Abstract—Hybrid beamforming is a promising concept to achieve high data rate transmission at millimeter waves. To implement it in a transceiver, many references optimally adapt to a high-dimensional multi-antenna channel but more or less ignore the complexity of the channel estimation. Realizing that received coupling coefficients of the channel and pairs of possible analog beamforming vectors can be used for analog beam selection, we further propose a low-complexity scheme that exploits the coupling coefficients to implement hybrid beamforming. Essentially, the coupling coefficients can be regarded as implicit channel state information (CSI), and the estimates of these coupling coefficients yield alternatives of effective channel matrices of much lower dimension. After calculating the Frobenius norm of these effective channel matrices, it turns out that the effective channel having the largest value of the Frobenius norm provides the solution to hybrid beamforming problem.

I. INTRODUCTION

With the rapid increase of data rates in wireless communications, bandwidth shortage is getting more critical. Therefore, there is a growing interest in using millimeter wave (mmWave) for future wireless communications taking advantage of the enormous amount of available spectrum [1]. When systems operate at mmWave frequency bands, a combination of analog beamforming [2], [3] and digital beamforming [4] can be one of the low-cost solutions, which is commonly called hybrid beamforming [5]-[8]. Unquestionably, it is intractable to deal with hybrid beamforming at a transmitter and a receiver simultaneously because there are four unknown matrices (two analog and two digital beamforming matrices). To avoid dealing with all these issues at the same time, one can simplify the problem by initially assuming that the channel matrix is known. Then the problem of hybrid beamforming on both sides (i.e., finding the precoder and combiner) can be solved by utilizing the singular value decomposition (SVD) of the channel matrix [5]-[8]. Unfortunately, the overhead of some preliminary work, such as channel estimation for large-scale antenna arrays [9], [10], makes the previously proposed solutions difficult to obtain.

Our previous work in [11] explains that the analog beam selection based on the power estimates is equivalent to the beam selection by the orthogonal matching pursuit (OMP) algorithm [12] when the analog beamforming vectors are selected from orthogonal codebooks. However, the result of the analog beam selection technique based on the received power can be further improved because the factor dominating the performance of hybrid beamforming is the singular values of the effective channel $H_E$ rather than the received power. A simple cure for the problem is that one can reserve a few more candidates of the analog beamforming vectors associated with the large received power levels. Then, the subset of these candidates yielding maximum throughput will provide the optimal solution to the hybrid beamforming. Again it is evident that the computational complexity exponentially increases as the size of the enlarged candidate set. Consequently, we have a strong motivation to find a relationship between the observations for the analog beam selection and a key parameter of the hybrid beamforming gain, and then use the relationship to facilitate the analog beam selection and the corresponding optimal digital beamforming. In the low SNR regime, we find that the Frobenius norm of $H_E$ is the key parameter of the hybrid beamforming gain. Moreover, the observations for the analog beam selection can be used to generate $H_E$. Accordingly, the collected observations yielding the maximum value of the key parameter gives us the necessary information to optimally select the analog as well as the corresponding digital beamforming matrices.

We use the following notations throughout this paper. $a$ is a scalar, $a$ is a column vector, and $A$ is a matrix. $a_n$ denotes the $n^{th}$ column vector of $A$: $[A]_{n,:}$ denotes the $n^{th}$ diagonal element of $A$: $[A]_{1:N,:}$ denotes the first $N$ column vectors of $A$: $[A]_{1:N,1:N}$ denotes the $N \times N$ submatrix extracted from the upper-left corner of $A$. $A^H$, and $A^T$ denote the complex conjugate, Hermitian transpose and transpose of $A$. $\|A\|_F$ denotes the Frobenius norm of $A$. $I_N$ and $O_{N\times M}$ denote the $N \times N$ identity and $N \times M$ zero matrices.

II. SYSTEM MODEL

A system having a transmitter with an $N_T$-element uniform linear antenna array (ULA) communicate $N_S$ OFDM data streams to a receiver with an $N_R$-element ULA as shown in Fig. 1. At the transmitter, a precoder $F_P F_B[k]$ at each subcarrier $k = 1, \cdots, K$ includes a analog beamforming matrix $F_P \in \mathbb{C}^{N_T \times N_{RF}}$ and a digital beamforming matrix $F_B[k] \in \mathbb{C}^{N_{RF} \times N_S}$. The $N_{RF}$ analog beamforming vectors in matrix $F_P$ are selected from a predefined codebook $\mathcal{F}_P = \{f_{n_f} \in \mathbb{C}^{N_T \times 1} | n_f = 1, \cdots, N_F \}$ with the $n_f^{th}$ member represented as [3]

$$f_{n_f} = \frac{1}{\sqrt{N_T}} \left[ 1, e^{j \frac{2\pi}{N_T} \sin(\phi_{n_f}) \Delta_d}, \cdots, e^{j \frac{2\pi}{N_T} \sin(\phi_{n_f}) (N_T-1) \Delta_d} \right]^T,$$

(1)
where $\phi_{T,n_f}$ stands for the $n_f^\text{th}$ candidate of the steering angles at the transmitter, $\Delta_d = \lambda_0/2$ is the distance between two neighboring antennas, and $\lambda_0$ is the wavelength at the carrier frequency. At the receiver, the combiner $W_p W_B[k]$ has a similar structure as the precoder. $W_p \in \mathbb{C}^{N_R \times N_{RF}}$ and $W_B[k] \in \mathbb{C}^{N_{RF} \times N_s}$ are the analog and digital beamforming matrices, respectively. Also, the columns of $W_p$ are selected from the other codebook $\mathcal{V} = \{\mathbf{w}_{n_{w}c} \in \mathbb{C}^{N_{RF} \times 1} : n_w = 1, \cdots, N_W\}$, where the members can be generated by the same rule as (1). Due to hardware constraints, the analog beamforming matrices are constant within one OFDM symbol duration.

Via a coupling of the precoder, combiner, and a frequency-selective fading channel $\mathbf{H}[k] \in \mathbb{C}^{N_{RF} \times N_T}$, the received signal $r[k] \in \mathbb{C}^{N_s \times 1}$ at subcarrier $k$ can be written as

$$r[k] = \sqrt{\rho} \cdot W_B^H[k] W_B^H \mathbf{H}[k] \mathbf{F}_p \mathbf{F}_B[k] \mathbf{s}[k] + W_B^H[k] W_B^H \mathbf{n}[k] = \sqrt{\rho} \cdot W_B^H[k] W_B^H \mathbf{H}[k] \mathbf{x}[k] + \mathbf{z}[k],$$

where $\rho$ stands for the average received power including the transmit power, transmit antenna gain, receive antenna gain, and path loss, $\mathbf{s}[k] \in \mathbb{C}^{N_{RF} \times 1}$ is the transmitted pilot or data vector whose covariance matrix $\mathbf{R}_s = \mathbb{E}[\mathbf{s}[k] \mathbf{s}^H[k]]$ is a diagonal matrix, and $\mathbf{n}[k] \in \mathbb{C}^{N_{RF} \times 1}$ is an $N_0$-dimensional independent and identically distributed (i.i.d.) complex Gaussian random vector, i.e., $\mathbf{n}[k] \sim \mathcal{CN}(0, N_0)$. Furthermore, the precoded transmitted signal $\mathbf{x}[k]$ and combined noise vector $\mathbf{z}[k]$ are enforced to satisfy the following two conditions: constant transmit power at each subcarrier and that the combined noise vector $\mathbf{z}[k]$ remains i.i.d.,

$$\begin{align*}
\mathbf{R}_x &= \mathbb{E}[\mathbf{F}_p \mathbf{F}_B[k] \mathbf{R}_s \mathbf{F}_B^H[k] \mathbf{F}_p^H] = \mathbf{R}_s, \\
\mathbf{R}_z &= \sigma_n^2 W_B^H[k] W_B^H \mathbf{R}_p W_B[k] = \sigma_n^2 \mathbf{I}_{N_s},
\end{align*}$$

which can also be regarded as power constraints on the precoder and combiner.

The properties of mmWave channels have been widely studied recently [13], [14]. Based on the references, a simplified cluster-based frequency-selective fading channel has $C$ clusters and $R$ rays of each cluster. At subcarrier $k$, the channel matrix can be written as

$$\mathbf{H}[k] = \sum_{c=1}^{C} \sum_{r=1}^{R} \alpha_{c,r} e^{-j 2\pi k l_{c,r} / \lambda} \mathbf{a}_A(\phi_{A,c,r}) \mathbf{a}_D(\phi_{D,c,r})^H,$$

where the channel characteristics are given by the following parameters: $\alpha_{c,r} \in \mathbb{R}_{>0}$ describes the inter- and intra-cluster path power and $\sum_{c=1}^{C} \sum_{r=1}^{R} |\alpha_{c,r}|^2 = 1$. $l_{c,r} \in \mathbb{R}_{>0}$ stands for the delay index measured in unit of the sampling interval. Intra-cluster angle of departure (AoD) $\phi_{D,c,r} = \phi_{D,c} + \epsilon_{D,c}$, where the mean $\phi_{D,c} \sim \mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2})$, and the other two factors (the angle spread $\epsilon_D$ and the offset angle $\Delta_{\phi}$) are given in [14] Table 7.5-3 and Table 7.5-6. In the same way, one can generate intra-cluster angle of arrival (AoA) $\phi_{A,c,r}$. The departure array response vector $\mathbf{a}_D(\phi_{D,c,r})$ has entries of equal magnitude and is a function of AoD represented as

$$\mathbf{a}_D(\phi_{D,c,r}) = \frac{1}{\sqrt{N_T}} \left[ 1, e^{j \frac{2 \pi}{\lambda_0} \sin(\phi_{D,c,r}) \Delta_d}, \cdots, e^{j \frac{2 \pi}{\lambda_0} \sin(\phi_{D,c,r}) (N_T - 1) \Delta_d} \right]^T,$$

and the arrival array response vector $\mathbf{a}_A(\phi_{A,c,r})$ has a similar form as (5).

### III. Problem Statement

The objective of the precoder $\mathbf{F}_p \mathbf{F}_B[k] \forall k$ and the associated combiner $\mathbf{W}_p \mathbf{W}_B[k] \forall k$ is to achieve the maximum throughput across all subcarriers subject to the power constraints. That is, we seek matrices that solve

$$\begin{align*}
\max_{\mathbf{F}_p, \mathbf{W}_p, \{\mathbf{F}_B[k], \mathbf{W}_B[k]\} \forall k} \sum_{k=0}^{K-1} I(\mathbf{F}_p, \mathbf{W}_p, \mathbf{F}_B[k], \mathbf{W}_B[k]), \\
\text{subject to} & \quad \mathbf{F}_p \mathbf{F}_B[k] \mathbf{W}_B[k] \mathbf{W}_B[k] \mathbf{F}_B[k] = \mathbf{I}_{N_s}, \\
& \quad \mathbf{F}_p \mathbf{F}_B[k] = \mathbf{f}_{p,n_f} \mathbf{w}_{n_f} \forall f, \\
& \quad \mathbf{W}_B[k] \mathbf{W}_B[k] \mathbf{W}_B[k] = \mathbf{I}_{N_s}, \\
& \quad \mathbf{F}_p \mathbf{F}_B[k] \mathbf{W}_B[k] = \mathbf{f}_{p,n_f} \mathbf{w}_{n_f} \forall k, \\
& \quad \mathbf{W}_B[k] \mathbf{W}_B[k] \mathbf{W}_B[k] = \mathbf{I}_{N_s}, \\
& \quad \mathbf{F}_p \mathbf{F}_B[k] \mathbf{W}_B[k] = \mathbf{f}_{p,n_f} \mathbf{w}_{n_f} \forall k,
\end{align*}$$

where $\mathbf{f}_{p,n_f}$ and $\mathbf{w}_{n_f}$ are respectively the $n_f^\text{th}$ column vectors of $\mathbf{F}_p$ and $\mathbf{W}_p$. The last two constraints are the
consequences of (3) and the throughput at subcarrier $k$ is defined as [6]

$$I(F_P, W_P, F_B[k], W_B[k]) = \log_2 \det (I_{N_R} + \rho R_s^{-1} (W_B[k] H[k] F_P F_B[k]) \cdot R_s (W_B[k] H[k] F_P F_B[k])^H) \cdot R_s \cdot R_s.$$  (7)

The solution to (6) is denoted as $(F_P, \text{Opt}, W_P, \text{Opt}, (F_B[k], W_B[k]) \forall k)$. If explicit channel state information (CSI), $H[k], k = 0, \ldots, K - 1$, is available, the problem of the precoder and combiner can be solved according to the references [5], [6]. In this paper, we consider a more pragmatic approach that channel knowledge is neither given nor estimated. To efficiently get the solution to (6), we try an alternative expression of (6) which has less probability to obtain the optimal solution as follows: given two sets $\mathcal{I}_F$ and $\mathcal{I}_W$ containing the candidates of $F_P$ and $W_P$, the achievable data rate in (6) is greater than or equal to

$$\max_{F_P \in \mathcal{I}_F} \max_{W_P \in \mathcal{I}_W} \sum_{k=0}^{K-1} I(F_P, W_P, F_B[k], W_B[k]) \text{s.t. } \begin{cases} \text{W}_B[k] H[k] F_P F_B[k] = I_{N_R} \forall k \\ \text{W}_B[k] H[k] F_P F_B[k] = \text{tr}(R_s) \forall k \end{cases}$$

$$I_{LM}(F_P, W_P): \text{local maximum throughput}$$  (8)

These two versions will have the same data rate if $\mathcal{I}_F$ and $\mathcal{I}_W$ include $F_P, \text{Opt}$ and $W_P, \text{Opt}$ respectively.

The reformulated problem in (8) becomes simpler because, given $F_P$ and $W_P$, the inner problem (to obtain the local maximum throughput $I_{LM}(F_P, W_P)$) is similar to conventional fully digital beamforming designs subject to different power constraints [4], [15]. In simpler words, the critical issue of hybrid beamforming is to solve the outer problem by an additional maximization over all members of $\mathcal{I}_F$ and $\mathcal{I}_W$. Therefore, the motivation is to find $\mathcal{I}_F$ and $\mathcal{I}_W$, which ideally include $F_P, \text{Opt}$, $W_P, \text{Opt}$, and perhaps few other candidates, and then select a pair $(F_P, W_P)$ from $\mathcal{I}_F$ and $\mathcal{I}_W$ leading to the maximum throughput.

IV. HYBRID BEAMFORMING ALGORITHM BASED ON IMPPLICIT CSI

A. Initial analog beam selection

To begin with, let us see how to obtain the sets $\mathcal{I}_F$ and $\mathcal{I}_W$ in (8) from the given codebooks $\mathcal{F}$ and $\mathcal{W}$. We call this step initial analog beam selection. Since the hardware-constrained analog beamforming matrices, $F_P$ and $W_P$, cannot be replaced by the identity matrices, we have to train all or some of the columns of the codebooks $\mathcal{F} = \{ \tilde{F}_{ij}, n_f = 1, \ldots, N_F \}$ and $\mathcal{W} = \{ \tilde{w}_{nw}, n_u = 1, \ldots, N_W \}$ by transmitting known pilot signals \{(s[k])\}_{k=0}^{K-1} satisfying $|s[k]|^2 = 1 \forall k$.

Then, an observation used for the analog beam selection at subcarrier $k$ with respect to a trained beam pair $(\tilde{F}_{ij}, \tilde{w}_{nw})$ is obtained by correlating the received signal with its transmitted pilot

$$y_{nw, n_f}[k] = s^*[k] \begin{pmatrix} \sqrt{\rho \tilde{w}_n H[k] \tilde{F}_{ij} s[k]} + \tilde{w}_n [k] \end{pmatrix}$$

$$= \sqrt{\rho \tilde{w}_n H[k] \tilde{F}_{ij} + z_{nw, n_f}[k]}$$  (9)

where the effective noise $z_{nw, n_f}[k] \sim \mathcal{CN}(0, \sigma_n^2)$ still has a Gaussian distribution with mean zero and variance $\sigma_n^2$ and similar observations become available on all subcarriers $k = 0, \ldots, K - 1$. The observation $y_{nw, n_f}[k]$ can also be viewed as a coupling coefficient of the channel and the beam pair $(\tilde{F}_{ij}, \tilde{w}_{nw})$.

Borrowing the idea from our previous works in [11], [16], it shows that if the columns of $\mathcal{F}$ and $\mathcal{W}$ are orthogonal, the sum of the power of $K$ observations in one OFDM symbol can be directly used for the analog beam selection. As a result, $M$ (assume $M \geq N_{RF}$) analog beam pairs can be selected individually and sequentially according to the sorted received power estimates

$$\{\bar{F}_{pm}, \tilde{w}_pm\} = \arg \max \sum_{k=0}^{K-1} |y_{nw, n_f}[k]|^2$$  (10)

where $m = 1, \ldots, M$, $\mathcal{F}' = \{ \bar{F}_{pm}, n = 1, \ldots, m - 1 \}$ and $\mathcal{W}' = \{ \tilde{w}_pm, n = 1, \ldots, m - 1 \}$ are the sets consisting of the selected analog beamforming vectors from iteration 1 to $m - 1$. We assume that $M \geq N_{RF}$ for the reason that the first $N_{RF}$ selected analog beam pairs according to the sorted magnitude of $\sum_{k=0}^{K-1} |y_{nw, n_f}[k]|^2$ may not always lead to a good solution because we do not yet consider the effect of digital beamforming during the analog beam selection phase. To find the optimal solution, we can further take into account the linear combination of $N_{RF}$ analog beamforming vectors selected from $\{\bar{F}_{pm}, \tilde{w}_pm\} \forall m$ with coefficients in digital beamforming.

After selecting $M$ analog beam pairs, we define two sets $\mathcal{I}_F$ and $\mathcal{I}_W$ consisting of all the combinations of any $N_{RF}$ distinct analog beamforming vectors selected from $\{\bar{F}_{pm}, m = 1, \ldots, M \}$ and $\{ \tilde{w}_pm, m = 1, \ldots, M \}$, respectively, which can be written as

$$\mathcal{I}_F = \{\bar{F}_{pij}, i_f = 1, \ldots, I_F\}$$

$$\mathcal{I}_W = \{\tilde{w}_{piw}, i_w = 1, \ldots, I_W\}$$  (11)

where the cardinality $I_F = I_W = \binom{M}{N_{RF}}$ of both sets is given by the binomial coefficient. The notation $\bar{F}_{pij}$ and $\tilde{w}_{piw}$ respectively denote the $i_f$th and $i_w$th candidates of the analog beamforming matrices $\bar{F}_p$ and $\tilde{w}_p$. When $M$ becomes large, there is a high probability that $\mathcal{I}_F$ and $\mathcal{I}_W$ include the optimal solution $F_P, \text{Opt}$ and $W_P, \text{Opt}$.

Schematic example: let us consider a scenario with $N_T = N_R = 8$ antenna elements, codebook sizes $N_F = N_W = 8$, orthogonal codebooks $\mathcal{F} = \mathcal{W}$ with steering angles given by $\{ -90^\circ \text{or} 90^\circ \}, -48.59^\circ, -30^\circ, -14.48^\circ, 0^\circ, 14.48^\circ, 30^\circ$.
one can try all 36 pairs of the members of \( F \), and the difference in path attenuation between path one and two amounts to 10 dB.

However, which beamforming strategy has higher throughput

there will be a competition between spatial multiplexing gain and choose the beam pairs yielding the maximum throughput, which will be detailed in the following subsections. In general, and enumerate them explicitly as

\[
I_F = \{ F_{P,1}, i_f = 1, \ldots, 6 \},
\]

\[
I_W = \{ W_{P,i_w}, i_w = 1, \ldots, 6 \}.
\]

For instance, \( F_{P,1} = [\tilde{f}_5, \tilde{f}_6] \) and \( W_{P,1} = [\tilde{w}_3, \tilde{w}_4] \). Then, one can try all 36 pairs of the members of \( I_F \) and \( I_W \) to determine the optimal weightings in digital beamforming and choose the beam pairs yielding the maximum throughput, which will be detailed in the following subsections. In general, there will be a competition between spatial multiplexing gain over different propagation paths and power gain available from the dominant path. In this case, as shown in Fig. 2(b) with the beams highlighted in blue, the two analog beam pairs steering to the dominant path lead to higher spectral efficiency. However, which beamforming strategy has higher throughput in any specific case is not clear beforehand.

**B. Digital beamforming**

After the initial analog beam selection, we are in possession of the two sets \( I_F \) and \( I_W \) that contain the candidates of \( F_P \) and \( W_P \), and the objective is to rapidly find which pair is the optimal. Before going into the detail of our proposed scheme, let us review the relationship between the analog and digital beamforming. Given one particular choice \( (F_{P,i_f}, W_{P,i_w}) \) selected from the candidate sets \( I_F \) and \( I_W \), it is clear that the goal of digital beamforming is to maximize the local maximum throughput \( I_{LM}(F_{P,i_f}, W_{P,i_w}) \) as defined in (8). As a consequence, for each subcarrier the digital beamforming problem can be formulated as a throughput maximization problem subject to the power constraints, which can be stated as

\[
(F_{B,i}[k], W_{B,i}[k]) = \arg \max \{ I(F_{P,i_f}, W_{P,i_w}, F_B[k], W_B[k]) \}
\]

s.t. \( W_B[k] W_{P,i_w}^H = \text{tr}(R) = 1 \)

where \( i = (i_f - 1)I_W + i_w \) is the index specifying the combined elements from \( I_F \) and \( I_W \). In this problem, given \((F_{P,i_f}, W_{P,i_w})\), the corresponding optimal digital beamforming matrices are given by (a detailed description of this part can be found in the journal version of this paper [17])

\[
F_{B,i}[k] = (F_{P,i_f})^{-0.5} [V_{E,i}[k]]_{1:N_S},
\]

\[
W_{B,i}[k] = (W_{P,i_w})^{-0.5} [U_{E,i}[k]]_{1:N_S},
\]

where the columns of \( V_{E,i}[k] \) and \( U_{E,i}[k] \) are respectively the right and left singular vectors of the effective channel

\[
H_{E,i}[k] = (W_{P,i_w}^H)^{-0.5} H[k] (F_{P,i_f}^H)^{-0.5} = \text{SVD} \ U_{E,i}[k] \Sigma_{E,i}[k] V_{E,i}[k].
\]

**C. Key parameter of hybrid beamforming gain**

Given a pair of elements selected from \( I_F \) and \( I_W \) and the corresponding optimal digital beamforming matrices across all
subcarriers, we can evaluate the local maximum throughput as

$$I_{LM}(F_{P,i}, W_{P,iw}) = \sum_{k=0}^{N_s} \sum_{n_s=1}^{N_c} \log_2 \left( 1 + \frac{\rho}{\sigma_n^2} \left[ \Sigma_{E,i}[k] \right]_{n_s,n_s} \right),$$

(15)

where the diagonal elements of $\Sigma_{E,i}[k]$ are the singular values of the effective channel $H_{E,i}[k]$. Based on the candidate set \{\((F_{P,i}, W_{P,iw})| \forall i_f, i_w\}\), the pair leading to the maximum throughput provides the best approximation of the global optimal analog beamforming matrices, denoted as (\(F_P, W_P\)), as well as the solution to the hybrid beamforming problem in (8).

\[
(\hat{F}_P, W_P) = \arg \max_{F_{P,i}, W_{P,iw} \in \mathcal{P}} I_{LM}(F_{P,i}, W_{P,iw}).
\]

(16)

However, this way of solving the problem requires the SVD of \{\(H_{E,i}[k] \}_{k=0}^{K-1}\) to obtain \(I_{LM}(F_{P,i}, W_{P,iw})\) for each pair, which means that we have to repeat the calculation as many as \(\binom{M}{2}\) times.

To reduce the potentially large computational burden, we ask ourselves what are the crucial parameter(s) or indicator(s) that actually determine the throughput. To answer this question, let \(R_s = \frac{1}{N_s} I_{N_s}\) (equal power allocation) so that the maximum achievable throughput at subcarrier \(k\) becomes

\[
I(F_{P,i}, W_{P,iw}, F_{B,i}[k], W_{B,i}[k]) = \sum_{n_s=1}^{N_s} \log_2 \left( 1 + \gamma \left[ \Sigma_{E,i}[k] \right]_{n_s,n_s} \right),
\]

(17)

where \(\gamma = \frac{\rho}{\sigma_n^2}\) is the SNR. To find the key parameter of the hybrid beamforming gain, we focus on the low SNR regime. Using the fact that \(\log(1 + \gamma x) \approx \gamma x\) as \(\gamma \to 0\), the achievable data rate in (17) can be approximated by

\[
I(F_{P,i}, W_{P,iw}, F_{B,i}[k], W_{B,i}[k]) \approx 0 \sum_{n_s=1}^{N_s} \left[ \Sigma_{E,i}[k] \right]_{n_s,n_s}
\]

\[
\approx \sum_{n_s=1}^{N_s} \left[ \text{tr}(\Sigma_{E,i}[k]) \right]_{n_s,n_s}
\]

\[
\lesssim \|H_{E,i}[k]\|_F^2
\]

(18)

with equality iff \(N_{RF} = N_s\). For the case of \(N_{RF} > N_s\), \(\|H_{E,i}[k]\|_F^2\) corresponds to the sum of all \(N_{RF}\) eigenvalues of \(H_{E,i}[k]H_{E,i}[k]^H\). Assuming that the sum of the weaker \(N_{RF} - N_s\) eigenvalues of \(H_{E,i}[k]H_{E,i}[k]\) is small, the approximation of \(\sum_{n_s=1}^{N_s} \left[ \Sigma_{E,i}[k] \right]_{n_s,n_s}\) by \(\|H_{E,i}[k]\|_F^2\) seems to be valid for most cases of interest.

Fortunately, the matrix $H_{E,i}[k]$ can be easily obtained from the original observations \{\(y_{n_w,n_f}[k]| \forall n_w, n_f, k\}\) [16]. Let us show the effective channel in (14) again and approximate the matrix $W_{P,iw} H_{E,i}[k] F_{P,i}$ by $Y_i[k]$ (the elements of $Y_i[k]$ can be collected from the observations),

$$H_{E,i}[k] \triangleq \left( W_{P,iw} H_{E,i}[k] F_{P,i} \right)^{-0.5} \left( W_{P,iw} H_{E,i}[k] F_{P,i} \right)^{-0.5} \approx \left( Y_i[k] \right)^{-0.5} \left( Y_i[k] \right)^{-0.5} \triangleq H_{E,i}[k].$$

(19)

For example, if $F_{P,i} = [\hat{f}_{P,1}, \cdots, \hat{f}_{P,NRF}]$ and $W_{P,iw} = [w_{P,1}, \cdots, w_{P,NRF}]$, where $\hat{f}_{P,n_f}$ is the $n_f^{th}$ column of $F$ and $\hat{w}_{P,n_{RF}}$ is the $n_{RF}^{th}$ column of $W$, one has

$$Y_i[k] = \begin{bmatrix}
    y_{1,1}[k] & \cdots & y_{1,NRF}[k] \\
    \vdots & \ddots & \vdots \\
    y_{N_{RF},1}[k] & \cdots & y_{N_{RF},N_{RF}}[k]
\end{bmatrix}$$

(20)

Therefore, given a pair $(F_{P,i}, W_{P,iw})$ selected from $\mathcal{F}_F$ and $\mathcal{F}_W$, we can rapidly obtain the approximation of $H_{E,i}[k]$, denoted as $\hat{H}_{E,i}[k]$ in (19).

To conclude, the proposed solution can be stated as: first obtain the candidate sets $(\mathcal{F}_F$ and $\mathcal{F}_W)$ and the approximation of $H_{E,i}[k]$ ($\hat{H}_{E,i}[k]$) from the observations \{\(y_{n_w,n_f}[k]| \forall n_w, n_f, k\)$ and then solve the maximization problem in (16), which can be rewritten as

$$\left( i_f, i_w \right) = \arg \max_{F_{P,i}, W_{P,iw} \in \mathcal{F}_F, \mathcal{F}_W} I_{LM}(F_{P,i}, W_{P,iw}) \approx \max_{F_{P,i}, W_{P,iw} \in \mathcal{F}_F, \mathcal{F}_W, i = (i_f - 1)I_W + i_w} \sum_{k=0}^{K-1} \|\hat{H}_{E,i}[k]\|^2_F.$$ 

(21)

According to the index pair $(i_f, i_w)$, we have the selected analog and corresponding digital beamforming matrices given by

$$\hat{F}_P = F_{P,i},$$

$$\hat{W}_P = W_{P,iw},$$

$$\hat{F}_{B}[k] = (\hat{F}_P^H \hat{F}_P)^{-0.5} \left[ \hat{Y}_{E,i}[k] \right]_{1:1:N_S},$$

$$\hat{W}_{B}[k] = (\hat{W}_P^H \hat{W}_P)^{-0.5} \left[ \hat{U}_{E,i}[k] \right]_{1:1:N_S},$$

(22)

\(\hat{y} = (\hat{i}_f - 1)I_W + \hat{i}_w\) and $\hat{U}_{E,i}[k] \hat{Y}_{E,i}[k] \hat{Y}_{E,i}[k]^H$ is SVD$(\hat{H}_{E,i}[k])$. The complete proposed algorithm of the hybrid beamforming implementation based on implicit CSI (i.e., the coupling coefficients \{\(y_{n_w,n_f}[k]| \forall n_w, n_f, k\)$) is shown in Algorithm 1, where $f(\hat{H}_{E,i}[k])$ in Step 4 denotes the analog beam selection criterion by using (17) or (18) with the argument $\hat{H}_{E,i}[k]$ as

\[
f(\hat{H}_{E,i}[k]) = \begin{cases}
    \sum_{n_s=1}^{N_s} \log_2 \left( 1 + \gamma \left[ \Sigma_{E,i}[k] \right]_{n_s,n_s} \right), & \text{w/o approx.} \\
    \|\hat{H}_{E,i}[k]\|_F^2, & \text{w/ approx.}
\end{cases}
\]

(23)
Algorithm 1: Hybrid beamforming based on implicit CSI

Input: \( \{y_{n,w,n_I}|v_{n,w,n_I}\} \vphantom{\sum_{n_k}} \)

Output: \( \{F_P, W_P, (\tilde{F}_B[k], \tilde{W}_B[k])\} \forall k \)

1. Given \( \{y_{n,w,n_I}|v_{n,w,n_I}\} \), select \( M \) analog beam pairs \( \{f_{m,n_I}|\tilde{w}_{m,n_I}\} \), where \( m = 1, \cdots, M \), according to (10).

2. Generate two candidate sets \( I_F \) and \( I_W \) based on \( \{f_{m,n_I}|\tilde{w}_{m,n_I}\} \) and \( \{W_{m,n_I}|\tilde{w}_{m,n_I}\} \), respectively.

3. \( H_{E,i}[k] = (\tilde{W}_P^H[k] \tilde{W}_P[k])^{-0.5} Y_i[k] (\tilde{F}_P[k] \tilde{F}_{P,i}^{-1})^{-0.5} \),

where \( \tilde{F}_P, \tilde{F}_{P,i} \) \( \in I_F \), \( \tilde{W}_P, \tilde{W}_{P,i} \) \( \in I_W \), and the entries of \( Y_i[k] \) are collected from \( \{y_{n,w,n_I}|v_{n,w,n_I}\} \).

4. \( (i_f, i_w) = \arg \max_i \sum_{k=0}^{K-1} f(H_{E,i}[k]) \),

where \( f(H_{E,i}[k]) \) is given in (23).

5. Output: \( F_P, W_P \) and \( \tilde{F}_B[k], \tilde{W}_B[k] \).

6. \( U_{E,i}[k] = V_{E,i}[k] (\tilde{W}_P^H[k] \tilde{W}_P[k])^{-0.5} (\tilde{F}_P[k] \tilde{F}_{P,i}^{-1})^{-0.5} \),

where \( i = (i_f - 1)I_W + i_w \).

7. Output: \( \vphantom{\sum_{n_k}} \)

\[ \begin{align*}
\tilde{F}_B[k] &= (\tilde{W}_P^H[k] \tilde{F}_P[k])^{-0.5} [\tilde{V}_{E,i}[k]]_{1:1,N_d}, \vphantom{\sum_{n_k}} \\
\tilde{W}_B[k] &= (\tilde{W}_P^H[k] \tilde{W}_P[k])^{-0.5} [\tilde{U}_{E,i}[k]]_{1:1,N_d}
\end{align*} \]

The advantages of the proposed algorithm are summarized as follows: (1) we can omit high-dimensional channel estimation problems, and (2) even though the cardinalities of \( I_F \) and \( I_W \) are large, the computational overhead is minor because we just need to calculate the Frobenius norm of the effective channel matrices, whose elements can be easily obtained from the observations \( \{y_{n,w,n_I}|v_{n,w,n_I}\} \).

In (17), we simply assume that the transmit power is equally allocated to \( N_S \) data streams to facilitate the process of finding the best value of the key parameter. Once the analog and digital beamforming matrices are selected, the global maximum throughput can be further improved by optimizing the power allocation (i.e., by a water-filling power allocation scheme [15]) for \( N_C \) data streams according to the effective channel condition \( H_{E,i}[k] \) at each subcarrier.

V. SIMULATION RESULTS

The system has \( N_T = N_R = 32 \) antennas, \( N_{RF} = 2 \) RF chains, \( N_S = 2 \) data streams, and \( K = 512 \) subcarriers. The cluster-based channel model has \( C = 5 \) clusters including one line-of-sight (LoS) and four NLoS clusters, and each cluster has \( R = 8 \) rays. The codebooks \( F \) and \( W \) have the same number of members, \( N_F = N_W = 32 \), and 32 steering angle candidates are: \( \{ \frac{\pi}{2} \sin^{-1} \left( \frac{(16 - 16)}{16} \right) \} \), \( n_f = 1, \cdots, 32 \), which yield orthogonal codebooks [18].

We chose the work in [5] that implements hybrid beamforming based on the singular vectors of \( H[k] \) \( \forall k \) (i.e., explicit CSI) as a reference method for comparison and extended it from single carrier to multiple carriers. Different to the reference scheme, Algorithm 1 uses the received coupling coefficients (i.e., implicit CSI) as the observations for the hybrid beamforming implementation. The received coupling coefficients are commonly used for channel estimation [9], [19], but in this paper we use them to directly implement the hybrid beamforming on both sides. As a result, we can get rid of the overhead of channel estimation.

To clearly present the difference in throughput by using different methods, the calculated throughput values are normalized to the throughput achieved by fully digital beamforming\(^1\).

Fig. 4 shows the achievable data rates with different methods, the calculated throughput values are normalized to the throughput achieved by fully digital beamforming\(^1\).

The data rates shown in \( I_{Pro}(Eig, M = 3) \) and its approximation achieved by using the key parameter.

\(^1\)The data rates achieved by fully digital beamforming used for the normalization from \( SNR = -20 \) dB to \( 20 \) dB (step by \( 5 \) dB) are: \( \{0.05, 0.14, 0.41, 1.03, 2.17, 3.77, 5.79, 8.17, 10.91\} \) in bit/s/Hz.
The research leading to these results has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 671551 (5G-X-Haul) and the TUD-NEC project “mmWave Antenna Array Concept Study”, a cooperation project between Technische Universität Dresden (TUD), Germany, and NEC, Japan.

REFERENCES