at the average link rate in Fig. 5 (b), we can see a more noticeable increase with \( B = 6 \) for the same reason, as now there are many more BS to benefit from the better channels that are shared. Turning to the CDFs in 6 (a) and (b), we find that the probability of having more available links is still much greater with RAPID, in particular at a high SNR.

V. CONCLUSION

In this paper we proposed have a cooperative mmWave beam training scheme in which multiple network BS share information to enhance the channel estimation accuracy of one another—and therefore the network performance as whole. In order to combine shared information, we proposed a Ray-of-Arrival Passing for In-Direct (RAPID) algorithm in which the probability of each directional path can be conditionally considered by multiple BS. By leveraging the derived statistical relations, it was proposed that each BS need only share
information with other BS that are able to utilize it, thus reducing the communication overhead involved in this information sharing. The presented results established that the BS link which has the worst quality benefits the most from the scheme. Furthermore, by considering a minimum rate threshold for communication, we demonstrated that RAPID is able to increase significantly the probability that one or more links are available to a user at any given time.

REFERENCES

[1] X. Ge, S. Tu, G. Mao, C.-X. Wang, and T. Han, “5G ultra-dense cellular networks,” IEEE Trans. Wireless Commun., vol. 23, no. 1, pp. 72–79, 2016.

[2] R. Baldemair, T. Irnich, K. Balachandran, E. Dahlman, G. Mildh, Y. Selén, S. Parkvall, M. Meyer, and A. Osseiran, “Ultra-dense networks in millimeter-wave frequencies,” IEEE Commun. Mag., vol. 53, no. 1, pp. 202–208, 01 2015.

[3] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, “User-centric ultra-dense networks for 5G: challenges, methodologies, and directions,” IEEE Trans. on Wireless Commun., vol. 23, no. 2, pp. 78–85, April 2016.

[4] C. Yang, J. Li, and M. Guizani, “Cooperation for spectral and energy efficiency in ultra-dense small cell networks,” IEEE Trans. on Wireless Commun., vol. 23, no. 1, pp. 64–71, February 2016.

[5] T. Bai, A. Alkhateeb, and R. Heath, “Coverage and capacity of millimeter-wave cellular networks,” IEEE Commun. Mag., vol. 52, no. 9, pp. 70–77, 2014.

[6] M. N. Kulkarni, S. Singh, and J. G. Andrews, “Coverage and rate trends in dense urban mmwave cellular networks,” IEEE GLOBECOM, pp. 3809–3814, 2014.

[7] T. Bai and R. W. Heath, “Coverage and rate analysis for millimeter-wave cellular networks,” IEEE Trans. on Wireless Commun., pp. 1100–1114, 2015.

[8] 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.

[9] C. Dehos, J. L. González, A. De Domenico, D. Kténas, and L. Dussopt, “Millimeter-wave access and backhauling: the solution to the exponential data traffic increase in 5G mobile communications systems?” IEEE Commun. Mag., vol. 52, no. 9, pp. 88–95, 2014.

[10] G. MacCartney and T. Rappaport, “73 GHz millimeter wave propagation measurements for outdoor urban mobile and backhaul communications in New York City,” in IEEE Int. Conf. on Commun. (ICC), June 2014, pp. 4862–4867.

[11] P. Popovski, “Ultra-reliable communication in 5G wireless systems,” in Int. Conf. on 5G for Ubiquitous Connectivity, Nov 2014, pp. 146–151.

[12] T. Levanen, J. Pirskanen, and M. Valkama, “Radio interface design for ultra-low latency millimeter-wave communications in 5G era,” in IEEE GLOBECOM, Dec 2014, pp. 1420–1426.

[13] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, “Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results,” IEEE Commun. Mag., vol. 52, no. 2, pp. 106–113, 2014.

[14] F. Khan, Z. Pi, and S. Rajagopal, “Millimeter-wave mobile broadband with large scale spatial processing for 5G mobile communication,” in 2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Oct 2012, pp. 1517–1523.

[15] H. Shokri-Ghadikolaei, F. Boccardi, E. Erkip, C. Fischione, G. Fodor, M. Kountouris, P. Popovski, and M. Zorzi, “The impact of beamforming and coordination on spectrum pooling in mmwave cellular networks,” in 2016 50th Asilomar Conference on Signals, Systems and Computers, Nov 2016, pp. 21–26.

[16] A. Sayeed, “Deconstructing multiantenna fading channels,” IEEE Trans. Sig. Process., vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
[17] J. Brady, N. Behdad, and A. Sayeed, “Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements,” *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, July 2013.

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

[19] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. Azizur Rahman, R. Funada, F. Kojima, I. Lakkis, H. Harada, and S. Kato, “Beam codebook based beamforming protocol for multi-gbps millimeter-wave WPAN systems,” in *IEEE J. Select. Areas Commun.*, 2009, pp. 1390–1399.

[20] Z. Xiao, T. He, P. Xia, and X.-G. Xia, “Hierarchical codebook design for beamforming training in millimeter-wave communication,” *IEEE Trans. Commun.*, vol. 15, no. 5, pp. 3380–3392, 2016.

[21] M. Kokshoorn, P. Wang, Y. Li, and B. Vucetic, “Fast channel estimation for millimetre wave wireless systems using overlapped beam patterns,” in *IEEE Int. Conf. on Commun. (ICC)*, June 2015, pp. 1304–1309.

[22] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic, “Millimeter wave MIMO channel estimation using overlapped beam patterns and rate adaptation,” *IEEE Trans. Signal Process.*, vol. PP, no. 99, pp. 1–1, 2016.

[23] M. Kokshoorn, H. Chen, Y. Li, and B. Vucetic, “RACE: A rate adaptive channel estimation approach for millimeter wave MIMO systems,” *IEEE GLOBECOM*, pp. 1–6, Dec 2016.

[24] ——, “Beam-on-graph: Simultaneous channel estimation for mmwave MIMO systems with multiple users,” *IEEE Trans. Commun.*, 2018.

[25] ——, “Fountain code-inspired channel estimation for multi-user millimeter wave MIMO systems,” in *IEEE Int. Conf. on Commun. (ICC)*, May 2017, to appear.

[26] A. Alkhateeb, G. Leus, and R. W. Heath, “Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?” in *2015 IEEE ICASSP*, 2015, pp. 2909–2913.

[27] M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, “Multi-connectivity in 5G mmwave cellular networks,” in *Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, June 2016, pp. 1–7.

[28] M. Giordani, M. Mezzavilla, and M. Zorzi, “Initial access in 5G mmwave cellular networks,” *IEEE J. Sel. Areas in Commun.*, vol. 35, no. 9, pp. 2069–2084, Sept 2017.

[29] C. K. Seow and S. Y. Tan, “Non-line-of-sight localization in multipath environments,” *IEEE Trans. on Mobile Comp.*, vol. 7, no. 5, pp. 647–660, May 2008.

[30] H. Deng and A. Sayeed, “Mm-wave MIMO channel modeling and user localization using sparse beamspace signatures,” in *IEEE SPAWC*, June 2014, pp. 130–134.

[31] A. Sayeed and V. Raghavan, “The ideal MIMO channel: Maximizing capacity in sparse multipath with reconfigurable arrays,” in *IEEE Proc. ISIT*, July 2006, pp. 1036–1040.

[32] S. Rangan, “Generalized approximate message passing for estimation with random linear mixing,” in *IEEE ISIT*, 2011, pp. 2168–2172.

[33] J. Vila and P. Schniter, “Expectation-maximization bernoulli-gaussian approximate message passing,” in *IEEE Conf. on Signals, Systems and Computers*, 2011, pp. 799–803.

[34] J. Mo, P. Schniter, N. Gonzalez Prelic, and R. Heath, “Channel estimation in millimeter wave MIMO systems with one-bit quantization,” in *Proc. Conf. Signals, Syst. Comput.*, Nov 2014, pp. 957–961.