Precoding Design and PMI Selection for BICM-MIMO Systems with 5G New Radio Type-I CSI

Marjan Maleki, Juening Jin, Hao Wang, Martin Haardt

March 25, 2024

***Abstract*—This paper proposes novel linear precoding algorithms for Multiple-Input Multiple-Output Bit-Interleaved Coded Modulation (MIMO-BICM) systems that maximize the achievable rate subject to power constraints. To overcome the nonlinear and nonconvex nature of the optimization problem, we rewrite the achievable rate in terms of the log-likelihood ratio (LLR) and introduce manifold-based gradient ascent (MGA) precoding and low-complexity non-iterative algorithms. Simulation results show significant gains in achievable rate and block error rate compared to existing techniques. Additionally, we extend our investigation to linear precoding with the constraint that the precoding matrix is selected from the codebook type-I adopted in Fifth-Generation New Radio (5G NR) networks. We propose heuristic algorithms that exploit the Kronecker and Discrete Fourier Transform (DFT) structure of the codebook and consider the singular vector decomposition (SVD) precoder as the optimal reference precoder. The traditional exhaustive search methods require a high complexity, especially for large codebook sizes. However, our proposed algorithms apply a combination of direct estimation and a low-dimensional search for deriving the indices, resulting in a reduced number of codebook precoder candidates. Simulation results show that our proposed low-complexity algorithms perform comparably to exhaustive search baselines.**

***Index Terms*—Bit-Interleaved Coded Modulation (BICM), 5G New Radio (NR), Precoder Matrix Index (PMI), channel state information (CSI), Type-I Codebook.**

I. Introduction

Bit-Interleaved Coded Modulation (BICM) is a reliable data transmission scheme that combines error-correction coding and modulation, separated by an interleaving permutation. BICM is widely adopted in modern communication systems such as wireless communications and satellite communications standards, including High Speed Packet Access (HSPA) , IEEE 802.11a/g/n, Long-Term Evolution (LTE), and the latest Digital Video Broadcasting (DVB) standards (DVB-T2/S2/C2). It has also been used in the design of Fifth-Generation New Radio (5G NR) systems and is likely to become the de facto choice for most future standards.

The BICM decoder treats the transmission as multiple parallel bit channels, with the symbol decoding metric being the product of the bit decoding metrics. The achievable rates can be evaluated using generalized mutual information (GMI), which is the sum of the individual bit channel's GMI [1]. The modulation and coding scheme (MCS) chosen for BICM affects the achievable rate, which is a trade-off between error-correction capability and data transmission rate [2]. In single-antenna systems, BICM with Gray labeling has been shown to significantly improve system performance and can approach the theoretical limit of channel capacity [3], [4]. The success of BICM in single-antenna systems motivated researchers to extend this technique to Multiple-Input Multiple-Output (MIMO) systems [5], [6]. Today, MIMO-BICM has become an essential component of 3GPP LTE and 5G NR systems, providing high data transmission rates and reliable communication over wireless channels.

Linear precoding is a crucial signal processing technique that enhances the performance of MIMO systems by leveraging channel knowledge to preprocess signals before transmission. Precoders are typically designed from an information-theoretical perspective, where the capacity or mutual information is utilized as the criteria for optimization. Current information-theoretical approaches for precoding design can be categorized into two groups: 1) capacity-based design, which uses the channel capacity for precoding design [7]. However, these designs are based on the impractical Gaussian input assumption, resulting in a significant performance loss when applied to practical systems that use discrete constellations, such as quadrature amplitude modulation (QAM). 2) Mutual information-based design, which maximizes the mutual information between a finite-alphabet input and the corresponding output [8]–[11]. While the maximum mutual information can be achieved by multi-level coding and multi-stage decoding techniques [12], the high complexity of multi-stage decoding renders it impractical for use in most applications. Therefore, for practical MIMO-BICM systems, the maximum input-output mutual information is not achievable, and the corresponding designed precoder based on the mutual information is far from optimal.

The area of precoding in MIMO-BICM systems has received relatively scant exploration. Key references, such as [13], have introduced linear precoding approaches for achieving full diversity while minimizing receiver complexity using an iterative decoder. Meanwhile, [14] introduced BICM Beamforming (BICMB), combining BICM with singular vector decomposition (SVD) beamforming under specific criteria on code rate and spatial stream to achieve full diversity. Authors in [15] demonstrated that by employing a constellation precoding technique, these conditions can be overridden, enabling the simultaneous achievement of full diversity and spatial multiplexing. Similarly, [16] designed a partial

M. Maleki and M. Haardt are with the Communications Research Laboratory, Ilmenau University of Technology, Ilmenau, Germany. E-mails: {marjan.maleki, martin.haardt}@tu-ilmenau.de.

J. Jin and H. Wang are with Huawei Technologies Co. Ltd., Beijing, China. E-mails:{jinjuening, hunter.wanghao}@hisilicon.com

algebraic precoder for transmitting multiple spatial streams with full diversity over a transmit beamformed MIMO-BICM channel. Our study innovates by presenting a novel method for designing optimal linear precoders for MIMO-BICM systems, focusing on maximizing the achievable BICM mutual information using finite-alphabet input constellations. This marks our work as a pioneering study from an information-theoretical perspective in this domain, to the best of our knowledge.

In Closed-loop MIMO systems, beamforming gains are achieved by leveraging channel state information (CSI) feedback from the user equipment (UE) to adapt precoding strategies based on the downlink CSI [17]. Codebook-based precoding, which involves reporting the precoder matrix indicator (PMI) to the gNodeB (gNB), is used in 5G NR systems to reduce the feedback complexity and the overhead. A brute-force search is typically used to select the precoder that maximizes the channel capacity or another related metric by searching through the codebook for the best PMI. However, this search becomes prohibitively complex as MIMO dimensions increase due to the codebook size scaling with the number of antennas.

Various PMI selection techniques have been proposed to reduce the complexity of exhaustive search methods. Some examples include a link adaptation and a PMI selection method based on mutual information maximization in [18], while [19] considers joint PMI and rank indicator (RI) selection using the minimum mean-square error (MMSE) equalizer. Another approach, presented in [20], estimates the optimal PMI directly from the singular vectors of the channel matrix by exploiting the Discrete Fourier Transform (DFT) structure, utilizing an iterative linear phase estimation (ILPE). However, these methods were designed for LTE and legacy systems and may not meet the requirements of 5G NR. To overcome this limitation, authors in [21] proposes a neural network-based approach for direct precoder matrix selection from the Type-I codebook in 5G NR, with the objective of maximizing the corresponding channel capacity. In a recent study [22], the proposed algorithms exploit the Kronecker and DFT structure of the codebook to efficiently find the optimal PMI indices in MIMO-BICM systems. The algorithms utilize a combination of direct estimation and low-dimensional search techniques, resulting in improved PMI selection performance.

Inspired by this, the aim of the second part of this paper is to expand the scope of precoding design in BICM-MIMO systems by introducing the constraint of selecting the precoding matrix from a predefined codebook. We present low-complexity algorithms that facilitate the selection of the optimal precoding matrix and its corresponding PMI indices from the Type-I codebook in 5G NR. Compared to the PMI selection proposed in [22], our method achieves lower computational complexity by employing a one-dimensional search for the singular-vector (SV)-based approach, as opposed to the two-dimensional search stage in [22]. In summary, the key contributions of this paper are as follows:

- A manifold-based gradient ascent (MGA) algorithm is proposed to maximize the achievable rate of MIMO-BICM under a power constraint. The algorithm takes into account the channel matrix, the signal-to-noise ratio (SNR), discrete alphabet constellation, and the soft MIMO detector used to compute the log-likelihood ratio (LLR).
- A non-iterative algorithm is introduced that utilizes the mathematical structure of the optimal linear precoder to design near-optimal precoders for MIMO-BICM with low complexity.
- The benefit of applying the proposed MGA and low-complexity algorithms in a MIMO system with LDPC codes is that maximizing the achievable MIMO-BICM rate is an effective approach to minimize the block error rate (BLER), leading to improved overall system performance.
- We introduce novel algorithms that efficiently select PMI from the Type-I codebook, aiming to maximize BICM rate while keeping computational complexity low. To address the challenges arising from discrete variables and the absence of a closed-form objective function, we propose a heuristic two-stage SV-based procedure for PMI selection, leveraging the channel covariance matrix.
- The proposed PMI selection methods are evaluated through computer simulations, employing MIMO correlated channels with cross-polarized antennas as per [23], and a realistic non-line-of-sight (NLOS) propagation model compliant with 3GPP TR 38.901 [24]. Performance comparisons are made with exhaustive search based on the BICM rate and MMSE capacity metrics.

The remainder of the paper is organized as follows. Section II presents the system model and problem statement. In Section III, we propose two precoding design techniques for MIMO-BICM systems without codebook constraints, namely the MGA and the low complexity precoding design. In Section IV, we extend our investigation to the codebook-based precoding design and discuss the Type-I codebook single-panel structure. Then, we present novel algorithms for selecting the best PMI from the Type-I codebook in terms of the achievable rate for MIMO-BICM systems. We also provide a computational complexity analysis of the proposed PMI selection algorithms. Section V presents simulation results, and finally, Section VI concludes the paper.

II. System model and problem statement

We consider a 5G NR system operating in the Frequency Range 1 (FR1), where a gNB with M_T antennas transmits N_L layers to a receiver equipped with M_R receive antennas. The antenna panel consists of N_1 horizontal antennas and N_2 vertical antennas with cross-polarization. Hence, the total number of transmit antenna ports for each panel is $M_T = 2N_1N_2$. The payload bit-stream is segmented and processed by a BICM scheme with a capacity-approaching code like LDPC. The encoded bit-stream is divided into code blocks (consisting of up to 8448 bits), interleaved, and then mapped to a symbol constellation, e.g., QAM via the Gray mapping. At the next stage, the data symbols are demultiplexed into N_L layers, and each layer is further mapped to the BS antennas by a precoding scheme. Assuming linear precoding, the frequency-domain received signal on subcarrier k can be expressed as

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{P}_k \boldsymbol{s}_k + \boldsymbol{n}_k \in \mathbb{C}^{M_R}, \tag{1}$$

where $\boldsymbol{H}_k \in \mathbb{C}^{M_R \times M_T}$ represents the MIMO channel matrix, $\boldsymbol{P}_k \in \mathbb{C}^{M_T \times N_L}$ is the linear precoding matrix, and $\boldsymbol{n}_k \in \mathbb{C}^{M_R}$ denotes the additive noise vector with zero mean and covariance matrix $\sigma^2 \boldsymbol{I}_{M_R}$, for $k \in \{1, ..., K\}$. The transmitted symbol vector $\boldsymbol{s}_k \in \mathcal{S}^{N_L}$ is taken from the symbol alphabet $\mathcal{S}$.

To account for the MCS in the precoder optimization, the optimization metrics, such as BLER and BICM mutual information, may be considered. We utilize mutual information as the optimization metric for efficient precoder design in MIMO-BICM systems. This metric offers a comprehensive evaluation, accounting for modulation schemes, channel conditions and noise levels, ensuring overall link capacity and robustness. The notable advantage of utilizing BICM mutual information is its effectiveness in assessing the BLER in a comprehensive communication system equipped with capacity-achieving error-correction codes. Through our simulations, we illustrate that selecting BICM mutual information as the optimization criterion ensures that the resulting precoder, is nearly optimal for practical coded systems, thereby minimizing BLER.

For notational simplicity, we drop the subscript k, in other words we consider a single carrier system without loss of generality and the generalization to multi-carrier transmission will be discussed later. By using the optimum soft detection at the receiver, we use the channel observation $\boldsymbol{y}$ along with a-priori information $P(\boldsymbol{s})$ to generate the intrinsic log-likelihood ratio (LLR) value for each b_{iq}, which denotes the q^{th} bit in the i^{th} stream. By defining the equivalent channel $\bar{\boldsymbol{H}} \triangleq \boldsymbol{H}\boldsymbol{P}$, for a given constellation $\mathcal{S}$, we have

$$L_{iq} = \ln \frac{p\left(b_{iq}=1 \middle| \boldsymbol{y}, \bar{\boldsymbol{H}}\right)}{p\left(b_{iq}=0 \middle| \boldsymbol{y}, \bar{\boldsymbol{H}}\right)} = \ln \frac{\sum\limits_{\boldsymbol{s}\in\mathcal{S}_{iq}^{1}} \exp\left(-\frac{\|\boldsymbol{y}-\bar{\boldsymbol{H}}\boldsymbol{s}\|^2}{\sigma^2}\right)}{\sum\limits_{\boldsymbol{s}\in\mathcal{S}_{iq}^{0}} \exp\left(-\frac{\|\boldsymbol{y}-\bar{\boldsymbol{H}}\boldsymbol{s}\|^2}{\sigma^2}\right)}, \quad (2)$$

where $p\left(b_{iq} \middle| \boldsymbol{y}, \bar{\boldsymbol{H}}\right)$ is the conditional probability mass function of b_{iq} given $\left(\boldsymbol{y}, \bar{\boldsymbol{H}}\right)$, $\mathcal{S}_{iq}^{b}$ denotes the set of transmit vectors given $b_{iq} = b \in \{0, 1\}$.

The calculation of (2) becomes exponentially complex as the number of data streams N_L increases. However, one can adopt the max-log approximation $\log \sum\limits_{k} \exp(x_k) \approx \max\limits_{k} x_k$, to simplify the LLR computation. This results in the following approximation of the LLR:

$$\tilde{L}_{iq} = \frac{1}{\sigma^2}\left[\min_{\boldsymbol{s}\in\mathcal{S}_{iq}^{0}} \left\|\boldsymbol{y}-\bar{\boldsymbol{H}}\boldsymbol{s}\right\|^2 - \min_{\boldsymbol{s}\in\mathcal{S}_{iq}^{1}} \left\|\boldsymbol{y}-\bar{\boldsymbol{H}}\boldsymbol{s}\right\|^2\right]. \quad (3)$$

By avoiding logarithmic and exponential operations, the approximate LLR in (3) can be computed efficiently using sphere decoding algorithms [25]- [26]. Afterwards, the estimated symbols based on LLR values are demodulated, deinterleaved, and decoded, yielding the estimated payload bitstream.

For a constellation with Q bits per symbol and a cardinality of $|\mathcal{S}| = 2^Q$, the MIMO-BICM system is represented as a set of $N_L Q$ independent parallel binary-input channels. Assuming i.i.d. uniform code bits, we obtain the mutual information between bits and the corresponding LLRs as

$$R_{\text{BICM}} = \sum_{i=1}^{N_L}\sum_{q=1}^{Q} \mathcal{I}(b_{iq}; \boldsymbol{y}|\bar{\boldsymbol{H}}) = \sum_{i=1}^{N_L}\sum_{q=1}^{Q} \left(1 - \mathcal{H}\left(b_{iq}\middle|\boldsymbol{y}, \bar{\boldsymbol{H}}\right)\right), \quad (4)$$

where $\mathcal{H}\left(b_{iq}\middle|\boldsymbol{y}, \bar{\boldsymbol{H}}\right)$ is the conditional entropy of b_{iq} given $\boldsymbol{y}$ and $\bar{\boldsymbol{H}}$. According to the definition in (2), the probability mass function $p(b_{iq}|\boldsymbol{y}, \bar{\boldsymbol{H}})$ can be expressed as the function of the LLR

$$p(b_{iq}|\boldsymbol{y}, \bar{\boldsymbol{H}}) = \frac{1}{1 + e^{(1-2b_{iq})L_{iq}}}. \quad (5)$$

Therefore, the conditional entropy is given as

$$\mathcal{H}\left(b_{iq}\middle|\boldsymbol{y}, \bar{\boldsymbol{H}}\right) = \mathbb{E}\left(\log_2\left(1 + e^{(1-2b_{iq})L_{iq}}\right)\right), \quad (6)$$

where the expectation is over (b_{iq}, L_{iq}). By substituting (6) into (4) and leveraging the approximate LLR provided in (3), we can derive the approximate achievable rate as follows:

$$\tilde{R}_{\text{BICM}} = N_L Q - \mathbb{E}\left(\sum_{i=1}^{N_L}\sum_{q=1}^{Q} \log_2\left(1 + e^{(1-2b_{iq})\tilde{L}_{iq}}\right)\right). \quad (7)$$

With the assumption of $E\left(\boldsymbol{s}\boldsymbol{s}^H\right) = \boldsymbol{I}_{N_L}$, the average transmit power at the transmitter will be $\mathbb{E}\left(\|\boldsymbol{P}\boldsymbol{s}\|^2\right) = \text{tr}\left(\boldsymbol{P}^H\boldsymbol{P}\right)$. The precoding design problem is approached with the goal of maximizing the achievable rate of MIMO-BICM system while maintaining a power constraint. The formulation, based on the achievable rate with approximate LLR as indicated in Eq. (7), is presented below:

$$\begin{aligned} \max_{\boldsymbol{P}} \quad & \tilde{R}_{\text{BICM}} \\ \text{s.t.} \quad & \text{tr}\left(\boldsymbol{P}^H\boldsymbol{P}\right) = \gamma \\ & \boldsymbol{P} \in \boldsymbol{\mathcal{W}}, \end{aligned} \quad (8)$$

where γ is the maximum transmit power and $\boldsymbol{\mathcal{W}}$ is the type-I codebook. We investigate the BICM maximization problem in two scenarios. First, we drop the codebook constraint and consider the problem in a general setting, as discussed in Section III. Then, in Section IV, we further extend our study by assuming that the precoder matrix is constrained to be chosen from the type-I codebook, and we aim to find the best PMI for low-dimensional CSI feedback under this constraint.

We can compare problem (8) with an existing MIMO precoding problem with finite-alphabet inputs [9]. In the coded modulation (CM) scheme, a linear precoder is designed by maximizing the input-output mutual information, denoted by R_{CM}, given by

$$R_{\text{CM}} = \mathcal{I}(\boldsymbol{s}; \boldsymbol{y}|\bar{\boldsymbol{H}}) = \mathcal{I}(\{b_{ik}\}_{ik}; \boldsymbol{y}|\bar{\boldsymbol{H}}). \quad (9)$$

Comparing (4) and (9), we can conclude that $R_{\text{BICM}} \leq R_{\text{CM}}$, as BICM receivers neglect the dependencies between code bits. The gap between R_{BICM} and R_{CM} increases with the condition number of the channel and the cardinality of the constellation, as the correlation between code bits increases. Therefore, the optimal precoder designed by maximizing R_{CM} may not be optimal for MIMO-BICM systems.

III. Precoder Design without Codebook Constraint

In this section, we present two approaches for solving the BICM rate maximization problem, to design the precoder without the codebook constraint. The first solution is the MGA approach, which can provide a high-accuracy solution at the cost of increased computational complexity. The second solution is based on the SVD of the channel matrix which provides a near-optimal performance while reducing the computational complexity.

A. Manifold-Based Gradient Ascent (MGA) Precoding Design

In this section, we will focus on developing numerical algorithms for problem (8) by first deriving the complex gradient of $\tilde{R}_{\text{BICM}}$ with respect to the precoder $\boldsymbol{P}$. To do this, we start by expressing the approximate LLR in (3) as the difference of two minimum distance problems. Let $\boldsymbol{s}_{iq}^0$ and $\boldsymbol{s}_{iq}^1$ be the solutions to the following minimum distance problems, respectively:

$$\begin{aligned}\boldsymbol{s}_{iq}^0 &= \underset{\boldsymbol{s}\in\mathcal{S}_{iq}^0}{\arg\min}\ \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}\|^2,\\ \boldsymbol{s}_{iq}^1 &= \underset{\boldsymbol{s}\in\mathcal{S}_{iq}^1}{\arg\min}\ \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}\|^2.\end{aligned} \tag{10}$$

Then, the approximate LLR value can be expressed as

$$\tilde{L}_{iq} = \frac{1}{\sigma^2}\left[\left\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^0\right\|^2 - \left\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^1\right\|^2\right]. \tag{11}$$

Suppose that the precoder is updated from $\boldsymbol{P}$ to $\boldsymbol{P}+\Delta\boldsymbol{P}$, and the solutions of minimum distance problems $\boldsymbol{s}_{iq}^0$ and $\boldsymbol{s}_{iq}^1$ remain unchanged as long as $\Delta\boldsymbol{P}\to\boldsymbol{0}$. Based on this assumption, we can derive the complex gradient of $\tilde{L}_{iq}$, which is provided in the following lemma.

Lemma 1: The complex gradient of $\tilde{L}_{iq}$ with respect to $\boldsymbol{P}$ is given by

$$\begin{aligned}\frac{\partial\tilde{L}_{iq}}{\partial\boldsymbol{P}^*} = \frac{1}{\sigma^2}\boldsymbol{H}^H\big[&(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^0)(\boldsymbol{s}-\boldsymbol{s}_{iq}^0)^H\\ &-(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^1)(\boldsymbol{s}-\boldsymbol{s}_{iq}^1)^H\big],\end{aligned} \tag{12}$$

where $\boldsymbol{s}$ is the transmit vector and $\boldsymbol{y}$ is the receive vector.

Proof: See Appendix A. ∎

The following theorem provides the complex gradient of $\tilde{R}_{\text{BICM}}$ with respect to $\boldsymbol{P}$, utilizing Lemma 1.

Theorem 1: The complex gradient of $\tilde{R}_{\text{BICM}}$ with respect to $\boldsymbol{P}$ is given by

$$\frac{\partial\tilde{R}_{\text{BICM}}}{\partial\boldsymbol{P}^*} = \mathbb{E}\left[\boldsymbol{G}(\boldsymbol{s},\boldsymbol{y})\right]. \tag{13}$$

Here, $\boldsymbol{G}(\boldsymbol{s},\boldsymbol{y})$ is defined as

$$\begin{aligned}\boldsymbol{G}(\boldsymbol{s},\boldsymbol{y}) = \frac{1}{\sigma^2}\boldsymbol{H}^H\Big[&(\boldsymbol{y}\boldsymbol{1}_{M_R}^T-\boldsymbol{H}\boldsymbol{P}\boldsymbol{S}^0)\boldsymbol{\Lambda}(\boldsymbol{s}\boldsymbol{1}_{M_R}^T-\boldsymbol{S}^0)^H\\ &-(\boldsymbol{y}\boldsymbol{1}_{M_R}^T-\boldsymbol{H}\boldsymbol{P}\boldsymbol{S}^1)\boldsymbol{\Lambda}(\boldsymbol{s}\boldsymbol{1}_{M_R}^T-\boldsymbol{S}^1)^H\Big],\end{aligned} \tag{14}$$

where $\boldsymbol{1}_{M_R}$ is the column vector of length M_R with all elements equal to one, $\boldsymbol{S}^0 = \left[\boldsymbol{s}_{11}^0,\boldsymbol{s}_{12}^0,...,\boldsymbol{s}_{N_{\text{L}}Q}^0\right]$, $\boldsymbol{S}^1 = \left[\boldsymbol{s}_{11}^1,\boldsymbol{s}_{12}^1,...,\boldsymbol{s}_{N_{\text{L}}Q}^1\right]$, and $\boldsymbol{\Lambda}$ is a diagonal matrix. Let $p=(i-1)Q+q$, the pth diagonal element of $\boldsymbol{\Lambda}$ is given as

$$\boldsymbol{\Lambda}_{pp} = -\frac{1}{\ln(2)}\cdot\frac{(1-2b_{iq})\exp\left[(1-2b_{iq})\tilde{L}_{iq}\right]}{1+\exp\left[(1-2b_{iq})\tilde{L}_{iq}\right]}. \tag{15}$$

Proof: To compute the complex gradient of $\tilde{R}_{\text{BICM}}$, we apply the chain rule in differentiation. Specifically, we first obtain the complex gradient of $\tilde{R}_{\text{BICM}}$ with respect to $\tilde{L}_{iq}$. Next, we utilize the complex gradient of $\tilde{L}_{iq}$ with respect to the precoder $\boldsymbol{P}$, which is obtained from Lemma 1. Therefore, the proof is concluded. ∎

The expectation in (13) does not have a closed-form expression, and therefore the gradient needs to be evaluated using the Monte Carlo method[1]. Given a large number of N independent samples $\{\boldsymbol{s}^n,\boldsymbol{y}^n\}_{n=1}^N$ generated by (1), an estimate of the complex gradient in (13), denoted as $\boldsymbol{\Omega}$, can be obtained by using the sample average as follows:

$$\boldsymbol{\Omega} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{G}(\boldsymbol{s}^n,\boldsymbol{y}^n). \tag{16}$$

Since $\mathbb{E}(\boldsymbol{\Omega}) = \frac{\partial\tilde{R}_{\text{BICM}}}{\partial\boldsymbol{P}^*}$ and the variance of $\boldsymbol{\Omega}$ decreases quadratically with increasing N due to the central limit theorem, we can use the noisy gradient $\boldsymbol{\Omega}$ to design an efficient algorithm.

To solve problem (8) using noisy gradient information, we propose a manifold-based gradient ascent (MGA) algorithm. At the ℓth iteration, the algorithm updates the current solution $\boldsymbol{P}_\ell$ to $\boldsymbol{P}_{\ell+1}$ using the following rule:

$$\boldsymbol{P}_{\ell+1} = \text{Proj}\left[\boldsymbol{P}_\ell + \frac{\mu}{\sqrt{\ell}}\nabla_{\boldsymbol{P}}\tilde{R}_{\text{BICM}}\right], \tag{17}$$

where $\frac{\mu}{\sqrt{\ell}}$ is the step-size, $\text{Proj}(\boldsymbol{P})$ is given by

$$\text{Proj}(\boldsymbol{P}) = \sqrt{\frac{\gamma}{\text{Tr}(\boldsymbol{P}^H\boldsymbol{P})}}\boldsymbol{P}, \tag{18}$$

and $\nabla_{\boldsymbol{P}}\tilde{R}_{\text{BICM}}$ is the gradient of $\tilde{R}_{\text{BICM}}$ with respect to $\boldsymbol{P}$ on the sphere manifold $\mathcal{M}=\{\boldsymbol{P}|\text{Tr}(\boldsymbol{P}^H\boldsymbol{P})=\gamma\}$. To obtain $\nabla_{\boldsymbol{P}}\tilde{R}_{\text{BICM}}$, we project $\boldsymbol{\Omega}$ onto the tangent space of $\mathcal{M}$. This can be achieved by solving the optimization problem:

$$\nabla_{\boldsymbol{P}}\tilde{R}_{\text{BICM}} = \underset{\boldsymbol{G}\in\mathcal{T}_{\mathcal{M}}}{\arg\min}\ \|\boldsymbol{\Omega}-\boldsymbol{G}\|^2, \tag{19}$$

where the tangent space $\mathcal{T}_{\mathcal{M}}$ is defined as $\mathcal{T}_{\mathcal{M}} = \left\{\boldsymbol{G}|\text{Tr}(\boldsymbol{G}^H\boldsymbol{P}+\boldsymbol{P}^H\boldsymbol{G})=0\right\}$. Using the Lagrangian multiplier method, we compute the closed-form expression of $\nabla_{\boldsymbol{P}}\tilde{R}_{\text{BICM}}$ as

$$\nabla_{\boldsymbol{P}}\tilde{R}_{\text{BICM}} = \boldsymbol{\Omega} - \frac{\text{Re}\left\{\text{Tr}\left[\boldsymbol{\Omega}^H\boldsymbol{P}\right]\right\}}{\gamma}\boldsymbol{P}. \tag{20}$$

The complete MGA algorithm is summarized in Algorithm 1.

[1]This process can occur at either the UE (FDD mode) or gNB (TDD mode). By generating multiple independent samples $\{\boldsymbol{s}_k\}$ and simultaneously computing $\{\boldsymbol{y}_k\}$ via equation (1), with independent Gaussian noises $\{\boldsymbol{n}_k\}$ generated by the Box-Muller method, there's no need to regenerate samples when the channel or precoder changes. Instead, the UE/gNB can store and reuse independent samples $\{\boldsymbol{s}_k,\boldsymbol{n}_k\}$ to compute $\{\boldsymbol{y}_k\}$.

Algorithm 1 MGA Precoding

1. **Initialization**: Generate an initial precoder $\boldsymbol{P}_1$.
2. **for** $\ell = 1$ **to** $\ell_{\max}$ **do**
   - Generate N independent samples $\{\boldsymbol{s}^n, \boldsymbol{y}^n\}$ with $\boldsymbol{P}_\ell$.
   - Compute $\boldsymbol{\Omega}$ at $\boldsymbol{P}_\ell$ by (16).
   - Compute $\nabla_{\boldsymbol{P}} \tilde{R}_{\mathrm{BICM}}$ by (20).
   - Update $\boldsymbol{P}_\ell$ to $\boldsymbol{P}_{\ell+1}$ according to (17).

   end for
3. **Output**: $\boldsymbol{P}_{\ell_{\max}}$.

B. Low Complexity Precoding Design

In this subsection, we explore the structure of the optimal precoder and propose a low-complexity non-iterative algorithm to design a near-optimal precoder. Given that optimal precoders with discrete input alphabets are non-diagonal, requiring precoding across all subchannels [8], our approach involves a two-stage precoding strategy. Initially, we utilize a reference SVD-based precoder known for maximizing capacity. Subsequently, we refine it with a second precoding submatrix, implemented through rotation matrices. We specifically consider the X-code structure for this second stage, enhancing diversity and optimizing it to maximize the BICM rate.

Let the singular value decomposition (SVD) of $\boldsymbol{H}$ be

$$\boldsymbol{H} = \boldsymbol{K}\boldsymbol{\Sigma}\boldsymbol{L}^H, \tag{21}$$

where $\boldsymbol{K} \in \mathbb{C}^{M_R \times \mathrm{rank}(\boldsymbol{H})}$ is a unitary matrix with left singular vectors, $\boldsymbol{\Sigma} \in \mathbb{C}^{\mathrm{rank}(\boldsymbol{H}) \times \mathrm{rank}(\boldsymbol{H})}$ is a diagonal matrix with singular values arranged in decreasing order, and $\boldsymbol{L} \in \mathbb{C}^{M_T \times \mathrm{rank}(\boldsymbol{H})}$ is a unitary matrix of right singular vectors. To start, we state the following proposition as a direct application of [27, Proposition 1].

Proposition 1: Given the SVD of the precoder matrix, the left singular vectors of the optimal precoder $\boldsymbol{P}$ can always be chosen to be the first N_L columns of $\boldsymbol{L}$.

Based on the above proposition, we factorize $\boldsymbol{P}$ into the following form

$$\boldsymbol{P} = \boldsymbol{L}(:, 1\!:\!N_L)\boldsymbol{\Theta}, \tag{22}$$

where $\boldsymbol{\Theta} \in \mathbb{R}^{N_L \times N_L}$ is an orthogonal matrix. Then, the equivalent system model is given by

$$\bar{\boldsymbol{y}} = \boldsymbol{K}^H \boldsymbol{y} = \boldsymbol{\Sigma}_{\mathrm{H}} \boldsymbol{\Theta} \boldsymbol{x} + \boldsymbol{K}^H \boldsymbol{n}. \tag{23}$$

The generation of the orthogonal matrix $\boldsymbol{\Theta}$ is based on the Givens rotation matrix, which is designed to decompose (23) into $\lfloor \frac{N_L}{2} \rfloor$ parallel 2×2 MIMO subsystems [28]. Inspired by optimal diversity order strategies in [28], our approach pairs the k'-th subchannel with the $(N_L - k' + 1)$-th subchannel for $k' \in \{1, 2, ..., \lfloor \frac{N_L}{2} \rfloor\}$. This strategy pairs a strong subchannel with a weaker one, followed by symbol rotation by an angle θ, to leverage the stronger channel for enhanced detection of symbols from the weaker channel. Therefore, each pair of (p, q) corresponds to a subsystem of (p, q)-th data streams which are precoded together by the submatrix $\boldsymbol{\Theta}^{(pq)} \in \mathbb{R}^{2\times 2}$. This submatrix is formed by selecting the (p, q)th rows and columns of $\boldsymbol{\Theta}$, i.e.,

$$\boldsymbol{\Theta}^{(pq)} = \begin{bmatrix} \Theta_{pp} & \Theta_{pq} \\ \Theta_{qp} & \Theta_{qq} \end{bmatrix}, \tag{24}$$

where $\boldsymbol{\Theta}_{pq}$ is the (p, q)th element of $\boldsymbol{\Theta}$. To parameterize the 2×2 orthogonal submatrix $\boldsymbol{\Theta}^{(pq)}$, we use the following form:

$$\boldsymbol{\Theta}^{(pq)} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \tag{25}$$

where $\theta \in [0, \frac{\pi}{4}]$ is a parameter that needs to be optimized. Therefore, the orthogonal matrix $\boldsymbol{\Theta}$ depends on a single parameter θ. To derive the optimal θ for a 2^Q–QAM constellation, a set of 2×2 channels and noise variances, denoted as $\{\boldsymbol{H}_i, \sigma_i^2\}_{i=1}^N$, is generated. Subsequently, for each $\{\boldsymbol{H}_i, \sigma_i^2\}$, a one-dimensional search is employed to find the optimal θ that maximizes the BICM mutual information:

$$\theta_i^\star = \arg\max_{\theta_i} \quad \tilde{R}_{\mathrm{BICM}}(\theta_i | \boldsymbol{H}_i, \sigma_i^2). \tag{26}$$

Here the notation $\tilde{R}_{\mathrm{BICM}}(\theta_i | \boldsymbol{H}_i, \sigma_i^2)$ denotes that $\tilde{R}_{\mathrm{BICM}}$ is a function of θ_i, given $\{\boldsymbol{H}_i, \sigma_i^2\}$. The corresponding optimal code rate is then defined as:

$$r_i^\star = \frac{\tilde{R}_{\mathrm{BICM}}(\theta_i^\star | \boldsymbol{H}_i, \sigma_i^2)}{2Q}.$$

Analyzing the data $\{\theta_i^\star, r_i^\star\}_{i=1}^N$, three key observations are made:

- When $r_i^\star$ is below a threshold β_1, the optimal angle $\theta_i^\star$ is nearly zero.
- When $r_i^\star$ is above a threshold β_2, the optimal angle $\theta_i^\star$ is nearly a constant.
- When $\beta_1 < r_i^\star < \beta_2$, the optimal angle $\theta_i^\star$ is monotonically increasing with respect to $r_i^\star$.

Based on these observations, we determine β_1 and β_2 from the data. A quadratic equation models the relationship between $\theta_i^\star$ and $r_i^\star$ when $\beta_1 < r_i^\star < \beta_2$, given by:

$$\theta_i^\star = a_1 (r_i^\star)^2 + a_2 r_i^\star + a_3, \quad i = 1, 2, ..., N$$

Expressing these equations in the following matrix form allows us to determine the unknown parameters $\{a_1, a_2, a_3\}$ directly through the least squares method:

$$\begin{bmatrix} \theta_1^\star \\ \vdots \\ \theta_N^\star \end{bmatrix} = \begin{bmatrix} (r_1^\star)^2 & r_1^\star & 1 \\ \vdots & \vdots & \vdots \\ (r_N^\star)^2 & r_N^\star & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}. \tag{27}$$

This shows that the optimal orthogonal matrix $\boldsymbol{\Theta}(\theta)$ is mainly determined by the modulation order and code rate. The optimal values of θ as a function of the code rate range, are given in Table I for different modulation schemes of QPSK, 16QAM, and 64QAM.

Given the code rate, we can compute the optimal θ and the corresponding near-optimal precoder $\boldsymbol{P} = \boldsymbol{L}(:, 1\!:\!N_L)\boldsymbol{\Theta}(\theta)$. If the code rate is unknown, we can use an initial $\theta_0 = 0.2$ to estimate the code rate $r_0 = \frac{\tilde{R}_{\mathrm{BICM}}}{N_L Q}$ at $\boldsymbol{P} = \boldsymbol{L}(:, 1\!:\!N_L)\boldsymbol{\Theta}(\theta_0)$, and then compute the optimal θ and the corresponding optimal precoder based on r_0.

Table I: Optimal value of θ for different code rates and modulation orders

QPSK		16QAM		64QAM	
$r \in [0, 0.4)$	$\theta = 0$	$r \in [0, 0.47)$	$\theta = 0$	$r \in [0, 0.58)$	$\theta = 0$
$r \in [0.4, 0.69]$	$\theta = -3.88r^2 + 5.43r - 1.46$	$r \in [0.47, 0.62]$	$\theta = -0.84r^2 + 2.44r - 0.95$	$r \in [0.58, 0.8]$	$\theta = -4.54r^2 + 7.3r - 2.7$
$r \in (0.69, 1]$	$\theta = 0.44$	$r \in (0.62, 1]$	$\theta = 0.24$	$r \in (0.8, 1]$	$\theta = 0.235$

IV. The codebook-based precoding design

To solve the optimization problem (8) with the codebook constraint, an exhaustive search over all precoder candidates in the codebook is required. However, evaluating the BICM rate using Monte Carlo simulations or numerical methods can be computationally expensive. To address this, capacity and signal-to-interference plus noise-ratio (SINR) are commonly used as quality indicators and the MIMO demodulators with lower complexity, consisting of a linear equalizer followed by per-layer scalar soft demodulators, can also be used for decoding. This approach has been studied using minimum mean-square error (MMSE) and zero-forcing (ZF) equalization [29]. We propose using MMSE equalization to decouple the layers and assess performance by considering the sum of mutual information over N_L sub-channels. Utilizing the orthogonality principle, the MMSE equalizer is given by

$$\boldsymbol{G} = \left(\boldsymbol{P}^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{P} + \sigma^2 \boldsymbol{I}_{N_L}\right)^{-1} \boldsymbol{P}^H \boldsymbol{H}^H, \tag{28}$$

and the corresponding post-MMSE SINR of the l^{th} output layer is expressed as [30]

$$\gamma_l = \frac{1}{\sigma^2 \left[\left(\boldsymbol{P}^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{P} + \sigma^2 \boldsymbol{I}_{N_L}\right)^{-1}\right]_{l,l}} - 1, \tag{29}$$

where, $[.]_{l,l}$ denotes l^{th} diagonal element of the inverse matrix. Therefore, the achievable capacity in the equalized l^{th} sub-channel is computed as

$$C_{\text{MMSE},l} = \log_2(1 + \gamma_l). \tag{30}$$

The sum of the achievable capacity over N_L sub-channels could serve as a cost function for searching the codebook, and our baseline method for selecting the precoder from the codebook can be given as

$$\begin{aligned} \max_{\boldsymbol{P}} \quad & \sum_{l=1}^{N_L} \log_2(1 + \gamma_l) \\ \text{s.t.} \quad & \text{tr}\left(\boldsymbol{P}^H \boldsymbol{P}\right) = \gamma \\ & \boldsymbol{P} \in \boldsymbol{\mathcal{W}}. \end{aligned} \tag{31}$$

The PMI of the optimal precoder, which maximizes the post-equalization mutual information, will be reported for the CSI feedback. We will consider the optimization problem with the BICM rate criterion maximization and propose a lower complexity solution in the next subsections. We will also compare the performance of our proposed method with the exhaustive search method used to solve the MMSE capacity maximization problem, which serves as a baseline.

A. The codebook type-I single-panel structure

The precoder matrix $\boldsymbol{W}$ for type-I single-panel CSI can be expressed as $\boldsymbol{W} = \boldsymbol{W}_1 \boldsymbol{W}_2$, with $\boldsymbol{W}_1 \in \mathbb{C}^{2N_1 N_2 \times 2L}$ and $\boldsymbol{W}_2 \in \mathbb{C}^{2L \times \nu}$, where $L = \lceil \frac{\nu}{2} \rceil$. The matrix $\boldsymbol{W}_1$ targets wideband and long-term channel properties, while $\boldsymbol{W}_2$ represents the subband and frequency dependent part of the channel. Type-I codebooks support precoding matrices up to rank 8, reported as the rank indicator $\nu \in \{1, 2, ..., 8\}$ to the gNB [31].

The matrix $\boldsymbol{W}_1$ defines a beam or group of beams pointing in various directions and can be expressed as $\boldsymbol{W}_1 = \boldsymbol{I}_2 \otimes \boldsymbol{B}$. The first column of $\boldsymbol{B} \in \mathbb{C}^{N_1 N_2 \times L}$, denoted by $\boldsymbol{v}_{lm}$ (or $\tilde{\boldsymbol{v}}_{lm}$ for $\nu \in \{3, 4\}$ and $M_T \geq 16$, can be expressed as the Kronecker product of two column vectors, given as

$$\begin{cases} \tilde{\boldsymbol{v}}_{l,m} = \tilde{\boldsymbol{v}}_l \otimes \boldsymbol{u}_m, & \text{if } \nu = 3 \text{ or } 4 \text{ and } P_{\text{CSI-RS}} \geq 16 \\ \boldsymbol{v}_{l,m} = \boldsymbol{v}_l \otimes \boldsymbol{u}_m, & \text{otherwise,} \end{cases} \tag{32}$$

where

$$\begin{aligned} \boldsymbol{v}_l &= \begin{bmatrix} 1 & e^{j\frac{2\pi l}{O_1 N_1}} & \dots & e^{j\frac{2\pi l(N_1-1)}{O_1 N_1}} \end{bmatrix}^T, \\ \boldsymbol{u}_m &= \begin{bmatrix} 1 & e^{j\frac{2\pi m}{O_2 N_2}} & \dots & e^{j\frac{2\pi m(N_2-1)}{O_2 N_2}} \end{bmatrix}^T, \\ \tilde{\boldsymbol{v}}_l &= \begin{bmatrix} 1 & e^{j\frac{4\pi l}{O_1 N_1}} & \dots & e^{j\frac{4\pi l(N_1/2-1)}{O_1 N_1}} \end{bmatrix}^T, \end{aligned} \tag{33}$$

and the values of l and m have been determined in [32]. The parameters O_1 and O_2 are the over-sampling factors for the horizontal and the vertical directions, respectively. For higher ranks, $\nu > 1$, the rest of the columns could be defined similarly as $\boldsymbol{v}_{l',m'} = \boldsymbol{v}_{l'} \otimes \boldsymbol{u}_{m'}$, $\boldsymbol{v}_{l'',m''} = \boldsymbol{v}_{l''} \otimes \boldsymbol{u}_{m''}$ and $\boldsymbol{v}_{l''',m'''} = \boldsymbol{v}_{l'''} \otimes \boldsymbol{u}_{m'''}$, using coefficients l', l'', l''' and m', m'', m''', respectively.

The codebook precoder is generated using predefined indices known as PMI, which consist of either three or four indices depending on the supported rank. The precoder construction involves using the indices $i_{1,1}$ and $i_{1,2}$ to determine l and m, respectively. Moreover, the index $i_{1,3}$ is mapped to k_1 and k_2 to derive additional parameters l' and m' for ranks 2, 3, and 4. Similarly, for $\nu > 4$, discrete Fourier transform coefficients are derived by adding constant values to l and m. Assuming that constants α and β are added to l and m, respectively, we define $\boldsymbol{v}_\alpha = \begin{bmatrix} 1 & e^{j\frac{2\pi\alpha}{O_1 N_1}} & \dots & e^{j\frac{2\pi\alpha(N_1-1)}{O_1 N_1}} \end{bmatrix}^T$ and $\boldsymbol{u}_\beta = \begin{bmatrix} 1 & e^{j\frac{2\pi\beta}{O_2 N_2}} & \dots & e^{j\frac{2\pi\beta(N_2-1)}{O_2 N_2}} \end{bmatrix}^T$. Then, the second and subsequent columns of $\boldsymbol{B}$ can be derived based on the first column by calculating the element-wise Hadamard product between $\boldsymbol{v}_l$ (or $\boldsymbol{u}_m$) and $\boldsymbol{v}_\alpha$ (or $\boldsymbol{u}_\beta$). Thus, we have

$$\begin{aligned} \boldsymbol{v}_x &= \boldsymbol{v}_l \odot \boldsymbol{v}_\alpha = \boldsymbol{D}_\alpha \boldsymbol{v}_l, \\ \boldsymbol{u}_y &= \boldsymbol{u}_m \odot \boldsymbol{u}_\beta = \boldsymbol{D}_\beta \boldsymbol{u}_m. \end{aligned} \tag{34}$$

Here, $x \in \{l', l'', l'''\}$, $y \in \{m', m'', m'''\}$, $\boldsymbol{D}_\alpha = \text{Diag}(\boldsymbol{v}_\alpha)$ and $\boldsymbol{D}_\beta = \text{Diag}(\boldsymbol{u}_\beta)$, where $\text{Diag}(.)$ creates a diagonal matrix by placing the vector argument along its main diagonal. Finally, the corresponding column of $\boldsymbol{B}$ can be obtained by taking the Kronecker product of these two vectors as

$$\boldsymbol{v}_x \otimes \boldsymbol{u}_y = (\boldsymbol{D}_\alpha \boldsymbol{v}_l) \otimes (\boldsymbol{D}_\beta \boldsymbol{u}_m) = \boldsymbol{D}_{\alpha\beta} \boldsymbol{v}_{l,m}, \tag{35}$$

where, $\boldsymbol{D}_{\alpha\beta} = \boldsymbol{D}_{\alpha} \otimes \boldsymbol{D}_{\beta}$. This shows the dependencies between the first column and the remaining columns of $\boldsymbol{B}$.

The $\boldsymbol{W}_2$ matrix chooses a group of DFT vectors from $\boldsymbol{W}_1$ and applies phase shifts over panels and/or polarizations. It can be expressed as the following Khatri-Rao product

$$\boldsymbol{W}_2 = \boldsymbol{\Psi} \diamond \boldsymbol{E}, \tag{36}$$

where the permutation matrix $\boldsymbol{E} \in \{0,1\}^{L\times\nu}$ selects the intended beam from the DFT column vectors of $\boldsymbol{B}$. The phase shifting between two polarizations is performed by $\boldsymbol{\Psi} = \begin{bmatrix} \mathbf{1}_{\nu}^{T} \\ \tilde{\boldsymbol{\varphi}}_{\nu}^{T} \end{bmatrix} \in \mathbb{C}^{2\times\nu}$, where $\mathbf{1}_{\nu}$ is a column vector of length ν with all elements equal to one, and another length-ν vector $\tilde{\boldsymbol{\varphi}}_{\nu}$ specified in [32], containing elements from the set $\{\pm 1, \pm\varphi_n\}$, with $\varphi_n = e^{\frac{j\pi n}{2}}$.

B. Proposed PMI selection algorithms for BICM rate maximization

To maximize the BICM rate in a codebook-based precoder selection scheme, the exhaustive search algorithm is the most straightforward but time-consuming and impractical for large system dimensions. Due to the discrete parameters and the constant-modulus constraints of DFT-based codebook precoders, the problem (8) becomes combinatorial, making it impractical to solve optimally. There are two solution methods for combinatorial optimization problems: exact and heuristic methods. However, even using the MMSE capacity cost function, an exact solution is difficult to obtain due to the non-convex nature of the problem and multiple variables. Numerical and optimization techniques used to solve these problems often require iterative procedures and may only converge to a local optimum. As such, a heuristic approach is necessary to balance exploration and exploitation and find optimal or near-optimal solutions.

In the absence of a per-antenna transmit power constraint, it is well-known that the optimal precoder can always be chosen to coincide with the right singular vectors of the channel matrix $\boldsymbol{H}$ through its SVD or equivalently found by an eigenvalue-decomposition (EVD) of $\boldsymbol{H}^H\boldsymbol{H}$ [27]. To minimize the mean squared error (MSE) and maximize the capacity, the optimal rank-N_L precoder can be constructed from the first N_L columns of $\boldsymbol{L}$, which correspond to the N_L largest eigenvalues of $\boldsymbol{H}^H\boldsymbol{H}$ [17]. Given this reference precoder, to approximate the closest precoder from the codebook for a given rank of $\nu = N_L$, it is sufficient to estimate its components, which are $\boldsymbol{v}_{l,m}$ ($\tilde{\boldsymbol{v}}_{l,m}$ when $\nu \in \{3,4\}$ and $M_T \geq 16$), the cophasing term φ_n, the third component which is $\boldsymbol{v}_{xy}$ for $y \in \{m', m'', m'''\}$, $x \in \{l', l'', l'''\}$ when $\nu > 1$, and the additional component θ_p when $\nu \in \{3,4\}$ and $M_T \geq 16$. These components are associated with the PMI indices $(i_{1,1}i_{1,2})$, i_2, and $i_{1,3}$. Therefore, the proposed approach to derive these constituent elements of the codebook precoder is to utilize the dominant right singular vector of the channel matrix. It is evident that the columns of the codebook precoder are interrelated, and performing a column-wise mapping of it to the optimal SV-based precoder is inefficient since it may converge to inconsistent solutions. Additionally, jointly optimizing these three constructing elements is challenging, and there is no explicit solution to this problem. To overcome this issue, we propose a simple and effective heuristic approach that involves a combination of searching over predefined values of one or two variables and mapping the remaining variables to the optimal singular-vector (SV)-based precoder.

For $\nu = 2$ or $\nu \in \{3,4\}$ with $M_T < 16$, deriving PMI indices for the codebook precoder involves estimating the vector $\boldsymbol{v}_{l,m}$ and the variables k_1 and k_2, which determine the DFT phases of the second (and fourth) column(s) of the codebook matrix precoder. This is a non-trivial task as the PMI index $i_{1,3}$ is mapped to the variables k_1 and k_2, but the reverse mapping from k_1 and k_2 to $i_{1,3}$ is not always unique for all antenna configurations. Furthermore, according to (35), if the value of the variable $i_{1,3}$ is given, the sub-vector $\boldsymbol{v}_{l',m'}$ can be constructed in terms of the sub-vector $\boldsymbol{v}_{l,m}$. In the case of $\nu \in \{3,4\}$ with $M_T > 16$, the PMI index $i_{1,3}$ determines the variable θ_p, which is a phase-shifting term in constructing the codebook precoder columns.

As there are only limited specified values for the variable $i_{1,3}$, a practical way to find near-optimal values of the PMI indices is to fix $i_{1,3}$ to one of its permitted values and then estimate $i_{1,1}$, $i_{1,2}$, and i_2 by mapping the first column of the codebook precoder to the dominant eigenvector.

Let $\boldsymbol{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2, ..., \boldsymbol{l}_{M_T}]$ represent the right singular vectors of the channel matrix, and let $\boldsymbol{w}_i$ denote the ith column of the codebook precoder $\boldsymbol{W}$. We assume that the first column $\boldsymbol{l}_1$ can be regarded as the first column of the codebook precoder matrix, denoted by $\boldsymbol{w}_1$, plus an error term due to the mapping to a unit-modulus codebook [20], and we have:

$$\boldsymbol{l}_1 = \boldsymbol{w}_1 + \boldsymbol{\epsilon}. \tag{37}$$

where $\boldsymbol{\epsilon}$ stands for the mapping error. The first column can be written as

$$\boldsymbol{w}_1 = \begin{bmatrix} \boldsymbol{v}_{l,m} \\ \varphi_n \boldsymbol{v}_{l,m} \end{bmatrix} = \begin{bmatrix} 1 \\ \varphi_n \end{bmatrix} \otimes \boldsymbol{v}_{l,m}. \tag{38}$$

The elements of $\boldsymbol{w}_1$ are unit-modulus DFT-like components with a Kronecker structure composed of three different DFT vectors: $\boldsymbol{v}_l$, $\boldsymbol{u}_m$, and $\begin{bmatrix} 1 & \varphi_n \end{bmatrix}^T$. By examining the relationship between the elements of $\boldsymbol{v}_{l,m}$ in

$$\boldsymbol{v}_{l,m} = \boldsymbol{v}_l \otimes \boldsymbol{u}_m = \begin{bmatrix} \boldsymbol{u}_m^T & e^{j\frac{2\pi l}{O_1 N_1}}\boldsymbol{u}_m^T & \cdots & e^{j\frac{2\pi l(N_1-1)}{O_1 N_1}}\boldsymbol{u}_m^T \end{bmatrix}^T, \tag{39}$$

we can see that starting from the first element, every N_2 consecutive elements are associated with $\boldsymbol{u}_m$, and we have N_1 and $2N_1$ copies of it along $\boldsymbol{v}_{l,m}$ and $\boldsymbol{w}_1$, respectively. Similarly, the first N_2 consecutive elements of $\boldsymbol{v}_{l,m}$ are multiplied by the first element of $\boldsymbol{v}_l$, the second N_2 consecutive elements of $\boldsymbol{v}_{l,m}$ are multiplied by the second element of $\boldsymbol{v}_l$, and so on until the end, which implies that there are N_2 and $2N_2$ copies of each element of $\boldsymbol{v}_l$ in $\boldsymbol{v}_{l,m}$ and $\boldsymbol{w}_1$, respectively. Moreover, the second N_1N_2 elements of $\boldsymbol{w}_1$ are φ_n phase-shifted versions of the first N_1N_2 elements.

Let us have N observations of a complex-valued sinusoidal signal buried in noise

$$x(k) = ae^{j\phi k} + \epsilon(k), \qquad k = 0, 1, ..., N-1 \tag{40}$$

where a is a complex-valued amplitude, and $\phi \in (-\pi, \pi)$ is the normalized phase and ϵ is the noise term. Using N samples of the noisy signal $x(k)$, the covariance method of linear prediction is employed for the phase estimation and the first order estimator is

$$\hat{\phi} = \angle \left(\sum_{k=0}^{N-2} x^*(k) x(k+1) \right) \mod (2\pi), \tag{41}$$

which is called linear frequency estimation in [33]- [34] and linear phase estimation (LPE) in [20]. In the subsequent equations based on LPE, the modulo operation has been omitted for brevity.

By defining $\theta_l \triangleq \frac{2\pi l}{N_1 O_1}$ and $\theta_m \triangleq \frac{2\pi m}{N_2 O_2}$, we can represent two consistent DFT vectors of $\boldsymbol{v}_{l,m}$, defined in (33), as $\boldsymbol{v}_l = \begin{bmatrix} 1 & e^{j\theta_l} & \cdots & e^{j(N_1-1)\theta_l} \end{bmatrix}^T$ and $\boldsymbol{u}_m = \begin{bmatrix} 1 & e^{j\theta_m} & \cdots & e^{j(N_2-1)\theta_m} \end{bmatrix}^T$. Thus, our objective is to estimate the unknown phases θ_l, θ_m and φ_n from the vector $\boldsymbol{l}_1$ in (37). To detect $\boldsymbol{w}_1$ from the perturbed vector $\boldsymbol{v}_{l,m}$, we utilize (41) to estimate each of the three phases from their relevant samples as described above. Although we have phase-shifted samples for each desired phase, (41) is successful in phase estimation because the common phase rotation of samples is canceled and does not affect the LPE result.

To illustrate the LPE technique for estimating the variable θ_m, we consider that there are $2N_1$ copies of $\boldsymbol{u}_m$ in $\boldsymbol{l}_1$, and applying LPE to these relevant samples yields:

$$\hat{\theta}_m = \angle \left(\sum_{j=1}^{2N_1} \sum_{k=1}^{N_2-1} \boldsymbol{l}_1^* \left((j-1)N_2 + k \right) \boldsymbol{l}_1 \left((j-1)N_2 + k + 1 \right) \right). \tag{42}$$

To estimate the unknown phases θ_l and φ_n, we adopt the same approach as for θ_m by considering the appropriate subvectors that contain the corresponding DFT vector elements with different rotations. There are $2N_2$ copies of each element of $\boldsymbol{v}_l$ with different phase shifts, and we can combine them using LPE to estimate θ_l, which yields

$$\hat{\theta}_l = \angle \left(\sum_{\substack{j=1 \\ j \neq N_1}}^{2N_1-1} \sum_{k=1}^{N_2} \boldsymbol{l}_1^* \left((j-1)N_2 + k \right) \boldsymbol{l}_1 \left(jN_2 + k \right) \right). \tag{43}$$

Considering the structure of $\boldsymbol{w}_1$, we observe that the phase difference between the first $N_1 N_2$ elements and the second $N_1 N_2$ elements is equal to φ_n. Therefore, we can estimate the phase φ_n using LPE as follows:

$$\hat{\varphi}_n = \angle \left(\sum_{k=1}^{N_1 N_2} \boldsymbol{l}_1^* (k) \, \boldsymbol{l}_1 \left(N_1 N_2 + k \right) \right). \tag{44}$$

For each phase variable, we have multiple copies of the corresponding DFT vector, which are combined in the last three equations. Instead of utilizing all the available samples for phase detection, we can reduce the computational complexity by using a fewer number of observations. This results in a lower cost and the following equations:

$$\begin{aligned} \hat{\theta}_m &= \angle \left(\sum_{k=1}^{N_2-1} \boldsymbol{l}_1^* (k) \, \boldsymbol{l}_1 (k+1) \right), \\ \hat{\theta}_l &= \angle \left(\sum_{k=1}^{N_2} \boldsymbol{l}_1^* (k) \, \boldsymbol{l}_1 \left(N_2 + k \right) \right), \\ \hat{\varphi}_n &= \angle \left(\sum_{k=1}^{N_1} \boldsymbol{l}_1^* (k) \, \boldsymbol{l}_1 \left(N_1 N_2 + k \right) \right). \end{aligned} \tag{45}$$

Having estimated the phases, the DFT coefficients can be obtained through hard decision as follows:

$$\begin{aligned} \hat{m} &= \text{round} \left(\frac{O_2 N_2}{2\pi} \hat{\theta}_m \right), \; \hat{l} = \text{round} \left(\frac{O_1 N_1}{2\pi} \hat{\theta}_l \right), \\ \hat{n} &= \text{round} \left(\frac{2}{\pi} \hat{\varphi}_n \right). \end{aligned} \tag{46}$$

To map the estimated DFT coefficients $\hat{l}$, $\hat{m}$ and $\hat{n}$ to the corresponding PMI indices $i_{1,1}$, $i_{1,2}$ and i_2, we need to follow a mapping rule that depends on the number of transmit layers and the codebook mode. The mapping rules for different ranks of $N_L \in \{1, 2, 3, 4\}$ and antenna configurations are presented in Table II.

After applying the aforementioned procedure for each permitted value of $i_{1,3}$, along with the estimated parameters $\hat{i}_{1,1}$, $\hat{i}_{1,2}$, and $\hat{i}_2$, the respective codebook precoder $\hat{\boldsymbol{P}}_{i_{1,3}}$ can be found. The selection of the best $i_{1,3}$ to achieve the optimal BICM rate involves assessing the BICM rate for each permissible $i_{1,3}$. However, as the BICM rate cannot be directly evaluated with a closed-form formulation, alternative cost functions are proposed to reduce computational complexity. Given that BICM mutual information characterizes achievable rates based on specific coding and modulation, we consider MIMO capacity—representing the upper limit of mutual information—as a significant precoding design criterion. This criterion has been extensively utilized in the context of spatial multiplexing and MIMO systems for precoder selection [17], [27], [35]- [36]. When the transmitter precodes the information with $\boldsymbol{P}$ before transmission, the mutual information assuming an uncorrelated complex Gaussian source, given $\boldsymbol{H}$ and a fixed $\boldsymbol{P}$, is

$$I(\boldsymbol{P}) = \log_2 \det \left(\boldsymbol{I}_{N_L} + \frac{1}{\sigma^2} \boldsymbol{P}^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{P} \right). \tag{47}$$

One criterion inspired by capacity is to maximize the mutual information by finding $i_{1,3} \in \{1, ..., I_{1,3}\}$ that maximizes the determinant of the effective channel matrix, given by

$$\hat{i}_{1,3} = \arg \max_{0 \leq i_{1,3} \leq I_{1,3}} \det \left(\hat{\boldsymbol{P}}_{i_{1,3}}^H \boldsymbol{H}^H \boldsymbol{H} \hat{\boldsymbol{P}}_{i_{1,3}} \right). \tag{48}$$

Alternatively, for certain system configurations, another appropriate selection metric could be the maximization of transmit power. In this case, the index $i_{1,3}$ that maximizes the transmit power can be chosen based on:

$$\hat{i}_{1,3} = \arg \max_{0 \leq i_{1,3} \leq I_{1,3}} \left\| \hat{\boldsymbol{P}}_{i_{1,3}}^H \boldsymbol{H}^H \boldsymbol{H} \hat{\boldsymbol{P}}_{i_{1,3}} \right\|_F. \tag{49}$$

Based on the steps outlined above, the proposed method is able to determine the best PMI indices without the need

Table II: Mapping from $\hat{l}$, $\hat{m}$ and $\hat{n}$ to $i_{1,1}$, $i_{1,2}$ and i_2 for different ranks

Config.	$i_{1,1}$	$i_{1,2}$	i_2
$N_L = 1$ Mode = 1	$\hat{l} \bmod (N_1 O_1)$	$\hat{m} \bmod (N_2 O_2)$	$\hat{n} \bmod 4$
$N_L = 1$ Mode = 2, $N_2 > 1$	$\lfloor \frac{\hat{l}}{2} \rfloor \bmod \left(\frac{N_1 O_1}{2} \right)$	$\lfloor \frac{\hat{m}}{2} \rfloor \bmod \left(\frac{N_2 O_2}{2} \right)$	$4\left(\hat{l} \bmod 2\right) + 8\left(\hat{m} \bmod 2\right) + \hat{n}$
$N_L = 1$ or 2 Mode = 2, $N_2 = 1$	$\lfloor \frac{\hat{l}}{2} \rfloor \left(\hat{l} = 0 \text{ or } 1\right)$, $\lfloor \frac{\hat{l}}{2} \rfloor - 1 \left(\hat{l} = N_1 O_1 \text{ or } N_1 O_1 + 1\right)$, $\lfloor \frac{\hat{l}}{2} \rfloor$ or $\lfloor \frac{\hat{l}}{2} \rfloor - 1$ (Oth.)	$\hat{m} = 0$	$4\left(\left(\hat{l} - 2 i_{1,1}\right) \bmod 4\right) + \hat{n}$
$N_L = 2$ Mode = 1	$\hat{l} \bmod (N_1 O_1)$	$\hat{m} \bmod (N_2 O_2)$	$\hat{n} \bmod 2$
$N_L = 2$ Mode = 2, $N_2 > 1$	$\lfloor \frac{\hat{l}}{2} \rfloor \bmod \left(\frac{N_1 O_1}{2} \right)$	$\lfloor \frac{\hat{m}}{2} \rfloor \bmod \left(\frac{N_2 O_2}{2} \right)$	$\hat{n} \bmod 2$
$N_L = 3$ or 4 $M_T < 16$	$\hat{l} \bmod (N_1 O_1)$	$\hat{m} \bmod (N_2 O_2)$	$\hat{n} \bmod 2$
$N_L = 3$ or 4 $M_T \geq 16$	$\hat{l} \bmod \left(\frac{N_1 O_1}{2} \right)$	$\hat{m} \bmod (N_2 O_2)$	$\hat{n} \bmod 2$

to search over all candidate precoders in the codebook and calculate the metric for each one. This eliminates the search phase completely or to a large extent, by directly deriving the PMI indices or reducing the number of search parameters to one index. To summarize, Algorithm 2 describes the steps of the proposed procedure for finding the best PMI indices.

Algorithm 2 SV-aided One-dimensional Search

1. **Input**: The right singular vectors of $\boldsymbol{H}$ denoted by $\boldsymbol{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2, ..., \boldsymbol{l}_{M_T}]$
2. Map $\boldsymbol{l}_1$ to $\begin{bmatrix} \boldsymbol{v}_{l,m} \\ \varphi_n \boldsymbol{v}_{l,m} \end{bmatrix}$ and estimate $\hat{\theta}_l$, $\hat{\theta}_m$ and $\hat{\varphi}_n$ using (45)
3. Obtain $\hat{m}$, $\hat{l}$ and $\hat{n}$ using (46)
4. Using Table II to map $\hat{l}$, $\hat{m}$ and $\hat{n}$ to $\hat{i}_{1,1}$, $\hat{i}_{1,2}$ and $\hat{i}_2$, respectively
5. **for** $i_{1,3} = 0$ **to** $I_{1,3}$ **do**
   - Find the precoder $\boldsymbol{P}_{i_{1,3}}$ corresponding to $\hat{i}_{1,1}$, $\hat{i}_{1,2}$ and $i_{1,3}$ in the codebook
   - Compute the cost function $C_{i_{1,3}}$, using the BICM rate or the criteria in (48) or (49)

   end for
6. $\hat{i}_{1,3} = \arg \max_{0 \leq i_{1,3} \leq I_{1,3}} C_{i_{1,3}}$
7. **Output**: $\boldsymbol{P}_{\text{opt}} = \boldsymbol{P}_{(\hat{i}_{1,1}, \hat{i}_{1,2}, \hat{i}_2, \hat{i}_{1,3})}$

For $\nu \in \{3, 4\}$ and $M_T \geq 16$, the rank-ν codebook precoder has a structure where all columns are constructed using the same sub-vector $\tilde{\boldsymbol{v}}_{l,m}$. This suggests combining the N_c columns of the rank N_L SV-based precoder for a given $i_{1,3}$ and estimating the other three PMI indices using a similar approach to Algorithm 2. We can compute three unknown phases θ_l, θ_m, and φ_n using the equations below, instead of (42)- (44):

$$\hat{\theta}_m = \angle \left(\sum_{i=1}^{N_c} \sum_{k=1}^{N_2 - 1} \boldsymbol{l}_i^* (k) \, \boldsymbol{l}_i (k+1) \right), \tag{50}$$

$$\hat{\theta}_l = \angle \left(\sum_{i=1}^{N_c} \sum_{k=1}^{N_2} \boldsymbol{l}_i^* (k) \, \boldsymbol{l}_i (N_2 + k) \right),$$

$$\hat{\varphi}_n = \angle \left(\sum_{i=1}^{N_c} \sum_{k=1}^{N_1} \boldsymbol{l}_i^* (k) \, \boldsymbol{l}_i (N_1 N_2 + k) \right).$$

All dominant N_c right singular vectors contribute in the phase estimation process, which leads to a larger number of samples used for phase calculation with the LPE method. The remaining steps, including mapping estimated DFT coefficients and selecting $i_{1,3}$ by computing a selection criterion, are identical to Algorithm 2. The steps are detailed in Algorithm 3.

Algorithm 3 Combined-SV aided One-dimensional Search

1. **Input**: The right singular vectors of $\boldsymbol{H}$ denoted by $\boldsymbol{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2, ..., \boldsymbol{l}_{M_T}]$
2. Exploiting $[\boldsymbol{l}_1, \boldsymbol{l}_2, ..., \boldsymbol{l}_{N_c}]$ to estimate $\hat{\theta}_l$, $\hat{\theta}_m$ and $\hat{\varphi}_n$ using (50)
3. Obtain $\hat{m}$, $\hat{l}$ and $\hat{n}$ using (46)
4. Using Table II to map $\hat{l}$, $\hat{m}$ and $\hat{n}$ to $\hat{i}_{1,1}$, $\hat{i}_{1,2}$ and $\hat{i}_2$, respectively
5. **for** $i_{1,3} = 0$ **to** $I_{1,3}$ **do**
   - Find the precoder $\boldsymbol{P}_{i_{1,3}}$ corresponding to $\hat{i}_{1,1}$, $\hat{i}_{1,2}$, $\hat{i}_2$ and $i_{1,3}$ in the codebook
   - Compute the cost function $C_{i_{1,3}}$, using the BICM rate or the criteria in (48) or (49)

   end for
6. $\hat{i}_{1,3} = \arg \max_{0 \leq i_{1,3} \leq I_{1,3}} C_{i_{1,3}}$
7. **Output**: $\boldsymbol{P}_{\text{opt}} = \boldsymbol{P}_{(\hat{i}_{1,1}, \hat{i}_{1,2}, \hat{i}_2, \hat{i}_{1,3})}$

Algorithm 2 estimates the first column of the codebook precoder using the dominant eigenvector, but this may not always lead to optimal choices for the corresponding indices in NLoS channel models. To address this, we propose two modified algorithms. The first considers N_c dominant eigenvectors and estimates the parameters for each column through Algorithm 2. The optimal eigenvector and $i_{1,3}$ are then determined by computing the cost function for the codebook precoder corresponding to the estimated $\hat{i}_{1,1}$, $\hat{i}_{1,2}$, and $\hat{i}_2$ for each specified eigenvector and different values of $i_{1,3}$.

In the second modification, i_2 is considered as an additional search parameter to further enhance the performance. Given the limited number of allowed values for i_2, determining the optimal value involves computing the cost function across combinations of its different values with the other variables, including $i_{1,3}$ and the first N_c eigenvectors. The detailed

steps of these two algorithms are outlined in a comprehensive manner in Algorithm 4. The nature of the search step is contingent on whether both i_2 and $i_{1,3}$ are explored or if the search is solely focused on $i_{1,3}$. This distinction leads to either a two-dimensional or three-dimensional search, respectively. In the two-dimensional case, the inner loop involving the variable i_2 within the nested loops with variables $i_{1,3}$ and i_2 is omitted.

Algorithm 4 SV-aided Two/Three-dimensional Search

1. **Input**: The right singular vectors of $\boldsymbol{H}$ denoted by $\boldsymbol{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2, ..., \boldsymbol{l}_{M_T}]$
2. **for** $n_c = 1$ **to** N_c **do**
   - Map the first $N_1 N_2$ elements of $\boldsymbol{l}_{n_c}$ to $\boldsymbol{v}_{l,m}$ and estimate $\hat{\theta}_l$, and $\hat{\theta}_m$ as in (45)
   - Obtain $\hat{m}$ and $\hat{l}$ using (46)
   - Using Table II to map $\hat{l}$ and $\hat{m}$ to $\hat{i}_{1,1}$ and $\hat{i}_{1,2}$, respectively
   - **for** $i_{1,3} = 0$ **to** $I_{1,3}$ **do**
     - **for** $i_2 = 0$ **to** I_2 **do**
       - Find the codebook precoder $\boldsymbol{P}$ corresponding to $\{\hat{i}_{1,1}, \hat{i}_{1,2}, i_2, i_{1,3}\}$
       - Compute the cost function of $C_{n_c, i_2, i_{1,3}}$
     - **end for**
   - **end for**

   end for
3. $\hat{n}_c, \hat{i}_2, \hat{i}_{1,3} = \arg \max_{\substack{0 \leqslant i_2 \leqslant I_2, \\ 0 \leqslant i_{1,3} \leqslant I_{1,3} \\ 0 \leqslant n_c \leqslant N_c}} C_{(n_c, i_2, i_{1,3})}$
4. **Output**: $\boldsymbol{P}_{\text{opt}} = \boldsymbol{P}_{(\hat{i}_{1,1}, \hat{i}_{1,2}, \hat{i}_2, \hat{i}_{1,3})}$

C. Computational Complexity Analysis

In this section, we analyze the computational complexity of the proposed low-dimensional search algorithms and compare them to the widely used exhaustive search in the codebook, based on the MMSE capacity as the baseline, since there is no closed-form solution for the BICM rate evaluation. In the four proposed SV-based algorithms, we first need to compute the SVD of the channel matrix or the EVD of the Gram matrix, which is proportional to $\mathcal{O}(M_T^3)$. Calculating the DFT coefficients l, m, and n according to (42)-(44) requires $2N_1(N_2-1)$, $2(N_1-1)N_2$, and N_1N_2 complex multiplications and $2N_1(N_2-1)-1$, $2(N_1-1)N_2-1$, and N_1N_2-1 complex additions, respectively. By exploiting the alternative formulations in (45), the number of operations is reduced to N_2-1, N_2, and N_1 complex multiplications and N_2-2, N_2-1, and N_1-1 complex additions, respectively. To find the best value of the fixed parameters or the best column among the first N_c eigenvectors, we need to calculate the cost function of $\det\left(\boldsymbol{P}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{P}\right)$. Since matrix multiplication is associative, we can first compute $\boldsymbol{H}^H\boldsymbol{H}$ once, which requires $M_T M_R^2$ multiplications. For each candidate precoder, we then need to multiply the left and right sides by $\boldsymbol{P}^H$ and $\boldsymbol{P}$, respectively, each of which needs $N_L M_T^2$ multiplications. If we perform the matrix products from the left to the right side, we will have a total of $2M_TM_RN_L + M_TN_L^2$ multiplications. In this way, we have to perform three matrix products for all precoder candidates. The complexity of computing the determinant of the resulting matrix is proportional to $\mathcal{O}\left(N_L^3\right)$. The computational complexity of the main steps of Algorithms 2-4 and the number of times each step is performed for different algorithms are represented in Tables III and IV, respectively. One significant advantage of our proposed algorithms is the significant reduction in the number of search parameters, leading to much lower complexity.

Table III: Computation complexity of the main steps of SV-based solution

Step	(Mult., Add.) or Com. Order
EVD of $\boldsymbol{H}^H\boldsymbol{H}$	$\mathcal{O}\left(M_T^3\right)$
Deriving $\hat{l}$ and $\hat{m}$	$(2N_2-1, 2N_2-3)$
Deriving $\hat{n}$	$(2N_1-1, 2N_1-3)$
$\boldsymbol{P}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{P}$	$\begin{pmatrix} 2M_TM_RN_L + \\ M_TN_L^2 \end{pmatrix}, \begin{matrix} 2M_TM_RN_L - \\ N_L(M_R+M_T+N_L) \end{matrix}$
$\det\left(\boldsymbol{P}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{P}\right)$	$\mathcal{O}\left(N_L^3\right)$

Table IV: Repetition times of the SV-based solution's main steps for different algorithms

Step	Alg. 1	Alg. 2	Alg. 3	Alg. 4
EVD of $\boldsymbol{H}^H\boldsymbol{H}$	1	1	1	1
Estimating $\hat{l}$ and $\hat{m}$	1	N_c	N_c	N_c
Estimating $\hat{n}$	1	N_c	N_c	-
$\boldsymbol{P}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{P}$	$I_{1,3}$	$I_{1,3}$	$N_cI_{1,3}$	$N_cI_2I_{1,3}$
$\det\left(\boldsymbol{P}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{P}\right)$	$I_{1,3}$	$I_{1,3}$	$N_cI_{1,3}$	$N_cI_2I_{1,3}$

To compare the computational complexity of different methods, we use the exhaustive search approach based on the MMSE capacity over different precoders in the codebook as the baseline method. This involves computing the matrix-by-matrix products for the term $\boldsymbol{P}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{P} + \sigma^2\boldsymbol{I}_{N_L}$ using $2M_TM_RN_L + M_TN_L^2$ multiplications and a matrix inversion proportional to $\mathcal{O}\left(N_L^3\right)$. These operations must be executed for all the existing precoder candidates in the codebook, which for each particular number of layers depends on the number of transmit antennas and their configuration, to find the best precoder that maximizes the MMSE capacity.

One significant advantage of our proposed algorithms is the significant reduction in the number of search parameters, leading to much lower complexity. For example, in a 12×4 MIMO system, the complexity is reduced by up to 100 times compared to the baseline exhaustive search solution.

V. Simulation Results

A. Non-Codebook Based Precoding

In this section, we provide numerical examples to compare the performances of the proposed linear precoding algorithms in terms of the achievable BICM rate and BLER through two channel examples. The SNR is defined as $\text{SNR} = \frac{\gamma}{\sigma^2}$.

Example 1

We investigate the performance of our proposed precoding algorithms for a 2×2 static MIMO system with channel

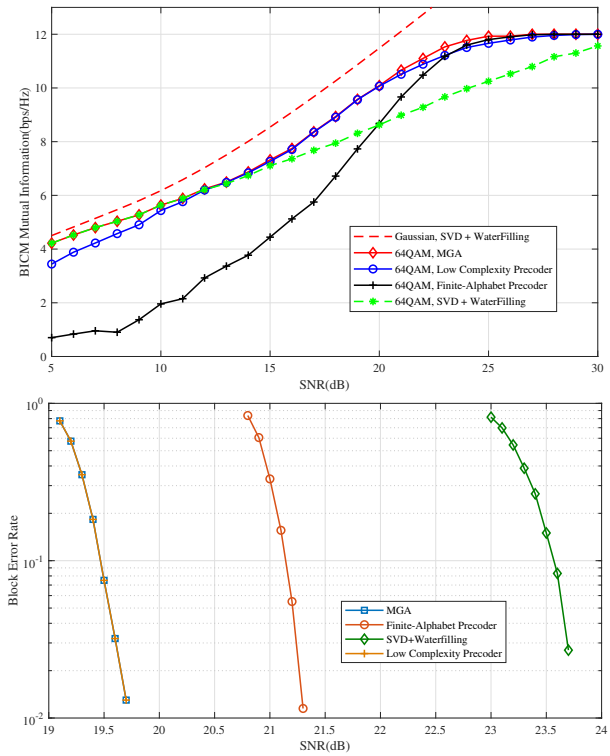

Figure 1: Achievable rate and BLER performance of a 2×2 MIMO-BICM system with 64QAM

matrix $\boldsymbol{H} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, where the input signal is drawn from 64QAM and $\gamma = 2$. To evaluate the performance, we compare our proposed algorithms with the precoding algorithm with finite-alphabet inputs in [9] and the capacity-achieving SVD precoder with water-filling power allocation. Fig. 1 shows that, the proposed algorithms outperform the existing finite-alphabet precoding algorithm across the entire range of SNRs. Moreover, at high SNR values, our algorithms surpass the SVD precoder with water-filling power allocation. The dashed line curve represents the upper bound of the achievable BICM rate, which corresponds to the MIMO capacity with Gaussian inputs. The low complexity precoding has performance loss of about 2-dB in low SNR compared with the optimal MGA precoding, but this loss can be fully compensated by rank adaptation algorithms which select the number of transmit data streams N_L according to the channel condition. In the medium to high SNR regimes, our proposed low-complexity precoding approaches the optimal performance of MGA precoding with significantly reduced complexity.

From the Fig. 1, it is evident that our MGA algorithm achieves 9 bps/Hz at SNR = 18 dB. In comparison, the finite-alphabet precoder and SVD with water-filling precoder achieve the same BICM mutual information at SNR = 20.3 dB and SNR = 21 dB, respectively. We have evaluated the BLER performance for MGA precoder at 18 dB, finite-alphabet precoder at 20.3 dB, and SVD with water-filling precoder at 21 dB using an LDPC encoder with coderate = 0.75 and code length $L = 12000$ bits. The receiver utilizes the normalized min-sum decoding algorithm for LDPC codes with 20 iterations. Our proposed MGA and low-complexity precoding methods demonstrate comparable performance and surpass both the SVD precoder with water-filling power allocation and the finite-alphabet precoding method, which aims to maximize the mutual information of coded modulation. Focusing on the SNR corresponding to BLER = 0.1, the figure illustrates that the MGA precoder exhibits approximately 1.7 dB and 4.2 dB gain compared to the finite-alphabet precoder and SVD with waterfilling precoder, respectively. Hence, maximizing BICM mutual information proves to be an excellent criterion for minimizing the block error rate.

Example 2

We now investigate a 4×4 complex-valued channel matrix, represented as follows:

$$\boldsymbol{H} = \begin{bmatrix} 1.3+0.8i & 1.3-1.4i & -1.2-1.5i & 0.8+0.3i \\ 0.1+1.3i & 1.4+1.9i & -0.3-0.8i & 0.9+0.7i \\ -0.9-0.2i & -0.8+0.7i & -0.8+0.8i & 0.1+0.9i \\ -0.2+0.2i & 0.4-0.6i & -1.7-0.8i & 0.0+0.4i \end{bmatrix}$$

where the input signal is drawn from 16QAM and we assume $\gamma = 4$. In Fig. 2, we compare our proposed algorithms with the capacity-achieving SVD with waterfilling power allocation and finite-alphabet precoding. The figure demonstrates that our proposed MGA and low-complexity precoder algorithms achieve 10 bps/Hz at SNR = 9dB and SNR = 10 dB, respectively. In comparison, the finite-alphabet precoder and the SVD with waterfilling precoder attain the same BICM mutual information at SNR = 12.4 dB and SNR = 11 dB, respectively. The corresponding precoders are denoted by $\boldsymbol{P}_{\mathrm{MGA,9dB}}$, $\boldsymbol{P}_{\mathrm{LCP,10dB}}$, $\boldsymbol{P}_{\mathrm{FA,12.4dB}}$ and $\boldsymbol{P}_{\mathrm{SVDWF,11dB}}$. at this point, the dimensions of $\boldsymbol{P}_{\mathrm{MGA,9dB}}$ and $\boldsymbol{P}_{\mathrm{SVDWF,11dB}}$ are 4×3, indicating transmitting 3 data streams and a corresponding code rate of 0.83. Also, the dimensions of $\boldsymbol{P}_{\mathrm{LCP,10dB}}$ and $\boldsymbol{P}_{\mathrm{FA,12.4dB}}$ are 4×4, with a corresponding code rate of 0.625.

We proceed to evaluate the block error rate performance with a spectral efficiency of 10 bps/Hz under four system settings:

- $\boldsymbol{P}_{\mathrm{MGA,9dB}}$, 16QAM, 3 data streams and 0.83 code rate;
- $\boldsymbol{P}_{\mathrm{LCP,10dB}}$, 16QAM, 4 data streams, and 0.625 code rate;
- $\boldsymbol{P}_{\mathrm{FA,12.4dB}}$, 16QAM, 4 data streams and 0.625 code rate;
- $\boldsymbol{P}_{\mathrm{SVDWF,11dB}}$, 16QAM, 3 data streams, and 0.83 code rate.

We focus on the SNR corresponding to BLER = 0.1. The figure illustrates that the MGA algorithm has approximately 1.4 dB, 2.7 dB, and 3.5 dB gains compared to the low-complexity precoder, the SVD with waterfilling precoder, and the finite-alphabet precoder, respectively. This validates the efficacy of maximizing the BICM mutual information for minimizing the block error rate. The running time of the proposed MGA and the low-complexity precoding algorithms is subject to various factors, including modulation order, the number of data streams, iterations, and the number of samples used for computing BICM mutual information and its gradient. For instance, in a scenario with 16QAM modulation, 4 data streams, 1000 samples, and 8 iterations for MGA, the average running time is approximately 9.73 seconds. On the other

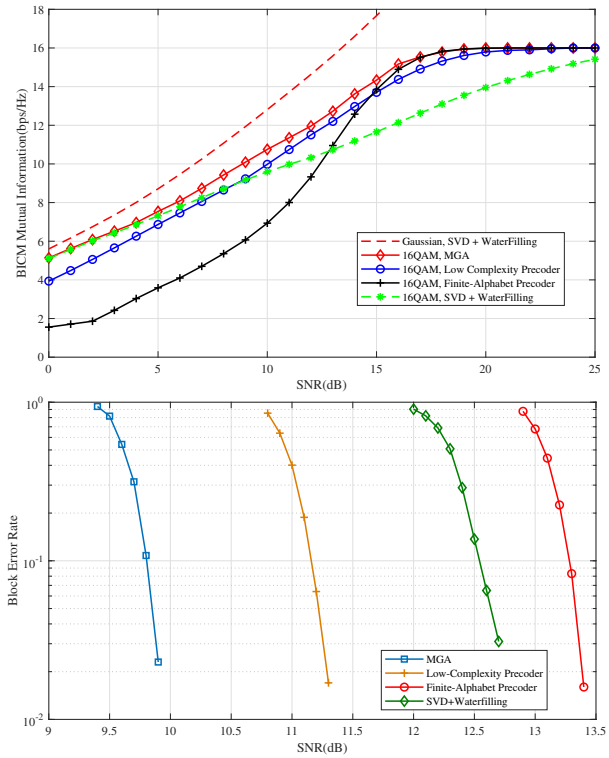

Figure 2: Achievable rate and BLER performance of a 4×4 MIMO-BICM system with 16QAM

hand, the low-complexity precoding algorithm, leveraging the simplicity of computing the SVD of the channel, exhibits a remarkably shorter running time of only 0.00083 seconds. This substantial difference underscores the efficiency and low computational cost of the low-complexity precoding algorithm, making it particularly attractive for scenarios with large MIMO systems.

B. Codebook Based Precoding

In this subsection, we present simulation results that assess the performance of our proposed PMI selection algorithms for BICM-MIMO systems, which use the type-I single-panel codebook. We compare our results with the exhaustive search methods used for selecting the optimal PMI that maximizes one of the two criteria of the achievable BICM rate or the MMSE capacity. We assume that the input signal vector elements are drawn from an M-QAM constellation, and we define the SNR as SNR = $\frac{\gamma}{\sigma^2}$, where $\gamma = 1$. The simulation scenarios consider a UE with $M_R = 4$ receive antennas, and we also use the channel capacity-related cost function, BICM rate, and transmit power as selection criteria in the search stage of the proposed algorithms.

We consider two scenarios to characterize the channel model. The first channel model is a highly correlated Kronecker Rayleigh fading channel model with two-dimensional cross-polarized antennas at the gNB and the UE, following the specifications in [23]. The gNB deploys cross-polarized antenna elements with $+/-45$ degrees polarization slant angles, while the UE deploys cross-polarized antenna elements with $+90/0$ degrees polarization slant angles.

The second channel model is the frequency-selective (multi-path) channel, where we use the CDL model for NLOS propagation described in 3GPP TR 38.901 [24]. To mitigate inter-symbol interference (ISI) and solve imperfections, we employ the orthogonal frequency division multiplexing (OFDM) technique with N sub-carriers. For this scenario, we use the QuaDRiGa5 simulator to generate a geometry-based stochastic channel model with simulation parameters listed in Table V.

Table V: Simulation OFDM framework parameters

Parameter	Value
Carrier frequency	3.5 GHz
OFDM subcarrier bandwidth	30 kHz
Delay spread (DS)	$0.35\mu s$
BS height	25 m
UE height	1.5 m
QuaDRiGa propagation scenario	3GPP 3D UMa NLOS
Number of multipaths	20
Antenna model	3GPP-3D
Antenna spacing	Half wavelength

In a MIMO-OFDM system, the sub-carriers are divided into groups of size K, and a precoder is designed for each group of adjacent sub-carriers. To this end, the Gram matrix of each group is calculated as $\tilde{\boldsymbol{H}} \triangleq \sum_{k=1}^{K} \boldsymbol{H}_k^H \boldsymbol{H}_k$, which serves as the input matrix for the proposed algorithms. The eigenvectors obtained from the EVD of the Gram matrix are used as inputs for the SV-based algorithms. We consider a transmission over 512 sub-carriers, and the same precoder is applied for all $K = 16$ sub-carriers of each resource block.

In the figure presented as Fig. 3, we demonstrate the achievable rate of Algorithms 2 and 3 in an 8×4 BICM-MIMO system with $(N_1, N_2) = (2, 2)$, for highly correlated channels. We observe that the performance gap for different numbers of layers compared to the BICM rate exhaustive search baseline method is less than 5%. Additionally, we find that the proposed SV-based methods outperform the MMSE capacity exhaustive search baseline for $N_L = 3, 4$ in the high SNR regime. In Fig. 4, we investigate a BICM-MIMO system an antenna configuration of $(N_1, N_2) = (4, 3)$. For each SNR value, we perform a comprehensive parameter tuning analysis to identify the optimal values for three key parameters: RI (determines the number of layers from $N_L \in \{1, 2, 3, 4\}$), modulation order (selected from QPSK, 16-QAM, and 64-QAM), and the PMI index (in type-I single panel codebook). The simulation results show that the proposed PMI selection algorithms can achieve a comparable performance to that of the exhaustive search method based on the BICM rate. Moreover, in the high-SNR range, the proposed algorithms outperform the MMSE capacity-based exhaustive search method. From the perspective of computational complexity, as detailed in Section IV-C, the exhaustive search for determining the best PMI, when $N_L = 2$, entails evaluating 512 and 1536 precoder candidates for 8 and 24 transmit antennas, respectively. Similarly, for $N_L = 3$ (or 4), the exhaustive search involves 384 and 768

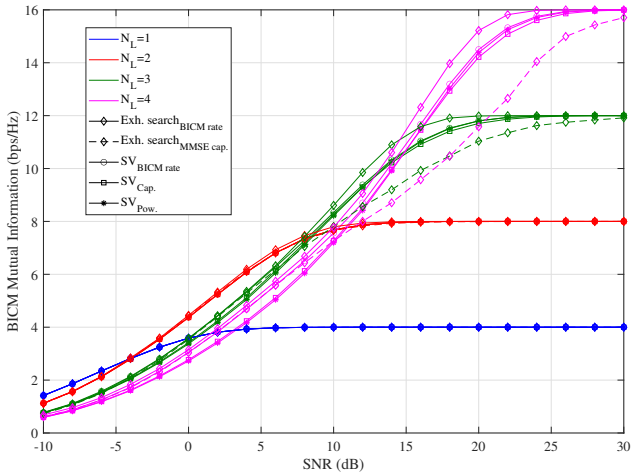

Figure 3: Achievable rate of the proposed SV-based algorithms for an 8×4 MIMO-BICM system with 16-QAM and highly correlated channels.

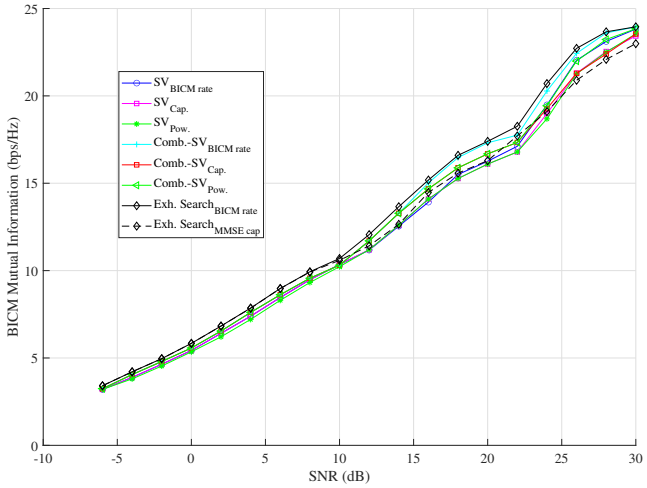

Figure 4: Achievable rate comparison of different algorithms for a 24×4 highly correlated MIMO-BICM system, with varying modulation orders and number of layers.

precoder candidates for 8 and 24 transmit antennas, respectively. In contrast, our algorithm exhibits a search complexity equivalent to $I_{1,3}$, which is set at4. Another advantage is that the proposed algorithms are applicable for each given RI and can be utilized for RI selection and link adaptation purposes. Additionally, by employing cost functions such as capacity, power instead of the BICM rate, the proposed algorithms do not depend on the SNR value, further reducing the computation complexity. In Fig. 5, we compare the performance of SV-based algorithms with the exhaustive search methods for multi-path fading channels in an 8×4 MIMO-OFDM system. To account for the high attenuation and weak signal of the NLoS channels, we examine more than one dominant eigenvector and select i_2 by evaluating its permitted values for the given selection cost function. Specifically, we apply Algorithm 4 with $N_c = 1$ for transmitting 3 layers in Fig. 5.

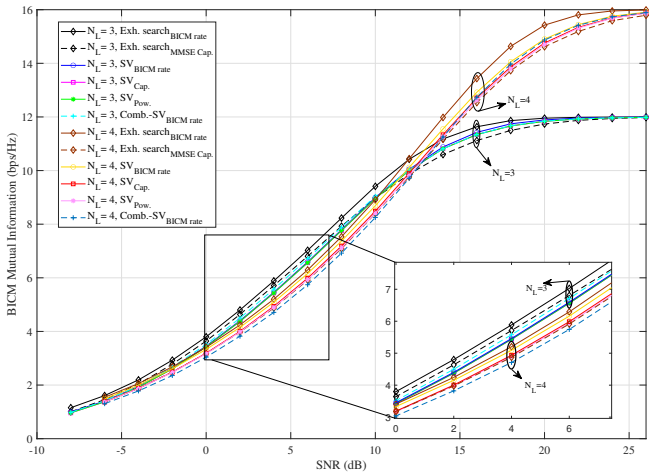

Figure 5: Achievable rate of the proposed SV-based algorithms for an 8×4 MIMO-BICM system with 16-QAM and $N_L = 3$ and $N_L = 4$ in multi-path NLoS channels.

As shown, Algorithm 4 achieves a performance close to that of the baseline methods by considering $N_c = 3$ (i.e., 3 dominant eigenvectors) for $N_L = 4$.

VI. Conclusion

This paper proposes several precoding algorithms for MIMO-BICM systems. Firstly, we propose two linear precoding algorithms for MIMO-BICM systems with a power constraint. The first algorithm, MGA precoding, is designed to maximize the achievable rate of the system. However, due to its high complexity, a low complexity non-iterative precoding algorithm is developed by studying the structure of the optimal precoder. Simulation results show that both proposed algorithms outperform the existing methods in terms of achievable rate, while the non-iterative algorithm significantly reduces the computational complexity.

Then, we focus on the MIMO-BICM system model based on 5G NR networks and introduce a precoding matrix selection technique to select the optimal PMI from the Type-I codebook in order to achieve the maximum BICM rate. Our proposed solution is based on the SVD of the channel matrix, and we put forward four different algorithms. In three of the designed algorithms, namely Algorithms 2 and 4, we derive the indices $i_{1,1}$ and $i_{1,2}$ from one of the dominant eigenvectors with low complexity. The difference among these algorithms is the number of dominant eigenvectors considered and whether the index i_2 is derived from the eigenvector or found by searching among its permitted values. On the other hand, our proposed SV-based Algorithm 3 utilizes a combination of N_c dominant eigenvectors to derive the indices $i_{1,1}$, $i_{1,2}$, and i_2. We have evaluated our proposed methods for single-carrier correlated channels, and as discussed in Section VI, they can be extended to facilitate OFDM multi-carrier transmission over non-line-of-sight (NLoS) channels with multi-path fading, which represents a worst-case scenario. Simulation results show that our proposed SV-based techniques can outperform the exhaustive search method based on the MMSE capacity and can reach

the performance of the BICM rate exhaustive search method in Rayleigh correlated channels. The complexity of our proposed methods is also lower than the exhaustive search baseline, which makes them more practical in real-world systems. Specifically, for an 8×4 and 12×4 MIMO systems, the complexity of our proposed methods is about 120 and 190 times less than the baseline, respectively. Moreover, in multi-path fading NLoS channels, our proposed SV-based techniques can achieve the BICM rate requirements with a lower complexity than the exhaustive search baseline. Therefore, our proposed algorithms can achieve a high performance while reducing a computational complexity in MIMO-BICM systems, rendering them suitable for practical implementations.

APPENDIX A
DERIVATION OF LEMMA 1

The differential of $\tilde{L}_{ik}$ is given by

$$d\tilde{L}_{iq} = \frac{1}{\sigma^2}\left(d\boldsymbol{z}_0^H\boldsymbol{z}_0 + \boldsymbol{z}_0^H d\boldsymbol{z}_0 - d\boldsymbol{z}_1^H\boldsymbol{z}_1 - \boldsymbol{z}_1^H d\boldsymbol{z}_1\right) \tag{51}$$

where $\boldsymbol{z}_0 = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^0$, $\boldsymbol{z}_1 = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^1$, and

$$\begin{aligned} d\boldsymbol{z}_0 &= d\boldsymbol{y} - \boldsymbol{H}d\boldsymbol{P}\boldsymbol{s}_{iq}^0 = \boldsymbol{H}d\boldsymbol{P}(\boldsymbol{s}-\boldsymbol{s}_{iq}^0) \\ d\boldsymbol{z}_1 &= d\boldsymbol{y} - \boldsymbol{H}d\boldsymbol{P}\boldsymbol{s}_{iq}^1 = \boldsymbol{H}d\boldsymbol{P}(\boldsymbol{s}-\boldsymbol{s}_{iq}^1) \end{aligned} \tag{52}$$

Inserting (52) into (51), we obtain

$$d\tilde{L}_{ik} = \frac{1}{\sigma^2}\left[d\boldsymbol{P}^H\frac{\partial\tilde{L}_{ik}}{\partial\boldsymbol{P}^*} + \left(\frac{\partial\tilde{L}_{ik}}{\partial\boldsymbol{P}^*}\right)^H d\boldsymbol{P}\right] \tag{53}$$

where

$$\begin{aligned} \frac{\partial\tilde{L}_{iq}}{\partial\boldsymbol{P}^*} = \frac{1}{\sigma^2}\boldsymbol{H}^H\big[&(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^0)(\boldsymbol{s}-\boldsymbol{s}_{iq}^0)^H \\ &-(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_{iq}^1)(\boldsymbol{s}-\boldsymbol{s}_{iq}^1)^H\big] \end{aligned} \tag{54}$$

This completes the proof.

REFERENCES

[1] A. Martinez, A. Guillen i Fabregas, G. Caire, and F. M. J. Willems, "Bit-interleaved coded modulation revisited: A mismatched decoding perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2756–2765, 2009.

[2] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[3] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, 1998.

[4] U. Wachsmann, R. Fischer, and J. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1361–1391, 1999.

[5] S. Muller-Weinfurtner, "Coding approaches for multiple antenna transmission in fast fading and OFDM," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2442–2450, 2002.

[6] E. Biglieri, G. Taricco, and E. Viterbo, "Bit-interleaved time-space codes for fading channels," in *Proc. Conference on Information Science and Systems*, 2000, pp. 15–17.

[7] M. Vu and A. Paulraj, "MIMO wireless linear precoding," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 86–105, 2007.

[8] F. Perez-Cruz, M. R. D. Rodrigues, and S. Verdu, "MIMO Gaussian Channels With Arbitrary Inputs: Optimal Precoding and Power Allocation," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1070–1084, 2010.

[9] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally Optimal Linear Precoders for Finite Alphabet Signals Over Complex Vector Gaussian Channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3301–3314, 2011.

[10] J. Jin, Y. R. Zheng, W. Chen, and C. Xiao, "Generalized Quadratic Matrix Programming: A Unified Framework for Linear Precoding With Arbitrary Input Distributions," *IEEE Transactions on Signal Processing*, vol. 65, no. 18, pp. 4887–4901, 2017.

[11] ——, "Hybrid Precoding for Millimeter Wave MIMO Systems: A Matrix Factorization Approach," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3327–3339, 2018.

[12] L.-J. Lampe, R. Schober, and R. Fischer, "Multilevel coding for multiple-antenna transmission," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 203–208, 2004.

[13] N. Gresset, J. J. Boutros, and L. Brunel, "Optimal linear precoding for BICM over MIMO channels," in *IEEE International Symposium on Information Theory*, 2004, pp. 66–66.

[14] E. Akay, E. Sengul, and E. Ayanoglu, "Bit interleaved coded multiple beamforming," *IEEE Transactions on Communications*, vol. 55, no. 9, pp. 1802–1811, 2007.

[15] H. J. Park, B. Li, and E. Ayanoglu, "Constellation precoded multiple beamforming," *IEEE Transactions on Communications*, vol. 59, no. 5, pp. 1275–1286, 2011.

[16] N. Gresset and M. Khanfouci, "Precoded BICM design for MIMO transmit beamforming and associated low-complexity algebraic receivers," in *Proc. IEEE GLOBECOM - Global Telecommunications Conference*, 2008, pp. 1–5.

[17] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[18] S. Schwarz, M. Wrulich, and M. Rupp, "Mutual information based calculation of the precoding matrix indicator for 3GPP UMTS/LTE," in *Proc. International ITG Workshop on Smart Antennas (WSA)*, 2010, pp. 52–58.

[19] W. Chen and S. Jin, "Performance evaluation of closed loop transmission for LTE-A uplink MIMO," in *Proc. 7th International Conference on Wireless Communications, Networking and Mobile Computing*, 2011, pp. 1–4.

[20] F. Penna, H. Cheng, and J. Lee, "A search-free algorithm for precoder selection in FD-MIMO systems with DFT-based codebooks," in *Proc. IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–6.

[21] T. Akyıldız and T. M. Duman, "Search-free precoder selection for 5G new radio using neural networks," in *Proc. IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2020, pp. 1–6.

[22] M. Maleki, J. Jin, and M. Haardt, "Low complexity PMI selection for BICM-MIMO rate maximization in 5G new radio systems," in *Proc. 31st European Signal Processing Conference (EUSIPCO)*, 2023.

[23] ETSI, "LTE; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception; part 101: User equipment (UE) radio transmission and reception (release 16)," ETSI, Tech. Rep. 136 101 v16.7.0, 2020. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/136100_136199/136101/16.07.00_60/ts_136101v160700p.pdf

[24] 3GPP, "TR 38.901 (v16.1.0): Study on channel model for frequencies from 0.5 to 100 GHz (release 16)," 3rd Generation Partnership Project (3GPP), Tech. Rep. 38.901, Nov. 2020. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/16.01.00_60/tr_138901v160100p.pdf

[25] C. Studer, A. Burg, and H. Bolcskei, "Soft-output Sphere Decoding: Algorithms and VLSI Implementation," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 290–300, 2008.

[26] Q. Qi and C. Chakrabarti, "Parallel High Throughput Soft-output Sphere Decoder," in *Proc. IEEE Workshop Signal Process. Syst.*, 2010, pp. 174–179.

[27] M. Payaro and D. P. Palomar, "On Optimal Precoding in Linear Vector Gaussian Channels With Arbitrary Input Distribution," in *Proc. IEEE Int. Symp. Inf. Theory*, 2009, pp. 1085–1089.

[28] S. K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "Precoding by Pairing Subchannels to Increase MIMO Capacity With Discrete Input Alphabets," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4156–4169, 2011.

[29] M. R. McKay and I. B. Collings, "Capacity and performance of MIMO-BICM with zero-forcing receivers," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 74–83, Jan. 2005.

[30] M. R. McKay, I. B. Collings, and A. M. Tulino, "Achievable sum rate of MIMO MMSE receivers: A general analytic framework," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 396–410, 2010.

[31] S. Ahmadi, *5G NR: Architecture, technology, implementation, and operation of 3GPP new radio standards*. Academic Press, 2019.

[32] 3rd Generation Partnership Project, "3GPP TS 38.214 V16.7.0 (2021-09): Technical Specification," 3rd Generation Partnership Project; Technical Specification Group Radio Access Network, Technical Specification TS 38.214, September 2021, release 16.

[33] L. Jackson, D. Tufts, F. Soong, and R. Rao, "Frequency estimation by linear prediction," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'78)*, vol. 3, 1978, pp. 352–356.

[34] P. Händel and I. Kiss, "On correlation based single tone frequency estimation," in *Proc. Finnish Signal Processing Symposium*, 1997, pp. 32–36.

[35] D. A. Gore, R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, 2000, pp. 2785–2788.

[36] S. Sandhu, R. U. Nabar, D. A. Gore, and A. Paulraj, "Near-optimal selection of transmit antennas for a MIMO channel based on shannon capacity," in *Proc. Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*, vol. 1, 2000, pp. 567–571.