Abstract—This letter investigates bit-interleaved coded multiple beamforming (BICMB) with perfect coding in millimeter-wave (mm-wave) multiple-input multiple-output (MIMO) systems to achieve the maximum diversity gain. Using perfect coding with BICMB enables us to do this. We show that by using BICMB and perfect coding, the diversity gain becomes independent of the number of transmitted data streams and the number of antennas in each remote antenna unit (RAU) at the transmitter and the receiver. The assumption is that the perfect channel state information (CSI) is known at both the transmitter and the receiver. With the assumption of the perfect CSI at the transmitter and the receiver, simulation results show that the use of BICMB with perfect coding results in the diversity gain values predicted by the analysis.

I. INTRODUCTION

Diversity gain analysis of a millimeter-wave (mm-wave) multiple-input multiple-output (MIMO) system with distributed antenna-subarray (DAS) architecture was first studied in [1]. The diversity gain calculated in [1] depends on the number of transmitted data streams in the system. This means by increasing the number of transmitted data streams, the diversity gain decreases. Furthermore, the diversity gain in [1] can be increased simply by increasing the number of antenna subarrays. Diversity gain analysis for the mm-wave MIMO systems is studied in [2].

Bit-interleaved coded modulation (BICM) was first introduced to increase the code diversity [3], [4]. Later on, bit-interleaved coded multiple beamforming (BICMB) was used to achieve full diversity gain and full multiplexing gain in MIMO systems [5], [6]. In this method, different codewords are interleaved among different subchannels with different diversity orders. To overcome this diversity degradation, in [7], we proved that by using BICMB in a mm-wave massive MIMO system with DAS architecture both full diversity gain and full multiplexing gain can be achieved.

Perfect space-time block codes (PSTBC) were studied in [8], [9] to achieve full rate and full diversity in any dimension. However, dimensions 2, 3, 4 and 6 are the only dimensions that can achieve an increase in the coding gain. In [10], perfect coding with multiple beamforming is used to achieve full diversity and full multiplexing in a MIMO system with less decoding complexity than a system employing PSTBC and full precoded multiple beamforming (FPMB). In [11], channel coding is added to the perfect coding and diversity gain analysis is carried out to prove that BICMB with perfect coding (BICMB-PC) achieves the full diversity order.

Space-time block codes (STBC) are studied in massive MIMO literature. In [12], space-time block codes are used to achieve full diversity gain in a flat fading non-coherent wireless communication system. Also, in [13] massive space-time block code (MaSTBC) is studied. Authors in [13] proposed a novel space-time modulation scheme with PSK modulated MaSTBC for multi-user massive MIMO uplink systems.

In this work, we use BICMB-PC to achieve full diversity gain in mm-wave MIMO systems. The diversity analysis for this system is carried out. We show that by using perfect coding in addition to convolutional coding, the diversity gain becomes independent of the number of transmitted data streams.

II. SYSTEM MODEL

One can approximate the average probability of bit error rate (BER) $P_E$ at high SNR regimes for both coded and uncoded systems as [14], [15]

$$P_E \approx (G_c \bar{\gamma})^{-G_d}, \quad (1)$$

where $G_c$ and $G_d$ are defined as coding gain and diversity gain, respectively. Note that diversity gain is not a property of high SNR regimes. Average SNR is shown by $\bar{\gamma}$. In a log-log scale, diversity gain $G_d$ determines the slope of the BER versus the average SNR curve in high SNR regime. Furthermore, changing $G_c$ leads to shift of the curve in SNR relative to a benchmark BER curve of $(\bar{\gamma}^{-G_d})$. In this work, our focus is on calculating the diversity gain and we leave the coding gain for future work.

We consider a single-user mm-wave MIMO scenario shown in Fig. 1, where the transmitter is equipped with $M_t$ RAUs, $N_t$ antennas at each RAU, and $N_{RF}$ RF chains. The receiver has $M_r$ RAUs, $N_r$ antennas at each RAU, and $N_{RF}$ RF chains. The transmitter sends $N_t = D$ data streams to the receiver. These data streams are generated as follows. First the bit codeword $c$ is generated through a convolutional encoder with code rate $R_c$. Then a random bit-interleaver is used to generate an interleaved sequence. The output of the interleaver is modulated by $M$-quadrature amplitude modulation (M-QAM). We define a one-to-one mapping from $X_k = [x_{1,k}, \ldots, x_{D,k}]$ to $Z_k$ as $Z_k = \mathbb{M} \{X_k\}$ where $\mathbb{M}$ denotes the PSTBC codeword.
generating function [11]. A PSTBC codeword, i.e., $Z_k$, is
generated by using $D^2$ consecutive complex-valued scalar
symbols [11]
\[
Z_k = \mathbb{M}(X_k) = \sum_{v=1}^{D} \text{diag}(Gx_{v,k}) E^v \text{E}^{-1},
\]
where $G$ is an $D \times D$ unitary matrix [8], $x_{v,k}$ is a $D \times 1$
vector whose elements are the $v$th $D$ input modulated scalar symbols
and $D \in \{2,3,4,6\}$. Matrix $E$ is defined as
\[
E = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 & 0 \\
0 & 0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 1 \\
g & 0 & 0 & \cdots & 0 & 0
\end{bmatrix},
\]
where
\[
g = \begin{cases}
i, & D = 2,4, \\
e^{-\frac{\pi}{4}}, & D = 3, \\
-\frac{\pi}{\sqrt{2}}, & D = 6.
\end{cases}
\]
As it can be seen from Fig. 1, the complex-valued matrix
$F_{BB} \in \mathbb{C}^{M_{RF} \times N_{RF}}$ is used for preprocessing at the baseband.
A set of $M_tN_t$ phase shifters is applied to the output of each RF
channel. As a result of this process, different beams are formed
in order to transmit the RF signals. We can model this process
with a complex-valued matrix
\[
\hat{M}_k = \sum_{v=1}^{D} \text{diag}(Gx_{v,k}) E^v \text{E}^{-1},
\]
where $L_{ij}$ is the number of propagation paths and $\alpha_{ij}^l$ is
the complex-gain of the $l$th ray which follows $CN(0,1)$, $\theta_{ij}^l \in [0,2\pi]$, $\phi_{ij}^l \in [0,2\pi]$, for all $i,j,l$, and the vectors $\mathbf{a}_r(\theta_{ij}^l)$ and $\mathbf{a}_i(\phi_{ij}^l)$ are the normalized array response at the receiver
and transmitter, respectively. In particular, this paper adopts a
uniform linear array (ULA) where both $\mathbf{a}_r(\theta_{ij}^l)$ and $\mathbf{a}_i(\phi_{ij}^l)$ are modeled as
\[
\mathbf{a}_{ULA}(\phi) = \frac{1}{\sqrt{N}} \begin{bmatrix}
1, e^{j\frac{2\pi}{\lambda}d \sin(\phi)}, \ldots, e^{j(N-1)\frac{2\pi}{\lambda}d \sin(\phi)}
\end{bmatrix}^T,
\]
where $\lambda$ is the transmission wavelength, and $d$ is the antenna
spacing.

The processed signal at the $k$th PSTBC codeword is
\[
Y_k = W_{BB}^{H} W_{RF}^{H} H F_{RF} \mathbf{Z}_k + W_{BB}^{H} W_{RF}^{H} \mathbf{n}_k,
\]
where $Y_k$ is an $D \times D$ complex-valued matrix, $\mathbf{n}_k$ is an $M_tN_t \times 1$
vector consisting of i.i.d. $CN(0,N_0)$ noise samples, $N_0 = \frac{N_{RF}}{SNR}$
where $SNR$ is defined as the signal-to-noise ratio (SNR),
$W_{RF}$ is the $M_tN_t \times N_{RF}$ RF combining matrix, and $W_{BB}$ is the
$N_{RF} \times N_{t}$ baseband combining matrix.

A solution based on singular value decomposition (SVD)
of the channel matrix $H = U \Lambda V^H$ can be derived for the
beamforming matrices [7]. By utilizing the optimal precoder
and combiner, one can write (8) as
\[
Y_k = \mathbf{A}Z_k + \tilde{\mathbf{n}}_k,
\]
where $\mathbf{n}_k = U^{H}_{(1:D)} \mathbf{n}_k$, and $U_{(1:D)}$ is the first $D$
columns of the unitary matrix $U$.

We model the PSTBC codeword sequence as $k' \rightarrow (k,(m,n),j)$,
where $k'$ represents the original ordering of the coded bits $c_{k'}$, $(k,(m,n),j)$ is the index of the PSTBC
codewords, the symbol position in $X_k$, and the bit position
on the label of the scalar symbol $x_{(m,n),k}$ respectively.
We define $\chi_b$ as the subset of all signals $x \in \chi$. Note that
the label has the value $b \in \{0,1\}$ in position $j$.

The maximum likelihood (ML) bit metrics for (8) can be
written as
\[
\gamma^{(m,n),j}(Y_k,c_{k'}) = \min_{X \in \chi_b} ||Y_k - A\mathbb{M}(X)||^2,
\]
where $\eta^{(m,n),j}_{c_{k'}}$ is defined as
\[
\eta^{(m,n),j}_{c_{k'}} = \{X : x_{(u,v)=(m,n)} \in \chi_{c_{k'}},
\text{and } x_{(u,v)=(m,n)} \in \chi\}.
\]
The ML decoder at the receiver makes decisions according to
the rule
\[
\hat{c} = \arg \min_{c} \sum_{k'} \gamma^{(m,n),j}(Y_k,c_{k'}).
\]

III. DIVERSITY GAIN ANALYSIS

In this section, the diversity gain is examined for
mm-wave MIMO systems employing DAS architecture employing
BICMB-PC. We show that the diversity gain becomes inde-
pendent of the number of transmitted streams, whereas in [1]
the diversity gain is dependent on the number of transmitted data streams. This will be done by computing an upper bound for the pairwise error probability (PEP).

**Theorem 1.** Suppose that $N_r$ and $N_t$ are sufficiently large [7]. Then, by utilizing BICMB-PC, mm-wave MIMO systems with DAS architecture can achieve a diversity gain of

$$G_d = \left( \frac{\sum_{i,j} \beta_{ij}}{\sum_{i,j} \beta_{ij}^2} \right)^2$$

for $i = 1, \ldots, M_r$ and $j = 1, \ldots, M_t$.

**Proof.** Assume that codeword $c$ is transmitted and codeword $\hat{c}$ is detected. Then one can write the PEP of $c$ and $\hat{c}$ as

$$P(c \rightarrow \hat{c}|H) = P \left( \sum_{i,j} \left| \mathbf{Y}_k - \mathbf{Z} \right|^2 \geq \sum_{i,j} \left| \mathbf{Y}_k - \mathbf{Z} \right|^2 \right) \leq \left( \frac{\sum_{i,j} \left| \mathbf{Y}_k - \mathbf{Z} \right|^2}{\sum_{i,j} \left| \mathbf{Y}_k - \mathbf{Z} \right|^2} \right) \leq \left( \frac{\sum_{i,j} \left| \mathbf{Y}_k - \mathbf{Z} \right|^2}{\sum_{i,j} \left| \mathbf{Y}_k - \mathbf{Z} \right|^2} \right)^2$$

where $\mathbf{Y} = \mathbf{X} - \mathbf{Z}$, $\mathbf{Z} = \mathbf{X} - \mathbf{Z}$, and $\mathbf{Z} = \mathbf{X} - \mathbf{Z}$.

By using an upper bound of the $Q$ function $Q(x) \leq \frac{1}{2} e^{-x^2}$, the average PEP in is upper bounded as

$$P(c \rightarrow \hat{c}) = E[P(c \rightarrow \hat{c}|H)] \leq \frac{1}{2} \mathbb{E} \left[ \exp \left( -\frac{\sum_{i,j} \left| \mathbf{Y}_k - \mathbf{Z} \right|^2}{2N_0} \right) \right].$$

By using [10] and rewrite (17) as

$$P(c \rightarrow \hat{c}) = \frac{1}{2} \mathbb{E} \left[ \exp \left( -\frac{\sum_{u=1}^{L t} A_u^2 \zeta_u}{2N_0} \right) \right],$$

where $\zeta_u = \sum_{k',d_H} \rho_{u,k}$ and $\rho_{u,k} = \sum_{d_H} \left| \mathbf{g}^{(d_H)}(\mathbf{x}_{c,k} - \hat{\mathbf{x}}_{c,k}) \right|^2$.

By defining $L_t = \sum_{i,j} \beta_{ij}$ as the rank of the channel matrix $H$, i.e., the number of singular values of $H$, we can write

$$\left( \frac{\zeta_{\min} \sum_{u=1}^{L t} A_u^2}{L_t} \right) \leq \left( \frac{\zeta_{\min} \sum_{u=1}^{D} A_u^2}{D} \right) \leq \left( \frac{\sum_{u=1}^{D} A_u^2 \zeta_u}{D} \right).$$

One can define

$$\Theta \triangleq \sum_{u=1}^{L t} A_u^2 \leq \sum_{i,j} \beta_{ij} \left| H_{ij} \right|^2 \leq \sum_{i,j} M_r M_t \beta_{ij} \left| H_{ij} \right|^2.$$  

When $N_r$ and $N_t$ are sufficiently large, the singular values of $H_{ij}$ converge to $\sqrt{N_r N_t} \alpha_{ij}^2$ in descending order [7]. Therefore, one can rewrite (20) as

$$\Theta = \sum_{i,j} \beta_{ij} \left| H_{ij} \right|^2 = N_r N_t \sum_{i,j} \beta_{ij} \left| H_{ij} \right|^2.$$  

Note that the random variable $\sum_{i,j} \alpha_{ij}^2$ has a chi-squared distribution with $2L_{ij}$ degrees of freedom, or equivalently a Gamma distribution with shape $L_{ij}$ and scale 2, i.e., $\mathcal{G}(L_{ij}, 2)$. Then, since $\beta_{ij} \leq 1$, $\Psi_{ij} = \mathcal{G}(L_{ij}, 2\beta_{ij}L_{ij}^{-1})$ [17]. The Welch-Satterthwaite equation is used here to approximate the shape and scale of the Gamma distribution. One can see that $\Theta$ is a linear combination of the independent random variables $\Psi_{ij}$ [18, p.4.1-1]. The shape and scale of $\Theta$ can be calculated as

$$\kappa = \frac{\sum_{i,j} \theta_{ij} \beta_{ij}^2}{\sum_{i,j} \beta_{ij}^2 L_{ij}^{-1}},$$

$$\theta = \frac{\sum_{i,j} \beta_{ij}^2 L_{ij}^{-1}}{\sum_{i,j} \beta_{ij}^2}.$$

By using (19) and the definition of the moment generating function (MGF) [19], we can upper-bound the PEP in (17) by

$$P(c \rightarrow \hat{c}) \leq \frac{1}{2} \mathbb{E} \left[ \exp \left( -\frac{\sum_{u=1}^{L t} A_u^2 \zeta_u}{2N_0} \right) \right].$$

By using MGF of $\Theta$, (24) can be written as

$$P(c \rightarrow \hat{c}) \leq \frac{1}{2} \left[ 1 + \theta \frac{\zeta_{\min} D N_t}{4L_t} \frac{1}{\Theta} \right]^{-\kappa}$$

for high SNR.

Hence, BICMB-PC achieves full diversity order of

$$G_d = \kappa = \left( \frac{\sum_{i,j} \beta_{ij}^2}{\sum_{i,j} \beta_{ij}^2 L_{ij}^{-1}} \right)$$

which is independent of the number of transmitted data streams.

**Remark 1.** By assuming that $L_{ij} = L$ and $\beta_{ij} = \beta$ for any $i \in \{1, \ldots, M_r\}$ and $j \in \{1, \ldots, M_t\}$, it can be seen easily that the mm-wave MIMO system with DAS architecture can achieve a diversity gain

$$G_d = M_r M_t L = D^2 L.$$  

(28)
One can compare this result with the diversity gain calculated for the single-user scenario in [1]. As it can be seen, similar to [7], the full diversity gain is achieved in this paper.

IV. Decoding

By replacing (2) in (9), we can rewrite (9) and show that each element of $AZ_k$ is related to only one of the $x_{v,k}$ [10], [11]. For the case $D = 3$, we can write

$$Y_k = \begin{bmatrix} \lambda_1 g_1^T x_{1,k} + n_1 \\ \lambda_2 g_2^T x_{2,k} + n_2 \\ \lambda_3 g_3^T x_{3,k} + n_3 \end{bmatrix}. \quad (29)$$

The processed signal in (29) can be divided into $D$ parts. Then one can write

$$y_{v,k} = \Omega_v \Lambda G x_{v,k} + \tilde{n}_{k,v} \quad (30)$$

where $v = 1, \ldots, D$ and $\Omega_v = \text{diag}(\omega_{v,1}, \ldots, \omega_{v,D})$. The elements of the matrix $\Omega_v$ are defined as

$$\omega_{v,u} = \begin{cases} 1, & 1 \leq u \leq D - v + 1 \\ g, & \text{otherwise} \end{cases} \quad (31)$$

One can simplify (30) by using the QR decomposition of $\Lambda G = QR$ as done in [11] to simplify the ML bit metrics defined in (10) as follows:

$$\gamma^{(m,n),j}(y_{k,c^v}) = \min_{x_{d+1} = \chi} ||\tilde{y}_{m,k} - Rx||^2, \quad (32)$$

where $\tilde{y}_{m,k} = Q^H \Omega^H y_{m,k}$, and $\rho_{n,v}^{n,j}$ is a subset of $\chi^D$. This subset is defined as

$$\rho_v^{n,j} = \{x = [x_1, \ldots, x_D]^T : x_{d+1} \in \chi^D, \text{and } x_{d+2} \in \chi \}. \quad (33)$$

As mentioned in [11], the complexity order of the simplified ML bit metrics (32) is proportional to $M^D$, i.e., $O(M^D)$. By using sphere decoding (SD), the average complexity reduces, but the worst-case scenario is still in the order of $M^D$, i.e., $O(M^D)$ [20]. Furthermore, for dimensions 2 and 4, the complexity can be reduced by separating the real part and imaginary part of $\tilde{y}_{m,k}$. Note that in these cases, the matrix $R$ is real. By doing this separation, the complexity order for the worst-case scenario reduces to $O(M^D - 0.5)$ [11].

V. Simulation Results

In the simulations, the industry standard 64-state 1/2-rate (133,171) $d_{free} = 10$ convolutional code is used. For BICMB, we separate the coded bits into different substreams of data and a random interleaver is used to interleave the bits in each substream. We assume that the number of RF chains in the receiver and transmitter are twice the number of data streams [7] (i.e., $Ntf = Nrf = 2Nc$). Also, each scale fading coefficient $\beta_{ij} = -20$ dB for all simulations, except for Fig. 5. At RAUs in both the transmitter and the receiver, ULA array configuration with $d = 0.5$ is considered. Information bits are mapped onto 16-QAM symbols in each subchannel.

Fig. 2 illustrates the results for BICMB-PC for different values of $D$ and $L_{ij}$ in a mm-wave MIMO system. Furthermore, we can see the comparison of the BICMB-PC with the BICMB results in [7]. Please note that for the sake of comparison, $M_r = M_t = D$ in simulations related to [7], where $M_r$ is the number of RAUs at the receiver side and $M_t$ is the number of RAUs at the transmitter side. The number of propagation paths are defined as $L = [L_{11}, L_{12}, L_{21}, L_{22}]$ and $L = [L_{11}, L_{12}, L_{21}, L_{22}, L_{23}, L_{31}, L_{32}, L_{33}]$. When $L = L$, all elements in $L$ are constant and equal to $L$. It can be seen that the diversity gain remains the same for different values for the number of propagation paths, as long as (27) returns the same value of $G_d$. For example for the dashed-dot line curves with triangle markers and circle markers, since $\beta_{ij} = \beta$, the diversity gains can be calculated using (27). These calculated values are $G_d = 2 \times 2 \times 2$ and $G_d = (2 \times 2)^2/(6^{-1} + 2^{-1} + 3^{-1} + 1^{-1})$, respectively. It can be seen that the BICMB curves in [7] which are shown in Fig. 2 with blue curves with no markers, have the same slope in high SNR, i.e., same diversity gain as the BICMB-PC for different setups.

It can be seen from Fig. 3 that changing the number of antennas at the RAUs does not affect the diversity gain. This confirms (27) where the diversity gain is independent of the number of antennas at each RAU. Furthermore, one can see that by doubling the number of resources here, i.e., the number of antennas at the transmitter or the receiver, the performance of the system gets better by a factor of 3 dB.

In Fig. 4 we compare the results of this paper with a con-
conventional MIMO system utilizing BICMB-PC with Rayleigh fading channel studied in [11]. In order to make these two comparable, we need to make the number of propagation paths in the channel to one, i.e., $L = 1$. Fig. 4 illustrates that for both cases when $D = 2$ and $D = 3$, the slope of the BER in high SNR for conventional MIMO system is the same as the the mm-wave model in (5).

Fig. 5 illustrates the effect of large scale fading coefficients on the diversity gain. We consider three different cases for BICMB-PC with $M_r = M_t = 2$, $L = 2$, and $N_g = 1$. Furthermore, in Case IV, we use BICMB in [7] where the diversity gain is $G_d \approx 4$ with $M_r = M_t = 1$, and $L = 4$. Let $B = [\beta_{ij}]$ where $\beta_{ij}$ expressed in dB, as the large scale fading coefficient matrix. We use the following matrices for the large scale fading coefficients for different cases:

$$B_1 = \begin{bmatrix} -23 & -23 \\ -23 & -23 \end{bmatrix}, \quad B_2 = \begin{bmatrix} -30 & -30 \\ -30 & -30 \end{bmatrix},$$

$$B_3 = \begin{bmatrix} 0 & -20 \\ -20 & -40 \end{bmatrix}, \quad B_4 = -20 \text{ dB}.$$

By using equation (27), one can see that the diversity gain for Case I and Case II is $G_d = 2 \times 2 \times 2 = 8$, whereas the diversity gain for Case III is $G_d = \left(10^{-7/2} \times 10^{-7/2} \times 10^{-7/2} \times 10^{-7/2}\right) \approx 4$. It can be seen from Fig. 5 that Case III has the same slope as Case IV, where both of them has a diversity gain of $G_d \approx 4$. One can see that in Case II, where the channel is inhomogeneous, the diversity gain decreases.

VI. CONCLUSION

In this work we showed that by utilizing BICMB-PC in a mm-wave MIMO system with DAS architecture, one can achieve full diversity gain. This means, the diversity gain is independent of the number of transmitted data streams and can be increased by increasing the number of RAUs at the transmitter or the receiver. We also show that the diversity gain is independent of the number of antennas at the RAUs in both the transmitter and the receiver. We leave the extension of the work to other STBC codes for future research.

REFERENCES