A Novel Channel Model and Optimal Beam Tracking Schemes for Mobile Millimeter-Wave Massive MIMO Communications

Joydev Ghosh [ORCID], Huiling Zhu, and Hüseyin Haci [ORCID]

Abstract—A novel channel model has been proposed for mobile millimeter-wave (mmWave) massive multiple-input multiple-output (MIMO) communications to evaluate the effect of end-user mobility. In this model the variance of number of clusters and number of rays generated from each cluster is taken into account that is novel and different from widely used channel models. Two optimum codebook based beam-tracking schemes-multi-objective joint optimization codebook (MJOC) and linear hybrid combiner (LHC)- have been proposed for the novel channel model and their performance for spectral efficiency (SE) is presented. Performance for the two most commonly used channel state information (CSI) estimation approaches is investigated. Finally, the relationship between the beamforming training blocks and optimal beam tracking scheme is presented.

Index Terms—millimeter-wave, MIMO, wireless communication.

I. Introduction

Shifting the wireless communication to millimetre wave (mmWave) from centimeter-wave (cmWave) spectrum can be a promising way to significantly improve the capacity. By decreasing the wavelength of transmitted signals, mmWave technology enables the realization of massive MIMO communications and beamforming technology. Achieving and operating with an accurate channel state information (CSI) is one of the main concerns for a transceiver architecture at mmWave, since it can greatly effect the performance. However, the perfect CSI estimation is becoming more difficult to achieve due to high user mobility and mmWave channel characteristics. In [1] a joint optimization algorithm has been proposed for hybrid precoding in mmWave MIMO systems, where precoding processing is divided into two parts as the baseband and radio frequency (RF) precoding and both processes are jointly optimized as a multi-objective codebook optimization problem. This approach helps the problem to be tractable. In another interesting research [2], at every stage of angle of arrival (AoA)/angle of departure (AoD) estimation, a problem of sparse signal reconstruction should be resolved by applying the convex joint precoder and combiner design (JPCD) scheme. However, this beam tracking scheme is limited to the application of small deviations of the AoAs/AoDs. By exploiting an iterative signal processing approach [3], the JPCD scheme decompose the given non-convex problem into a set of convex subproblems. Even though this scheme can perform effectively in the entire training blocks regime for perfect CSI, it cannot be practical due to low latency requirements of modern systems. Similar challenges to the above described

Joydev Ghosh is with the Department of Software Engineering, Institute of Cybernetics, Tomsk Polytechnic University, 634050 Tomsk, Russia (e-mail: joydev.ghosh.ece@gmail.com).

Huiling Zhu is with the School of Engineering and Digital Arts, University of Kent, Canterbury CT2 7NT, U.K. (e-mail: H.Zhu@kent.ac.uk).

Hüseyin Haci is with the Electric and Electronic Engineering, Near East University, Mersin 10, Nicosia 1403, Turkey (e-mail: huseyin.haci@neu.edu.tr).

can be seen in most of the optimum mean squared deviation (MSD) based hybrid combiner schemes [4]. The mmWave channel is sparse in space and beam misalignment may quickly lead to increased MSD, hence an algorithm called as linear hybrid codebook (LHC) scheme can be a good candidate to employ to align beams since it is shown to perform well in mobile scenarios. Another beam tracking technique that performs well for vehicle to-vehicle (V2V) communications is by [5], but it is not designed for mmWave communication.

Although interesting research, above schemes assume channel models where the number of clusters and number of rays generated from each cluster are fixed despite user mobility. This assumption is not practical. In practice, with the change in users' geographical position there will be a variance in the number of clusters and number of rays generated from each cluster during a communication. In case these characteristics are not considered and handled with an appropriate signal processing at the receiver, the received signal's quality may be significantly degraded. To address this open issue, this paper proposes a channel model that the variance of number of clusters and number of rays generated from each cluster is taken into account. Also, proposes two optimal codebook based beam-tracking schemes such as multi-objective joint optimization codebook (MJOC) and linear hybrid combiner (LHC) for the novel channel model and investigate their spectral efficiency performance to provide seamless connectivity to users.

II. System Model

A mmWave massive MIMO communication system is assumed, where a BS denoted as T_x is equipped with B_{T_x} number of antennas and B_{RF} number of radio frequency (RF) chains. The BS is performing a downlink (DL) communication to a user with B_{R_x} number of antennas and B_{RF} number of RF chains subject to $B_s \leq B_{RF} \leq B_{T_x}$ and $B_s \leq B_{RF} \leq B_{R_x}$, where B_s is data symbol blocks. The total number of subcarriers is represented by C and a data symbol carried at the cth subcarrier is denoted by $s[c]$ for $c = \{1, 2, \ldots, C\}$. At T_x, $s[c] \forall c$ are initially precoded by applying a baseband precoding matrix $F[c]$ with the dimension $B_{RF} \times B_s$, and passed through B_{RF} number of C-point inverse discrete Fourier transform (IDFT) to convert digital symbol blocks into time-domain. The l_c-length cyclic prefix (CP) is then appended to the time domain symbol blocks prior to applying the RF precoding matrix $F_{RF} \in \mathbb{C}^{B_{T_x} \times B_{RF}}$. At R_x, the received signal is initially coupled with an ideal carrier and frequency offset synchronizer by applying an RF combining matrix $G_{RF} \in \mathbb{C}^{B_{R_x} \times B_{RF}}$. Next, time-domain symbol block samples are transformed to the digital symbol blocks by removing the CP and applying C-point dicrete Fourier transform (DFT) together with the combining matrix $G[c] \in \mathbb{C}^{B_{RF} \times B_s}$. Thus, received signal at subcarrier c can be expressed as [1], [6],

$$X_j[c] = G^*[c]G^*_{RF}H_{i,j}[c]F[c]F_{RF}s[c] + G^*[c]G^*_{RF}\omega[c], \quad (1)$$

where $(.)^*$ represents optimum matrices of $G[c]$ and G_{RF} respectively from quantization codebooks to get the best signal power, $H_{i,j}(c)$ denotes the channel matrix for subcarrier c, where i represents the ith antenna element and j represents the jth user as the receiver. The additive white gaussian noise (AWGN) at the cth subcarrier of the receiver is defined by $\omega[c] \sim \mathcal{N}(0, \sigma^2 I)$. $w(c)$ is independent and identically distributed (i.i.d.), and has a zero-mean with σ^2 variance, where σ stands for standard deviation and I indicates the identity matrix.

Suppose the end users are not static and undergo a high speed mobility. The downlink (DL) signal from BS to the receivers will undergo a Doppler shift due to the receiver's mobility. Doppler shift effects have been extensively studied by widely used channel models. However, at massive MIMO systems user mobility will also have considerable effect on the number of clusters and number of rays generated from each cluster for a transmission. For a multi-antenna system with A antenna elements and separation distance of d_2 between two neighboring antenna elements, denote $Z(t)$ as the number of discovered clusters at time t and $z = \{1, 2, \ldots, Z(t)\}$ as the index of the cluster. The Doppler frequency shift due to velocity v in the context of antenna element $a \in \{1, 2, \ldots, A\}$ can be expressed by [7],

$$f(a, t, z) = \frac{f_c v}{c} \cos \theta(a, t, z), \quad (2)$$

where f_c denotes carrier frequency at which signal modulation occurs, c denotes speed of light, $\theta(a, t, z)$ is the angle between arriving wave via the zth cluster at the ath antenna element and the direction of motion, given by (see 3GPP TS36.104) $\cos \theta(a, t, z) = \frac{d_1 + a d_2 + \Delta z - vt}{\sqrt{(d_1 + a d_2 + \Delta z - vt)^2 + d_3^2}} \approx \frac{d_1 + \Delta z - vt}{\sqrt{(d_1 + \Delta z - vt)^2 + d_3^2}}$ for $d_1 >> a d_2$ which shows that $\cos \theta(a, t, z)$ will be almost same for all a in such a case. Δz is the zth cluster's angle of arrival offset with respect to an arriving wave with a single cluster that is taken as a perpendicular wave. $\Delta z \sim \mathbb{U}(1, 2\pi)$, where $\mathbb{U}$ denotes uniform distribution. d_1 denotes distance between the user and the first antenna element in a row, d_3 denotes distance between the user and the BS. Note that at (2), parameters a, t, and z are independent of each other and Doppler frequency shift can be averaged over them one by one. For the mathematical tractability of analysis, without loss of generality, let $\hat{f} = \mathbb{E}[f(a, t, z)]$ be the expected value of Doppler frequency shift with respect to z, where $\mathbb{E}[.]$ is the statistical expectation.

All the antenna elements apart from multi-antenna arrangement system can simultaneously capable of receiving the signal, hence the message signal can be modulated with the sampling period T_s keeping M symbols in a single data frame. Therefore, MIMO channel with Rician fading effect at the mth symbol can be modelled as [8, refer to (4)],

$$\mathcal{H}_{i,j}(m) = \frac{1}{A} \sum_{a=1}^{A} \frac{h_a}{d_a^\alpha} e^{j2\pi \hat{f} m T_s} H_{i,j}(\tau, t), \forall m = \{1, 2..M\}, \quad (3)$$

where $A = B_{T_x}$, α denotes Path-loss exponent, $d_a = d_1 + a d_2$, $\frac{1}{d_a^\alpha}$ denotes path-loss, $h_a = \sqrt{\frac{K}{K+1}} e^{j\varphi_a} + \sqrt{\frac{K}{K+1}} u_a = r_a e^{j\beta_a}$ in which $u_a \sim \mathcal{CN}(0, 1)$ and K denotes shape parameter. r_a follows the Rician distribution with the centre 1 and arbitrary phase shift β_a due to propagation of the signal is uniformly distributed over $[0, 2\pi]$

The probability density function (PDF) of $r_a, \forall a \in [0, N]$ can be expressed as,

$$\rho(r_a, \sigma) = \frac{r_a}{\sigma^2} \sqrt{\frac{K}{K+1}} \exp\left(-\frac{r_a^2 + \sqrt{\frac{K}{K+1}}}{2\sigma^2 \sqrt{\frac{K}{K+1}}}\right). \quad (4)$$

In order to consider the frequency spectrum and scattering cluster properties of mmWave, we introduce a wideband MIMO channel model with Z clusters. Suppose that, there is presence of Y rays between T_x and R_x for each of z clusters alongside a time delay τ_z. Besides, each ray associated with each of z clusters produces horizontal and vertical angle of arrival (AoA), $\{\phi_{R_x,zy}, \vartheta_{R_x,zy}\} \in [0, 2\pi]$, and horizontal and vertical angle of departure (AoD), $\{\phi_{T_x,zy}, \vartheta_{T_x,zy}\} \in [0, 2\pi]$, $\forall z = \{1, 2, \ldots, Z\}$ and $\forall y = \{1, 2, \ldots, Y\}$. We have used the notation τ_y

for the relative time delay for each of y rays. Therefore, the delay-τ channel matrix, denoted by $H_{i,j}(\tau) \in \mathbb{C}^{B_{T_x} \times B_{R_x}}$, can be expressed as, provided by (2) in [9],

$$H_{i,j}(\tau, t) = \frac{1}{\sqrt{Y}} \sum_{z=1}^{Z(t)} \sum_{y=1}^{Y_{Z(t)}} \varrho_y \Gamma(\tau t_s - \tau_z - \tau_y)$$
$$\nu_{R_x}(\phi_{R_x,zy}, \vartheta_{T_x,zy}) \nu_{T_x}^H(\phi_{T_x,zy}, \vartheta_{T_x,zy}), \quad (5)$$

where $(.)^H$ stands for Hermitian transpose operator, ϱ_y indicates the gain of zth cluster, $\Gamma(\tau t_s - \tau_z - \tau_y)$ denotes the raised cosine pulse for t_s-spaced sample signals determined at $(\tau t_s - \tau_z - \tau_y)$, $\nu_{R_x}(\phi_{R_x,z}) \in \mathbb{C}^{B_{R_x} \times 1}$ and $\nu_{T_x}(\phi_{T_x,z}) \in \mathbb{C}^{B_{T_x} \times 1}$ are the array response vectors for R_x and T_x antennas, respectively. At the places where the presence of a baseband signal avoiding outage was discovered by a heuristic process, the poisson distribution of the number of discovered clusters can be expressed as below [10, refer to section 'Cluster Numbers and Hybrid Beamforming Performance'],

$$Z(t) \sim \max\{Poisson(\zeta), 1\}, \quad (6)$$

where ζ set to heuristic mean of $Z(t)$.

The time delay, τ_z, associated with each cluster can be expressed as,

$$\tau_z = -\alpha_\tau \zeta^{-1} \log_e V_z, \quad (7)$$

where τ_z denotes delay scaling variable which signifies the impact on power fraction for delay among the clusters, V_z stand for a uniform random variable $V_z \sim V[0, 1]$.

The cluster z is considered to be reduced or increased in amount of power consumption according to the factor as given below [11],

$$\hat{P}_z = e^{\tau_z \left(\frac{\alpha_\tau - 1}{\alpha_\tau \zeta^{-1}}\right)} 10^{-0.1 K_z}, \text{ where } K_z \sim \mathcal{N}(0, \sigma^2) \quad (8)$$

The ultimate power fragments in individual clusters are served with the use of normalized $\hat{P}_z$ to unity, hence usage of power in the kth cluster can be presented as,

$$P_z = \frac{\hat{P}_z}{\sum_{i=1}^{Z} \hat{P}_z} \quad (9)$$

Now, by using (7), (8) can be re-expressed as,

$$\hat{P}_z = e^{[(1-\alpha_\tau)\log_e V_z]} 10^{-0.1 K_z}$$
$$= e^{\log_e V_z^{(1-\alpha_\tau)}} 10^{-0.1 K_z} \quad (10)$$
$$= V_z^{(1-\alpha_\tau)} 10^{-0.1 K_z}$$

The feasible number of rays that can be generated in the context of each cluster as [11],

$$Y_{Z(t)} \sim \mathbb{D}(1, Z), \quad (11)$$

where $\mathbb{D}$ denotes discrete uniform distribution (DUD)[1] with integers, 1 and Z. By following the properties, PMP and CDF of $Y_{Z(t)}$ can be written as $\frac{1}{Z}$ and $\frac{y}{Z}$.

Based on the MIMO channel response given in (1), the channel matrix between the ith antenna element and the jth user at subcarrier c can be expressed as,

$$H_{i,j}[c] = \sum_{\tau=0}^{l-1} \mathcal{H}_{i,j}(\tau) \exp\left\{-j\left(\frac{2\pi c}{C}\right)\tau\right\}, \quad (12)$$

where $l \leq l_c + 1$ is the lth tap in the mmWave channels.

---

[1]*Properties:* A discrete uniform variable $Y_{Z(t)}$ with integers x_1 and x_2, where $x_1 < x_2$, has the probability mass function (PMF) and the cumulative distribution function (CDF), respectively, as below: $f(y) = \frac{1}{x_2 - x_1 + 1}$, $\forall y = \{x_1, x_1 + 1, \ldots, x_2\}$ $F(y) = P(Y_{Z(t)} \leq y) = \frac{y - x_1 + 1}{x_2 - x_1 + 1}$, $\forall y = \{x_1, x_1 + 1, \ldots, x_2\}$.

A. Perfect CSI (Case Study I)

Under the ideal condition, the ergodic achievable sum-rate of the jth user for perfect CSI (PCSI) can be determined by [12],

$$R_{j,PCSI} = \mathbb{E}\left[\log_2 |I_j + V_j^{-1} Y_{j,j} Y_{j,j}^H|\right] \qquad (13)$$

where $I_j = U_j^H G_{i,j}^H F_{i,j}$, U_j denotes channel independent beam-former at the BS, $F_{i,j}$ denotes DL precoder at the jth user for the ith BS and is given by $F_{i,j} = \bar{G}_{i,j}(\bar{G}_{i,j}^H \bar{G}_{i,j})^{-1} U_j$, $\bar{G}_{i,j} = G_{i,j} + \sum_{\substack{x=1, \\ x\neq i}}^{n} \sqrt{\frac{P_x}{P_i}} G_{x,j} \Psi_x \Psi_i^H$, where $G_{i,j}$ denotes DL combiner at the jth user for the ith BS and $\bar{G}_{i,j}$ denotes the estimation of combining matrix $G_{i,j}$, Ψ_x pilot sequences transmitted from x^{th} BS. $V_j = \sigma^2 U_j^H U_j + \sum_{\substack{x=1, \\ x\neq j}}^{n} Y_{j,x} Y_{j,x}^H, Y_{j,x} = \sum_{i \in \mathcal{M}(x)} \sqrt{p_{x,i}} U_j^H G_{j,i}^H G_{x,i} (G_{x,i}^H G_{x,i})^{-1} U_x$, $p_{x,i}$ denotes the co-efficient that regulates the power transmitted by i^{th} BS to x^{th} user.

B. Imperfect CSI (Case Study II)

Under the non-ideal condition, the ergodic achievable sum-rate of the jth user for imperfect CSI (IPCSI) can be determined by [13],

$$R_{j,IPCSI} = \mathbb{E}\left[\log_2 |I_j + \bar{V}_j^{-1} \bar{Y}_j \bar{Y}_j^H|\right], \qquad (14)$$

where $\bar{Y}_j = \sum_{i \in \mathcal{M}(x)} \sqrt{p_{i,j}} \bar{Y}_{i,j}$, $\bar{Y}_{i,j} = \mathbb{E}[U_j^H G_{i,j}^H F_{i,j}]$, $V_j = \sigma^2 U_j^H U_j + \sum_{i \in \mathcal{M}(x)} \sum_{y \in \mathcal{M}(x)} \sqrt{p_{i,j}} \sqrt{p_{y,j}} M_{i,y,j} + \sum_{\substack{x=1, \\ x\neq j}}^{n} \sum_{i \in \mathcal{M}(x)} \sum_{y \in \mathcal{M}(x)} \sqrt{p_{i,x}} \sqrt{p_{y,x}} N_{i,j,x,y}$, $M_{i,y,j} = \mathbb{E}[(U_j^H G_{i,j}^H F_{i,j} - \mathbb{E}[U_j^H G_{i,j}^H F_{i,j}]) \times (U_j^H G_{y,j}^H F_{y,j} - \mathbb{E}[U_j^H G_{y,j}^H F_{y,j}])^H]$, $N_{i,j,x,y} = [U_j^H G_{i,j}^H F_{i,j} U_j G_{y,j} F_{i,x}^H]$.

III. CODEBOOK-BASED BEAM TRACKING SCHEMES:

In beamspace based massive MIMO communications, user-equipments' (UEs') locations in the orthogonal subspace is the main criteria to achieve the optimum response. In this section, two code-book based beam tracking schemes are proposed for the novel channel model.

A. Multiobjective Joint Optimization Codebook (MJOC) Scheme

MJOC scheme is performing a joint optimization of baseband pre-coding matrix $F[c]$ and RF precoding matrix $F_{RF}[c]$ subject to sum-rate and RF chain constraints. Multiobjective joint optimization with constraints are challenging problems that may need a number of problem transformations to become computationally feasible. To find a solution for MJOC, let us introduce three discrete relaxation parameters by the notations μ_j, δ_j and Φ_j with the assumption that $\gamma_j \geq \mu_j$, $R_j \geq \delta_j$ and Φ_j refers to the jth column of $F[c] \in \mathbb{C}^{B_{Tx} \times n}$, $j = \{1, \ldots, J\}$, where J represents the total number of users. $\gamma_j = 2^{2B_{Rx}}$ represents the highest achievable signal to noise ratio (SNR) given that the rank of the channel is not less than B_{Rx}. The approach using tailored big-M constraints [14, refer to sections 5.4 and 6.2.2] are employed to solve the discrete-continuous optimization problem. By following the upper bound big-M constraint, the optimal codebook design problem with limited sum-rate constraint can be formulated as a continuous relaxed parameter problem as,

$$\max_{\mu_j, \delta_j, \Phi_j} \sum_{j=1}^{J} \delta_j R_j, \qquad (15)$$

Subject to: $1 + \mu_j \geq e^{\delta_j}, \forall j,$ \qquad (15a)

$\gamma_j \geq \mu_j, \gamma_j \geq \hat{\delta}_j, \forall j,$ \qquad (15b)

$$\sum_{j=1}^{J} ||\Phi_j F_{RF}||_2^2 \leq P_{max}, ||\hat{\Phi}||_0 \leq B_{RF}, \qquad (15c)$$

where $\hat{\delta}_j = e^{\delta_j} - 1$, (15a) is a convex condition on $R_j{}^2$, P_{max} denotes the maximum feasible transmit power, $||\hat{\Phi}||_0 \leq B_{RF}$ denotes L_0-norm [3] constraint for RF chain that signifies the preferred codewords can not be higher than the available RF chains.

In order to deal with non-convex conditions given in (15b) and (15c), L_0-norm constraints are initially transformed into the following objective function,

$$\max_{\mu_j, \delta_j, \Phi_j} \sum_{j=1}^{J} \delta_j R_j + \Xi ||\hat{\Phi}||_0, \qquad (16)$$

Subject to: (15b) and $\sum_{j=1}^{J} ||\Phi_j F_{RF}||_2^2 \leq P_{max}$, where Ξ is applied as a co-efficient of L_0-norm to ensure the control over a sparse solution and satisfy the L_0-norm condition in subsequent steps.

Further to deal with inducing group-sparsity and quality sparse signal recovery from L_0-norm minimization problem, the square of convex L_∞-norm can be viewed as a relaxation of a non-convex L_0-norm mini-mization problem without changing the conditional properties. Besides, L_∞-norm offers sparse signal reconstruction and sparse solution from a few measurements.

Next, the non-convex L_0-norm problem can be re-expressed as a square of convex L_∞-norm problem as follows,

$$\max_{\mu_j, \delta_j, \Phi_j} \sum_{j=1}^{J} \delta_j R_j + \Xi ||F||_\infty^2, \qquad (17)$$

Subject to: (15b) and $\sum_{j=1}^{J} ||\Phi_j F_{RF}||_2^2 \leq P_{max}$.

Further, if the vector norm is a L_p-norm, then the induced matrix norm should satisfy the below property:

$$||F||_\infty = \sum_{i=1}^{B_{Tx}} \max_j |\Phi_j(i)|. \qquad (18)$$

Therefore,

$$||F||_\infty^2 = \left(\sum_{i=1}^{B_{Tx}} \max_j |\Phi_j(i)|\right)^2$$

$$= \sum_{i_1=1}^{B_{Tx}} \sum_{i_2=1}^{B_{Tx}} \left(\left(\max_{j \in \{1,\ldots,J\}} |\Phi_j(i_1)|\right) \right.$$

$$times \left.\left(\max_{j \in \{1,\ldots,J\}} |\Phi_j(i_2)|\right)\right)$$

$$= \sum_{x=1}^{B_{Tx}} \sum_{y=1}^{B_{Tx}} \max_{s_1,s_2 \in \{1,\ldots,J\}} |Q_{s_1,s_2}(x,y)|, \qquad (19)$$

---

[2] e^{ax} is convex on objective function $\mathbb{R}$, for any $a \in \mathbb{R}$ [15].
[3] L_p norm: $||\hat{\Phi}||_p \leq B_{RF}, p \geq 1$.

where $Q_{s_1,s_2} = \Phi_{s_1}\Phi_{s_2}^H, \forall\{s_1,s_2\}$ for $Q_{s_1,s_2} \succeq 0$ and rank $(Q_{s_1,s_2}) = 1, \forall\{s_1,s_2\}$.

Hence, (17) can be re-expressed as,

$$\max_{\mu_j,\delta_j,Q_{s_1,s_2}} \sum_{j=1}^{J} \delta_j R_j + \Xi ||F||_\infty^2, \qquad (20)$$

Subject to: $1 + \mu_j \geq e^{\delta_j}, \sum_{j=1}^{J} tr\left(\hat{F}_{RF}Q_{j,j}\right) \leq P_{max}.$ (20a)

$$\gamma_j \geq \mu_j, \gamma_j \geq \hat{\delta}_j, \forall j, Q_{j,j} \succeq 0 \text{ and rank }(Q_{s_1,s_2}) = 1,$$
$$\forall s_1, \forall s_2, \qquad (20b)$$

where $\hat{F}_{RF} = F_{RF}^H F_{RF}$. Substituting back into (1), it can be re-expressed as,

$$\hat{X}_j = \hat{G}^*\hat{G}_{RF}^*\hat{H}_{i,j}\hat{Q}_{j,j} + \hat{G}^*\hat{G}_{RF}^*\hat{\omega}. \qquad (21)$$

In order to improve the performance at the short training blocks regime keeping the target of lower-complexity receivers design in mind, we require a hybrid combiner with low mean squared deviation (MSD) between the transmitted and received signals.

It has been shown by [16] that MJOC's performance may be limited when the number of beamforming vectors at T_x and/or combining vectors at R_x is limited. That is the reason why we propose the LHC scheme next, as an ideal scheme for such scenarios.

B. Linear Hybrid Combiner (LHC) Scheme

We propose LHC scheme by following the MSD minimization as an optimization problem with a constant $F[c]F_{RF}$ and weight matrix ξ, and is given as,

$$(G^*, G_{RF}^*)$$
$$= \arg\min_{\substack{G\in\mathbb{C}^{B_{RF}\times B_s}, \\ G_{RF}\in\mathbb{C}^{B_{Rx}\times B_{RF}}}} \left[tr\{\xi\mathbb{E}\left[||s - G^H G_{RF}^H X||_2^2\right]\} - \log_2\xi\right].$$
$$(22)$$

By [17], the solution of (22) can be expressed as,

$$G_{MSD}^H = \mathbb{E}[sX^H]\mathbb{E}[XX^H]^{-1}$$
$$= \frac{1}{B_s}F^H F_{RF}^H H^H$$
$$\times \left(\frac{1}{B_s}FF_{RF}HF^H F_{RF}^H H^H + \sigma^2 I\right)^{-1}$$
$$= F^H F_{RF}^H H^H \left(FF_{RF}HF^H F_{RF}^H H^H + B_s\sigma^2 I\right)^{-1}. \quad (23)$$

Since this solution cannot be break down into a product of G^H and G_{RF}^H, we should further restate the problem of (23) by following the methodology applied in [18], given as,

$$\mathbb{E}\left[||s - G^H G_{RF}^H X||_2^2\right]$$
$$= \mathbb{E}\left[(s - G^H G_{RF}^H X)^H(s - G^H G_{RF}^H X)\right]$$
$$= \mathbb{E}\left[tr((s - G^H G_{RF}^H X)^H(s - G^H G_{RF}^H X))\right] \qquad (24)$$
$$= tr(\mathbb{E}[ss^H]) - 2\Re\{tr(\mathbb{E}[sX^H]GG_{RF})\}$$
$$+ tr(G^H G_{RF}^H \mathbb{E}[XX^H]GG_{RF}).$$

Now, by using (23) and Theorem 8.35 of [19], (24) can be re-expressed as,

$$\chi(G, G_{RF})$$
$$= tr\left(G_{MSD}^H\mathbb{E}[XX^H]G_{MSD}\right) - 2\Re\{tr(\mathbb{E}[sX^H]GG_{RF})\}$$
$$+ tr\left(G^H G_{RF}^H\mathbb{E}[XX^H]GG_{RF}\right)$$
$$= tr\left(G_{MSD}^H\mathbb{E}[XX^H]G_{MSD}\right)$$
$$- 2\Re\left\{tr(G_{MSD}^H\mathbb{E}[XX^H]GG_{RF})\right\}$$
$$+ tr\left(G^H G_{RF}^H\mathbb{E}[XX^H]GG_{RF}\right)$$
$$= tr\left(\left(G_{MSD}^H - G^H G_{RF}^H\right)\mathbb{E}[XX^H]\right.$$
$$\left.\times (G_{MSD}^H - G^H G_{RF}^H)^H\right)$$
$$= ||\mathbb{E}[XX^H]^{\frac{1}{2}}(G_{MSD}^H - G^H G_{RF}^H)||_F^2. \qquad (25)$$

Hence, the minimum MSD problem can be expressed as,

$$(G^*, G_{RF}^*) = \arg\min_{\substack{G\in\mathbb{C}^{B_{RF}\times B_s}, \\ G_{RF}\mathbb{C}^{B_{Rx}\times B_{RF}}}} \left[tr\left\{\xi||\mathbb{E}\left[XX^H\right]^{\frac{1}{2}}\right.\right.$$
$$\left.\left.\left(G_{MSD}^H - G^H G_{RF}^H\right)||_F^2\right\} - \log_2\xi\right], \qquad (26)$$

where $||.||_F^2$ denotes the standard Frobenius norm and (26) indicates to observe the prediction of G_{MSD} without a limitation over the cardinality of $G^H G_{RF}^H$ for $G_{RF} \in \mathbb{C}^{B_{Rx}\times B_{RF}}$. The MSD estimation problem is a joint estimation problem of multiple signal realizations, especially when the signal samples have a joint sparse support over a given constraint. Hence, address the problem by the product of F and F_{RF} as an optimal precoder and determine GG_{RF}. The sparse signal approximation for the best matching projection is given by Algorithm 1. The following optimization problem presents an algorithmic solution,

$$\min_{\tilde{G}} ||\sqrt{\xi}(G_{MSD} - \mathbb{E}\left[XX^H\right]G_{RF}\tilde{G})||_F^2, \qquad (27)$$

subject to: $||diag(\tilde{G}\tilde{G}^*)||_0 = B_{RF}$ and $||\nu_{Rx}^H\tilde{G}||_F^2 = P_{max}$.

IV. SIMULATION RESULTS AND DISCUSSION

In this section, numerical results are provided for the developed channel-model and beam tracking schemes. System performance has been evaluated based on the metrics given by (13) and (14) for the two common practices in channel estimation. The aim at both studies in the context of CSI is focus on the insightful performance comparison of ergodic achievable sum-rate maximization for benchmarking. CSI on a subcarrier with respect to each user can be estimated periodically through the pilot sequence (periodic reference signal). Discrete Fourier transform and discrete cosine transform based methods are popular family of channel estimation methods for multi-carrier systems [20]. [20] have shown that by employing the least squares and linear minimum mean square error (MMSE) techniques estimates for the CSI can be achieved with mean square error about 10^-6 for $SNR \geq 25$ dB. [21] proposed an adaptive estimation method to predict users' CSI for varying user mobility profiles (including velocity). The method is based on the concept of Gaussian process regression and a dual-control technique is used to determine the most appropriate prediction time to keep the packet loss rate below a pre-defined threshold with the given

Algorithm 1: Sparse Signal Approximation for Best Matching Projection.

1. Function: Sparse
   $\_T_X\_R_X(\nu_{R_x}, X, G_{MSD}, B_{RF}, P_{max})$
2. Inputs: ν_{R_x}, X, G_{MSD}, B_{RF}, and P_{max} (If applicable)
3. Initialization:
   $G_{RF} \leftarrow$ Empty Matrix, $G_{tem} \leftarrow G_{MSD}$
4. **while** $l \leq B_{RF}$ **do**
   | 5. $\varpi = \nu_{R_x}^H \mathbb{E}[XX^H]G_{tem}$
   | 6. $d = \arg\max_{\substack{z=\{1,\cdots,Z\} \\ y=\{1,\cdots,Y\}}} (\varpi\varpi^H)_{z,y}$
   | 7. $G_{RF} = [G_{RF}|\nu_{R_x}^d]$
   | 8.
   | $G = (G_{RF}^H \mathbb{E}[XX^H]G_{RF})^{-1}G_{RF}^H \mathbb{E}[XX^H]G_{MSD}$
   | 9. $G_{tem} = \frac{\sqrt{\xi}(G_{MSD} - \mathbb{E}[XX^H]G_{RF}G)}{||\sqrt{\xi}(G_{MSD} - \mathbb{E}[XX^H]G_{RF}G)||_F^2}$
10. Normalize $G = \sqrt{P_{max}}\frac{G}{||G_{RF}G||_F^2}$, If the power constraint is active.
11. Outputs: G_{RF}, G

TABLE I
THE MAIN SIMULATION PARAMETERS

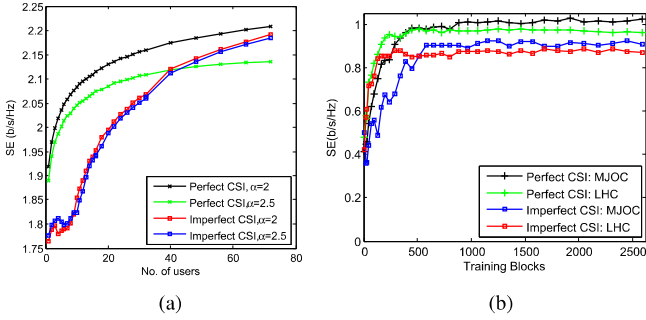| Parameter	Value	Parameter	Value
User's velocity	20m/s	Distance from BS to users	40m
Data rate	133 Mbps	Number of BSs	32
Wavelength	0.15m	Channel coherence	64 blocks
Doppler frequency	133Hz	Normalized doppler frequency	10^{-6}

(a)   (b)

Fig. 1. Figure 1: (a) SE performance versus the number of users at the system for PCSI and IPCSI cases with $\alpha \in \{2, 2.5\}$. (b) SE performance versus training blocks (i.e. number of beamforming vectors at Tx × number of combining vectors at Rx) for MJOC and LHC schemes.

user mobility. It has been shown that packet loss less than 1% can be achieved even with 60 km/hr velocity given that appropriate prediction time is set. Table I provides the details about the simulation parameters. As CSI is a crucial aspect of reliable high data-rate communications in MIMO channel, perfect and imperfect CSI, respectively, at T_x and R_x are assumed at the simulations.

Fig. 1(a) shows the spectral efficiency (SE) performance versus the total number of users for PCSI and IPCSI cases and path loss exponent $\alpha \in \{2, 2.5\}$ to simulate different level of signal degradation due to imperfect CSI estimation and increase in α. The aim of simulations is to investigate if one's effect can dominate the other. SE in this paper is defined as the uncoded bit rate per unit bandwidth. It can be seen

from the figure that the performance of both channel estimation cases for different values of α increases as the number of users and this is due to multi-user gain. The behavior of the curves corresponding to $\alpha = 2$ and $\alpha = 2.5$ for PCSI are similar with a large difference in performance as the number of users gets large. This shows that α has a significant effect on the performance since PCSI can be realized and this effect increases as the network gets densely populated. Plots for IPCSI with $\alpha = 2$ and $\alpha = 2.5$ show a similar response with marginal difference in performance for all the considered cases of number of users. This shows that α (for the considered values) do not have a significant effect on performance when IPCSI case is considered. The reason is that channel estimation error dominates the performance degradation due to increased value of α.

Fig. 1(b) shows the SE performance versus number of training blocks for MJOC and LHC schemes for the cases in which PCSI and IPCSI are applied. The performance for both schemes and CSI cases increases with respect to the increasing length of training blocks. Slight variations can be seen at the performances due to the assumptions/relaxation in optimization problem formulation of baseband matrices and MSD minimization. Performance variations of MJOC scheme is shown to be higher with respect to LHC, especially at the case of IPCSI and short training blocks. The main reason is that compared to LHC, joint optimization of baseband and RF precoding matrices makes MJOC more sensitive to the channel estimation error due to IPCSI and may cause more frequent deviations from the optimal solution. With the increase of the training blocks size to a medium and/or large length (e.g. 600 and higher), the impact of channel estimation error becomes insignificant and the variation of MJOC's performance reduce to a level similar to LHC. For the short training blocks regime (e.g. up to around 400), LHC scheme outperforms MJOC for both IPCSI and PCSI cases. For the medium to large training block length, it can be seen that MJOC scheme outperforms LHC scheme for the PCSI and IPCSI cases. Thus, it can be seen from the figure that the number of training blocks has a significant effect on the choice of beam-tracking scheme to be employed to achieve the best SE and interestingly the stability of the performance of schemes.

V. CONCLUSION

In this paper, a novel channel model for mmWave communications has been proposed to incorporate varying number of clusters and number of rays generated from each cluster due to user mobility. Two optimal beam-tracking schemes -MJOC and LHC - have been proposed for the novel channel model that exploits the channel correlation information among UEs and optimize the SE. Simulation results are presented to investigate the performance of two most common practices in CSI estimation, the PCSI and IPCSI, and effect of training block's length on the choice of beam-tracking schemes to achieve the best SE performance. It is shown that the LHC scheme is suitable when training block is short, while MJOC scheme is preferable for medium to long training blocks.

REFERENCES

[1] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.
[2] X. Chen, J. Lu, P. Fan, and K. B. Letaief,"Massive MIMO beamforming with transmit diversity for high mobility wireless communications," *IEEE Access*, vol. 5, pp. 23032–23045, 2017.
[3] B. Pierre, M. Kilinc, and J. Linderoth, "Algorithms and software for convex mixed integer nonlinear programs," *Mixed Integer Nonlinear Program.*, vol. 154, pp. 1–39, 2012.

[4] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no 5, pp. 3311–3326, May 2016.

[5] C. Perfecto, J. D. Ser, and M. Bennis, "Millimeter-wave V2V communications: Distributed association and beam alignment," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2148–2162, Sep. 2017.

[6] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432–1443, Jul. 2017.

[7] L. Liu *et al.*, "Position-based modeling for wireless channel on high-speed railway under a viaduct at 2.35 GHz," *IEEE J. Sel. Areas Commun.*, vol. 30, no 4, pp. 834–845, May 2012.

[8] R. W. Heath, N. G. Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no 3, pp. 436–453, Apr. 2016.

[9] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, 2014, pp. 273–277

[10] M. Shafi *et al.*, "Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 14–20, Dec. 2018.

[11] 3GPP TR 25.996 V11.0.0, "Spatial channel model for Multiple Input Multiple Output (MIMO) simulations, Sep. 2012.

[12] G. Caire, "On the ergodic rate lower bounds with applications to massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no 5, pp. 3258–3268, May 2018.

[13] A. Mukherjee and A. L, Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, Jan. 2011.

[14] O. Stein, J. Oldenburg, and W. Marquardt, "Continuous reformulations of discrete-continuous optimization problems," *Comput. Chem. Eng.*, vol. 28, pp. 1951–1966, 2004.

[15] *Convex Optimization:convex Functions—Boyd & Vandenberghe*, [Online]. Available: https://see.stanford.edu/materials/lsocoee364a/03ConvexFunctions.pdf

[16] O. Mehanna, N. Sidiropoulos, and G. Giannakis, "Joint multicast beamforming and antenna selection," *IEEE Trans. Sig. Proc.*, vol. 61, no. 10, pp. 2660–2674, May 2013.

[17] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, London, 2000, vol. 1.

[18] T. Michaeli and Y. Eldar, "Constrained nonlinear minimum MSE estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2008, pp. 3681–3684.

[19] G. Meurant, "Preconditioning," in *Studies in Mathematics and Its Applications.* Amsterdam, The Netherlands: Elsevier, pp. 397–540, 1999, doi: 10.1016/s0168-2024(99)80009-5.

[20] C. Chen and D. W. Lin, "Channel estimation for LTE and LTE-A MU-MIMO uplink with a narrow transmission band," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Florence, Italy, 2014, pp. 6484–6488.

[21] A. Chiumento *et al.* "Adaptive CSI and feedback estimation in LTE and beyond: A Gaussian process regression approach," *J. Wireless Commun. Netw.* vol. 168, no. 1, pp. 1–14, 2015.