Quadrature Permutation Matrix Modulation

Burak Ozpoyraz, *Graduate Student Member, IEEE*, Atalay Aydin, *Graduate Student Member, IEEE*, and Ertugrul Basar, *Fellow, IEEE*

***Abstract*—This paper introduces the in-phase and quadrature (IQ) extension to the permutation matrix modulation (PMM) technique and proposes the quadrature PMM (QPMM) scheme for higher spectral efficiency without any expense in RF chains. Assuming a single-user MIMO (SU-MIMO) system model with all antennas activated, the proposed QPMM scheme utilizes the permutation matrices as the spatial indexing unit for both the IQ components of the complex symbol vector. Furthermore, a low-complex detector, conditional maximum likelihood detector (C-MLD), that achieves the same bit error rate (BER) performance as the optimal joint MLD is presented. In addition to the low-complex C-MLD detector, a deep learning (DL) based detector, called FusionDet, is also proposed for the QPMM scheme. This DL-based detector provides a trade-off between complexity and BER performance. The BER performance of the QPMM scheme and the complexities of the detectors are examined. The provided computer simulations reveal that the proposed QPMM scheme outperforms the conventional PMM method for different MIMO setups and modulation levels. In addition, the complexity analysis shows that C-MLD attains the same BER performance as the optimal joint MLD with significantly lower complexity. Finally, FusionDet is observed to further decrease detector complexity in exchange for sub-optimal detection performance.**

***Index Terms*—Index modulation (IM), multiple-input multiple-output (MIMO), deep learning (DL), neural network (NN).**

I. Introduction

WIRELESS communication systems have been facing increasing demands for extreme data rates, substantial spectral efficiency, and considerably low hardware complexity. However, traditional communication schemes are restricted to transmitting data through the amplitude and phase of the transmitted signal, which limits the ability to meet the demands of modern communication systems. Index modulation (IM) provides a promising solution by exploiting the indices of the building blocks of wireless communication systems, such as transmit antennas, radio frequency (RF) mirrors, modulation types, and subcarriers, to convey additional information bits. Alternatively, IM schemes offer the capability to transmit information more efficiently regarding energy consumption. This energy efficiency is achieved by selectively deactivating specific key components of the system yet utilizing them for the purpose of data transmission. Therefore, IM is a revolutionary data transmission technique that enables the creation of entirely new dimensions for communication [1].

Despite the advantages associated with reduced power consumption and straightforward design, IM-aided multiple-input multiple-output (MIMO) schemes encounter various practical challenges that need to be resolved [2], [3]. One of the primary obstacles MIMO schemes employing IM face is the trade-off between detection complexity and accuracy. While there exist analytical solutions proposed for MIMO-IM detection, their complexity becomes increasingly impractical as the number of antennas grows, which is counterproductive. Given the modern trend toward designing communication systems with large MIMO configurations, the complexity issue associated with MIMO-IM becomes crucial. Consequently, instead of pursuing analytical approaches to address the detection challenge, deep learning (DL) solutions are devised to establish a data-driven algorithm that achieves satisfactory performance while maintaining low complexity, even in the context of large MIMO setups.

This paper proposes the quadrature permutation matrix modulation (quadrature PMM, QPMM) scheme. The QPMM scheme employs double indexing in the spatial domain, with the spatial bits modulated to two permutation matrices for the in-phase and quadrature (IQ) components of the transmit symbol vector in a single-user MIMO (SU-MIMO) system model. The remaining data bits generate the amplitude-phase modulated (APM) symbols. The IQ components of the transmit symbol vector are separately precoded using the singular values of the channel matrix and corresponding permutation matrix, similar to the conventional PMM method [4]. A low-complex detector, called conditional maximum likelihood detector (conditional MLD, C-MLD), is introduced, where APM symbols are decoded separately. In addition, a DL-based QPMM detector, named FusionDet, is presented to obtain a hybrid architecture that allows a trade-off between detection complexity and bit error rate (BER) performance. BER performance over Rayleigh fading channels is analyzed and compared to the conventional PMM method. Moreover, the BER performance of FusionDet is investigated considering different trade-off points depending on the FusionDet settings. In addition, the complexity of the proposed detectors are analyzed. The contributions of this paper are threefold. First, a novel IM-aided MIMO scheme, referred to as QPMM is introduced. Second, a low-complexity detector for both PMM and QPMM, named C-MLD is presented. Finally, a hybrid DL-based detector for QPMM, termed FusionDet, is proposed.

The rest of the paper is organized as follows. In Section II, a brief literature on IM-aided MIMO and DL-aided IM schemes

Received 25 March 2024; revised 2 September 2024, 6 January 2025, and 30 April 2025; accepted 12 June 2025. Date of publication 26 June 2025; date of current version 22 December 2025. The associate editor coordinating the review of this article and approving it for publication was A. Liu. *(Corresponding author: Ertugrul Basar.)*

Burak Ozpoyraz and Atalay Aydin are with the Department of Electrical and Electronics Engineering, Koç University, Sarıyer, 34450 Istanbul, Turkiye (e-mail: bozpoyraz20@ku.edu.tr; atalayaydin22@ku.edu.tr).

Ertugrul Basar is with the Department of Electrical Engineering, Tampere University, 33720 Tampere, Finland, on leave from the Department of Electrical and Electronics Engineering, Koç University, Sarıyer, 34450 Istanbul, Turkiye (e-mail: ertugrul.basar@tuni.fi).

Digital Object Identifier 10.1109/TWC.2025.3581483

© 2025 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

will be provided. Section III will present the proposed QPMM scheme along with the low-complex C-MLD detector. The detection methods for the PMM scheme will be given in Section IV. The hybrid structure of the data- and model-driven FusionDet detector will be provided in Section V. Section VI will analyze the theoretical BER upper bound of the QPMM scheme as well as the complexity of both PMM and QPMM detectors. Section VII will present the simulation results, and Section VIII will conclude the paper.

II. Related Works

A. MIMO-IM Literature

Numerous schemes have been proposed to leverage IM in MIMO systems by exploiting multiple antennas. Spatial modulation (SM) [5], which utilizes the indices of the transmit antennas of a MIMO system to convey additional information, is by far the most widely known application of MIMO-IM. Space shift keying (SSK) [6] is a simpler implementation of IM in MIMO systems than SM, as it only transmits information through the indices of the active transmit antennas. To further increase the spectral efficiency (SE) of SM, in generalized SM (GSM) [7], the number of active transmit antennas can vary beyond one, and the same data symbol is transmitted simultaneously over multiple selected active transmit antennas. The precoding-aided SM (PSM), also known as the receive SM (RSM), conveys spatial bits on receive antennas [8] which requires channel state information at transmitter (CSIT) instead of receiver. In a PSM-aided MIMO system, all transmit antennas are activated, and the APM symbol is beamformed to a single receive antenna to utilize receive antenna index for conveying information unlike SM and SSK schemes that exploit transmit antenna indices. The authors of [9] propose a generalized PSM (GPSM) scheme in which multiple receive antennas are selected, and a different APM symbol is beamformed to each receive antenna chosen for more efficient use of spatial resources. Precoding enables symbol-by-symbol detection for each APM symbol, reducing the complexity of the MLD without sacrificing the BER performance. Quadrature SM (QSM) [10] is another improvement on SM that independently maps the in-phase and quadrature (IQ) components of the RF carrier to different transmit antennas, depending on different bits of the source data. This technique enhances the overall throughput of SM with the extra spatial dimension. Enhanced SM (ESM) [11] is an extension of classical SM that conveys additional information by exploiting the signal constellations. The GPSM scheme proposed in [9] is combined with the QSM concept in [12]. A single receive antenna combination is selected for each IQ component of the APM symbols, doubling the spatial bits compared to GPSM. Another clever indexing method is presented in [13], where the authors consider the GPSM scheme in the power domain. APM symbols with a particular power of P_1 are steered to the selected receive antennas, whereas APM symbols with a power of $P_2 \neq P_1$ are beamformed to the remaining receive antennas. Thus, each receive antenna is used for data reception in this dual-layer transmission, increasing bandwidth efficiency over GPSM. The study in [14] conveys additional information bits on both transmit and receive antennas, called joint SM (JSM). Transmit antennas are divided into non-overlapping groups, and a group is selected for data transmission at each time slot. Receive antennas are divided into two groups: RSM and diversity antennas. A single receive antenna is selected from RSM antennas to beamform the APM symbol in each channel use. Diversity antennas receive the same APM symbol to introduce a receive diversity and improve the detection performance. The authors of [15] extend the concept of IM into multi-user scenarios and present a novel downlink multi-user (MU)-MIMO system based on GPSM of [9], the dual-layer transmission of [13], and JSM of [14]. One of the non-overlapping transmit antenna groups beamform the APM symbols with powers of P_1 and P_2 to the selected and remaining receive antennas, respectively, for each user. The study in [16] employs PSM in over-determined MIMO systems, where the number of receive antennas is higher than the number of transmit antennas. The linear precoding schemes, such as zero-forcing (ZF) and minimum mean squared error (MMSE) precoding, require that the number of transmit antennas is higher than or equal to the number of receive antennas. Therefore, the authors of [16] introduce the grouped PSM scheme in which the receive antennas are divided into non-overlapping groups, each with the antennas fewer than or equal to the number of transmit antennas. The information bits consist of three units for selecting the active group, the active receive antenna in that group, and the APM symbol. The grouped PSM scheme is further extended to the generalized grouped PSM, where multiple groups and multiple receive antennas from each group are selected. The conventional QSM scheme is combined with the space-time block codes (STBC) in [17] to introduce a transmit diversity. Despite using Alamouti's STBC algorithm, the proposed scheme is generalizable to other STBC methods. Information bits are mapped to the complete 3D constellation consisting of active antenna indices (spatial dimension) and the 2D M-ary symbols in [18], which makes it possible to vary the number of active antennas in each transmit vector. Authors in [19] proposes a novel domain, the spreading code domain, and introduces the code index modulation-spread spectrum (CIM-SS) system. In CIM-SS scheme, data is mapped to spreading codes and constellation symbols, allowing only half of the bits to be physically transmitted and the other half to be encoded in the spreading codes. In generalized CIM-SM (GCIM-SM) [20], SM is combined with CIM-SS, which uses incoming information bits to determine modulated symbols, active transmit antenna indices, and corresponding spreading code indices to enhance data rate. Last but not least, the advanced PSM scheme of [4], called PMM, employs spatial permutation matrices as additional information-carrying units. Singular value decomposition (SVD) based precoding is implemented rather than the well-known linear precoding methods such as ZF and MMSE. In addition, the authors present a low-complexity detector that can only work with the phase-shift keying (PSK) modulated symbols.

Our proposed scheme, QPMM, and PMM [4] are able to index transmit antennas while keeping them active, which is a prospective advantage over conventional SM approaches. Due to more efficient handling of spatial resources in QPMM,

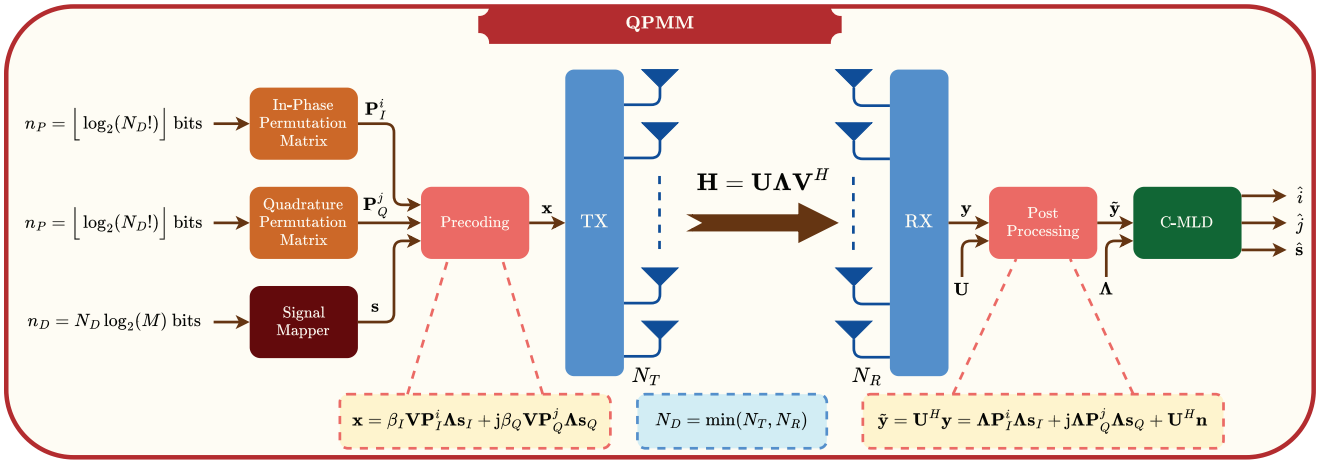


Fig. 1. The system model of the QPMM scheme.

a certain level of data rate can be achieved with a lower modulation order. Moreover, our scheme improves spectral efficiency without any expense in RF chains by introducing the IQ extension to the PMM [4] scheme. In addition to that, the proposed low complexity detectors in this paper work for both PSK and QAM modulation, while proposing more feasible system designs that also integrates DL into the solution.

B. DL-Based IM Literature

In addition to the comprehensive body of literature previously examined pertaining to the design of IM systems, it is noteworthy that DL-based architectures have emerged within the domain of IM, with a particular focus on addressing high-complexity tasks, such as detection. In [21], the authors propose a novel DL-based detector for MIMO-IM systems, called IMRecoNet, which formulates the detection process as a sparse reconstruction problem and designs a complex value neural network (NN) to realize the sparse reconstruction process in a computationally efficient manner. DeepSM [22] is a DL-based detector for SM that updates the channel state information (CSI) at each time slot before detecting the data, making it suitable for operation in time-varying channels. Note that, DL-based solutions are utilized in other IM schemes as well, such as OFDM-IM [23], [24]. In [25], authors use DL-aided signal detection for two-stage index modulated universal filtered multi-carrier systems for underwater acoustic (UWA) communication. A DL-assisted detection scheme for millimeter wave systems with IM is proposed in [26]. In [27], a receiver with DL detector is designed for X-transform time-domain synchronous IM (TDS-IM) OFDM.

III. QPMM System Model

A MIMO system with N_T transmit and N_R receive antennas is assumed, as given in Fig. 1. The system model indicates that the spatial bits consist of two permutation bit units to be modulated into permutation matrices for the IQ components of the complex APM symbols, which doubles the number of spatial bits in the conventional PMM method. The perfect CSI is assumed to be available at both sides for the SVD-based precoding, post-processing, and detection.

A. Transmitter

The incoming bits are divided into three building blocks. The first $n_P = \lfloor \log_2(N_D!) \rfloor$ permutation bits select $\mathbf{P}_I^i = \left[\mathbf{e}_1^i\ \mathbf{e}_2^i \dots \mathbf{e}_{N_D}^i\right]$ from the $\mathcal{P}^{N_P}$ set, where $N_D = \min(N_T, N_R)$ is the number of APM symbols transmitted in a time slot, $\mathbf{P}_I^i \in \mathbb{R}^{N_D \times N_D}$ is the in-phase permutation matrix corresponding to the i^{th} permutation, $\mathbf{p}_i$, of N_D, $\mathbf{e}_k^i \in \mathbb{R}^{N_D \times 1}$ is the standard basis column vector with zeros except one at the $\mathbf{p}_i(k)^{\text{th}}$ element, and $\mathcal{P}$ and N_P are the set and number of legitimate permutation matrices, respectively. The second n_P bits select $\mathbf{P}_Q^j = \left[\mathbf{e}_1^j\ \mathbf{e}_2^j \dots \mathbf{e}_{N_D}^j\right]$ from the $\mathcal{P}^{N_P}$ set, where $\mathbf{P}_Q^j \in \mathbb{R}^{N_D \times N_D}$ is the quadrature permutation matrix corresponding to $\mathbf{p}_j$. The remaining $n_D = N_D \log_2(M)$ bits construct the APM symbol vector $\mathbf{s} = \left[s_1\ s_2 \dots s_{N_D}\right]^T$ from $\mathcal{S}^{M^{N_D}}$ where M is the modulation level, s_k is the k^{th} APM symbol that is selected from an M-ary constellation set $\mathcal{C}$ consisting of unit power elements, e.g., $\mathbb{E}[|s_k|^2] = 1$, and $\mathcal{S}$ and M^{N_D} are the set and the number of all possible APM symbol vectors, respectively. An example of $\mathcal{P}$ for $N_D = 3$ is given in Table I. The number of the total transmitted bits during a time slot is given as follows:

$$\begin{aligned} l &= 2n_P + n_D, \\ &= 2\lfloor \log_2(N_D!) \rfloor + N_D \log_2(M). \end{aligned} \tag{1}$$

The channel matrix $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is assumed to follow flat fading and consists of independent and identically distributed (i.i.d.) complex Gaussian distributed random variables with zero-mean and unit variance. The SVD of $\mathbf{H}$ can be obtained as given below:

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H \tag{2}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\mathbf{\Lambda}$ is the diagonal matrix including the eigenvalues of $\mathbf{H}$. It should be noted that $\mathbf{U}$ and $\mathbf{V}$ are $(N_R \times N_T)$ and $(N_T \times N_T)$ matrices, respectively, for $N_T \leq N_R$, whereas they are $(N_R \times N_R)$ and $(N_R \times N_T)$ matrices, respectively, for $N_T > N_R$. The transmit symbol vector obtained after the pre-processing of the APM symbols using the singular values of the MIMO channel is expressed as follows:

$$\mathbf{x} = \beta_I \mathbf{Z}_I \mathbf{s}_I + j\beta_Q \mathbf{Z}_Q \mathbf{s}_Q,$$

TABLE I
THE SET OF LEGITIMATE PERMUTATION MATRICES $\mathcal{P}$ FOR $N_D = 3$

Permutation Bits	i	$\mathbf{p}_i$	$\mathbf{P}_i$
$[0 \quad 0]$	1	$\mathbf{p}_1 = [1 \quad 2 \quad 3]$	$\mathbf{P}_1 = \begin{bmatrix} \mathbf{e}_1^{1^T} \\ \mathbf{e}_2^{1^T} \\ \mathbf{e}_3^{1^T} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$[0 \quad 1]$	2	$\mathbf{p}_2 = [1 \quad 3 \quad 2]$	$\mathbf{P}_2 = \begin{bmatrix} \mathbf{e}_1^{2^T} \\ \mathbf{e}_2^{2^T} \\ \mathbf{e}_3^{2^T} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
$[1 \quad 0]$	3	$\mathbf{p}_3 = [2 \quad 1 \quad 3]$	$\mathbf{P}_3 = \begin{bmatrix} \mathbf{e}_1^{3^T} \\ \mathbf{e}_2^{3^T} \\ \mathbf{e}_3^{3^T} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$[1 \quad 1]$	4	$\mathbf{p}_4 = [2 \quad 3 \quad 1]$	$\mathbf{P}_4 = \begin{bmatrix} \mathbf{e}_1^{4^T} \\ \mathbf{e}_2^{4^T} \\ \mathbf{e}_3^{4^T} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

$$= \beta_I \mathbf{V}\mathbf{P}_I\mathbf{\Lambda}\mathbf{s}_I + j\beta_Q\mathbf{V}\mathbf{P}_Q\mathbf{\Lambda}\mathbf{s}_Q, \tag{3}$$

where $\mathbf{Z}_I$, $\mathbf{Z}_Q$, β_I, and β_Q are the precoding matrices and the power normalization factors of the IQ components $\mathbf{s}_I$ and $\mathbf{s}_Q$, respectively. The power normalization factors are obtained as below:

$$\beta_I = \sqrt{\frac{P_{tot}/2}{||\mathbf{Z}_I\mathbf{s}_I||^2}}, \quad \beta_Q = \sqrt{\frac{P_{tot}/2}{||\mathbf{Z}_Q\mathbf{s}_Q||^2}}, \tag{4}$$

where P_{tot} is the total transmitted power within a time slot.

B. Receiver

The received signal vector $\mathbf{y}$ is given as follows:

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{n}, \\ &= \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H[\beta_I\mathbf{Z}_I\mathbf{s}_I + j\beta_Q\mathbf{Z}_Q\mathbf{s}_Q] + \mathbf{n}, \\ &= \beta_I\mathbf{U}\mathbf{\Lambda}\mathbf{V}^H\mathbf{V}\mathbf{P}_I\mathbf{\Lambda}\mathbf{s}_I + j\beta_Q\mathbf{U}\mathbf{\Lambda}\mathbf{V}^H\mathbf{V}\mathbf{P}_Q\mathbf{\Lambda}\mathbf{s}_Q + \mathbf{n}, \\ &= \beta_I\mathbf{U}\mathbf{\Lambda}\mathbf{P}_I\mathbf{\Lambda}\mathbf{s}_I + j\beta_Q\mathbf{U}\mathbf{\Lambda}\mathbf{P}_Q\mathbf{\Lambda}\mathbf{s}_Q + \mathbf{n}, \end{aligned} \tag{5}$$

where $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is the additive white Gaussian noise (AWGN) vector consisting of complex Gaussian distributed random variables with zero-mean and N_0 variance. It should be noted that unitary matrices satisfy $\mathbf{U}^H\mathbf{U} = \mathbf{V}^H\mathbf{V} = \mathbf{I}$.

The received signal vector is post-processed to mitigate the interference between the receive antennas and allow symbol-by-symbol detection, as given below:

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{U}^H\mathbf{y}, \\ &= \beta_I\mathbf{U}^H\mathbf{U}\mathbf{\Lambda}\mathbf{P}_I\mathbf{\Lambda}\mathbf{s}_I + j\beta_Q\mathbf{U}^H\mathbf{U}\mathbf{\Lambda}\mathbf{P}_Q\mathbf{\Lambda}\mathbf{s}_Q + \mathbf{U}^H\mathbf{n}, \\ &= \beta_I\mathbf{\Lambda}\mathbf{P}_I\mathbf{\Lambda}\mathbf{s}_I + j\beta_Q\mathbf{\Lambda}\mathbf{P}_Q\mathbf{\Lambda}\mathbf{s}_Q + \tilde{\mathbf{n}}. \end{aligned} \tag{6}$$

Since $\mathbf{U}$ is a unitary matrix, the post-processing does not change the noise statistics.

C. Joint Maximum Likelihood Detector

The joint MLD searches for all combinations of the in-phase and quadrature permutation matrix and APM symbol vector, which can be expressed as given below:

$$\begin{aligned} [\hat{i}, \quad \hat{j}, \quad \hat{\mathbf{s}}] = \underset{\substack{i \in \{1,\ldots N_P\} \\ j \in \{1,\ldots N_P\} \\ \mathbf{s} \in \mathcal{S}}}{\arg\min} \quad & ||\mathbf{d}||^2, \\ \mathbf{d} = \tilde{\mathbf{y}} - \beta_I\mathbf{\Lambda}\mathbf{P}_I^i\mathbf{\Lambda}\mathbf{s}_I - j\beta_Q\mathbf{\Lambda}\mathbf{P}_Q^j\mathbf{\Lambda}\mathbf{s}_Q. & \end{aligned} \tag{7}$$

D. Conditional Maximum Likelihood Detector

The complexity of the joint MLD grows with the increase in the number of APM symbols and the channel dimensions. Therefore, the low-complex C-MLD is introduced, which decodes APM symbols separately, attaining the same BER performance as the joint MLD. C-MLD calculates the sum of each APM symbol's minimum decision metric for each (i, j) combination. The combination with the minimum metric sum is decoded as the spatial information. The decoded APM symbols correspond to the APM symbols that provide the minimum metrics for the decoded (i, j) combination.

The APM symbol that provides the minimum decision metric for s_k and any (i, j) combination is given as below:

$$\hat{s}_k^{i,j} = \underset{s \in \mathcal{C}}{\arg\min} \, |\bar{y}_k - (\beta_I r_I^{i,k} s_I + j\beta_Q r_Q^{j,k} s_Q)|^2, \tag{8}$$

where $\bar{y}_k$ is the received symbol corresponding to s_k, which can be obtained as

$$\bar{y}_k = \Re\{\tilde{y}_{\mathbf{p}_i(k)}\} + j\Im\{\tilde{y}_{\mathbf{p}_j(k)}\}, \tag{9}$$

where $\tilde{y}_a$ is the a^{th} element of $\tilde{\mathbf{y}}$. In addition, $r_I^{i,k}$ is the element at the $\mathbf{p}_i(k)^{\text{th}}$ row and the k^{th} column of $\mathbf{R}_I = \mathbf{\Lambda}\mathbf{P}_I^i\mathbf{\Lambda}$, and $r_Q^{j,k}$ is the element at the $\mathbf{p}_j(k)^{\text{th}}$ row and the k^{th} column

of $\mathbf{R}_Q = \mathbf{\Lambda}\mathbf{P}_Q^j\mathbf{\Lambda}$. The minimum decision metric of s_k is calculated as follows:

$$\epsilon_k^{i,j} = |\bar{y}_k - (\beta_I r_I^{i,k} \hat{s}_{k_I}^{i,j} + j\beta_Q r_Q^{j,k} \hat{s}_{k_Q}^{i,j})|^2. \tag{10}$$

The sum of each APM symbol's minimum decision metric for any (i, j) combination is obtained as given below:

$$\bar{\epsilon}^{i,j} = \sum_{k=1}^{N_D} \epsilon_k^{i,j}. \tag{11}$$

Afterward, the selected (i, j) combination is decoded as follows:

$$\left[\hat{i}, \quad \hat{j}\right] = \underset{i,j \in \{1,\ldots,N_P\}}{\arg\min} \bar{\epsilon}^{i,j}. \tag{12}$$

Finally, the k^{th} APM symbol is decoded according to the $(\hat{i}, \hat{j})$ combination as follows:

$$\hat{s}_k = \hat{s}_k^{\hat{i},\hat{j}}. \tag{13}$$

To further explain the detection procedure, a numerical example will be given. Let $N_T = N_R = N_D = 3$, $i = 2$, and $j = 3$. $\mathbf{P}_I$ and $\mathbf{P}_Q$ are selected from Table I according to i and j values as follows.

$$\mathbf{P}_I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{14}$$

$$\mathbf{P}_Q = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{15}$$

The post-processed received signal vector $\tilde{\mathbf{y}}$ under the assumption of $\beta_I = \beta_Q = 1$ and without noise is obtained as follows.

$$\tilde{\mathbf{y}} = \mathbf{\Lambda}\mathbf{P}_I\mathbf{\Lambda}\mathbf{s}_I + j\mathbf{\Lambda}\mathbf{P}_Q\mathbf{\Lambda}\mathbf{s}_Q, \tag{16}$$

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \tilde{y}_3 \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} s_{1I} \\ s_{2I} \\ s_{3I} \end{bmatrix} + j \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} s_{1Q} \\ s_{2Q} \\ s_{3Q} \end{bmatrix} \tag{17}$$

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \tilde{y}_3 \end{bmatrix} = \begin{bmatrix} a^2 & 0 & 0 \\ 0 & 0 & bc \\ 0 & bc & 0 \end{bmatrix} \begin{bmatrix} s_{1I} \\ s_{2I} \\ s_{3I} \end{bmatrix} + j \begin{bmatrix} 0 & ab & 0 \\ ab & 0 & 0 \\ 0 & 0 & c^2 \end{bmatrix} \begin{bmatrix} s_{1Q} \\ s_{2Q} \\ s_{3Q} \end{bmatrix}. \tag{18}$$

As a result of this derivation,

$$\tilde{y}_1 = a^2 s_{1I} + jabs_{2Q}. \tag{19}$$

$$\tilde{y}_2 = bcs_{3I} + jabs_{1Q}. \tag{20}$$

$$\tilde{y}_3 = bcs_{2I} + jc^2 s_{3Q}. \tag{21}$$

Note that at this point, symbol-by-symbol decoding is not possible. However, obtaining $\overline{\mathbf{y}}$ as follows will resolve this problem.

$$\begin{aligned} \overline{y}_1 &= \Re\left\{\tilde{y}_{\mathbf{p}_2(1)}\right\} + j\Im\left\{\tilde{y}_{\mathbf{p}_3(1)}\right\}, \\ &= \Re\{\tilde{y}_1\} + j\Im\{\tilde{y}_2\}, \\ &= a^2 s_{1I} + jabs_{1Q}. \end{aligned} \tag{22}$$

$$\begin{aligned} \overline{y}_2 &= \Re\left\{\tilde{y}_{\mathbf{p}_2(2)}\right\} + j\Im\left\{\tilde{y}_{\mathbf{p}_3(2)}\right\}, \\ &= \Re\{\tilde{y}_3\} + j\Im\{\tilde{y}_1\}, \\ &= bcs_{2I} + jabs_{2Q}. \end{aligned} \tag{23}$$

$$\begin{aligned} \overline{y}_3 &= \Re\left\{\tilde{y}_{\mathbf{p}_2(3)}\right\} + j\Im\left\{\tilde{y}_{\mathbf{p}_3(3)}\right\}, \\ &= \Re\{\tilde{y}_2\} + j\Im\{\tilde{y}_3\}, \\ &= bcs_{3I} + jc^2 s_{3Q}. \end{aligned} \tag{24}$$

Resulting $\overline{\mathbf{y}}$ is suitable for symbol-by-symbol decoding.

IV. PMM Detectors

As an alternative to the joint MLD for the conventional PMM scheme, the authors of [4] present a low-complexity detector, which will be called LCD throughout the simulations. The main drawback of LCD is the requirement of PSK modulation during transmission. The proposed C-MLD can also be used for the detection of the conventional PMM scheme, achieving the same BER performance as the joint MLD with considerably lower complexity, like in the QPMM scheme.

The signaling details of the conventional PMM method are given in [4]. Similar to the proposed QPMM scheme, the post-processed received symbol vector of the conventional PMM technique is expressed as below:

$$\tilde{\mathbf{y}} = \beta\mathbf{\Lambda}\mathbf{P}\mathbf{\Lambda}\mathbf{s} + \tilde{\mathbf{n}}, \tag{25}$$

where $\beta = \sqrt{P_{tot}/||\mathbf{V}\mathbf{P}\mathbf{\Lambda}||^2}$ is the power normalization factor and $\mathbf{P}$ is the permutation matrix.

A. Joint Maximum Likelihood Detector

As discussed for the proposed QPMM scheme, the joint MLD aims at finding the combination of the permutation matrix and APM symbol vector that gives the minimum decision metric, as below:

$$\left[\hat{i}, \quad \hat{\mathbf{s}}\right] = \underset{\substack{i \in \{1,\ldots,N_P\} \\ \mathbf{s} \in \mathcal{S}}}{\arg\min} ||\tilde{\mathbf{y}} - \beta\mathbf{\Lambda}\mathbf{P}_i\mathbf{\Lambda}\mathbf{s}||^2. \tag{26}$$

B. Low-Complexity PMM Detector

The authors of [4] propose a two-stage low-complexity detector. The index, *i*, of the permutation matrix is detected at the first stage as given below:

$$\hat{i} = \underset{i \in \{1,\ldots,N_P\}}{\arg\max} [(\mathbf{P}_i\mathbf{\Lambda}\mathbf{1})^T \text{abs}(\mathbf{\Lambda}^{-1}\tilde{\mathbf{y}})], \tag{27}$$

where $\mathbf{1}$ is an $(N_D \times 1)$ vector consisting of all ones and abs(.) is the element-wise absolute value operation.

The APM symbols are separately decoded at the second stage, which is given as follows:

$$\begin{aligned} \tilde{\mathbf{x}} &= \mathbf{P}_{\hat{i}}^T \tilde{\mathbf{y}}, \\ s_k &= \underset{s \in \mathcal{C}}{\arg\min} |s - \tilde{x}_k|^2, \quad \text{for } k \in \{1,\ldots,N_D\}, \end{aligned} \tag{28}$$

where $\tilde{x}_k$ is the k^{th} element of $\tilde{\mathbf{x}} \in \mathbb{C}^{N_D \times 1}$.

C. Conditional Maximum Likelihood Detector

The proposed low-complex C-MLD with a simple variation can be utilized as the conventional PMM detector. Similar to the C-MLD for QPMM, the APM symbols are decoded separately. Specifically, in the PMM system, C-MLD calculates the sum of each APM symbol's minimum decision metric for each permutation matrix. Subsequently, the permutation matrix index, i, with the minimum metric sum is decoded. The APM symbols yielding minimum metrics for the decoded permutation matrix are demodulated into the data bits. The APM symbol that provides the minimum decision metric for s_k and any permutation matrix index i is obtained as follows:

$$\hat{s}_k^i = \arg\min_{s\in\mathcal{C}} |\tilde{y}_{\mathbf{p}_i(k)} - \beta r^{i,k} s|^2, \tag{29}$$

where $r^{i,k}$ is the element at the $\mathbf{p}_i(k)^{\text{th}}$ row and the k^{th} column of $\mathbf{R} = \mathbf{\Lambda}\mathbf{P}_i\mathbf{\Lambda}$. The minimum decision metric of s_k is expressed as given below:

$$\epsilon_k^i = |\tilde{y}_{\mathbf{p}_i(k)} - \beta r^{i,k} \hat{s}_k^i|^2. \tag{30}$$

The sum of each APM symbol's minimum decision metric for any permutation matrix index, i, is calculated as follows:

$$\bar{\epsilon}^i = \sum_{k=1}^{N_D} \epsilon_k^i. \tag{31}$$

Finally, the selected permutation matrix index, i, and the transmitted APM symbols are decoded as below:

$$\hat{i} = \arg\min_{i\in\{1,\ldots,N_P\}} \bar{\epsilon}^i,$$
$$\hat{s}_k = \hat{s}_k^{\hat{i}}. \tag{32}$$

V. FusionDet: DL-Based QPMM Detector

Within the scope of this work, we introduce a DL-based detector, called FusionDet (FD), which is designed with the primary objective of yielding further improvements in detection complexity. This achievement is realized through the utilization of a trade-off between computational complexity and BER performance.

A. FusionDet High-Level Architecture

To further decrease the complexity associated with the detection process in the context of C-MLD, FusionDet is introduced as a novel DL-based index detector. FusionDet offers the advantage of facilitating index detection with significantly diminished computational complexity. The implementation of a low-complexity deep NN (DNN) for index detection circumvents the computationally expensive requirement of performing symbol detection for each conceivable index pair. Instead, the symbol detection process is exclusively executed for the index pair that has been successfully detected.

FusionDet is a fully connected (FC) DNN-based architecture that is trained for classifying only the indices of a received signal. The high-level architecture is given in Fig. 2, where $\tilde{\mathbf{y}}$ is the post-processed received signal vector, and $flat(\mathbf{H})$ is the row flattened channel matrix. The input feature vector, $\bar{\mathbf{x}}$, is obtained by concatenating real and imaginary parts of $\tilde{\mathbf{y}}$ and

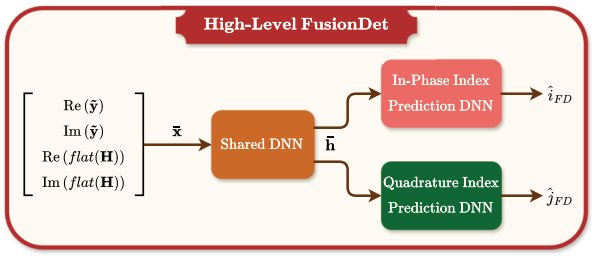


Fig. 2. The high-level architecture of FusionDet.

$flat(\mathbf{H})$. The DNN outputs are in-phase, $\hat{i}_{FD}$, and quadrature, $\hat{j}_{FD}$, index predictions of FusionDet for the received signal.

To scale the input feature vector $\bar{\mathbf{x}}$, min-max normalization (or feature scaling) is applied over the training dataset, as given in Eq. 33.

$$\bar{\mathbf{x}}_{scaled} = \frac{\bar{\mathbf{x}} - \bar{x}_{min}}{\bar{x}_{max} - \bar{x}_{min}}, \tag{33}$$

where $\bar{\mathbf{x}}_{scaled}$ is the scaled input feature vector, and $\bar{x}_{min}$ and $\bar{x}_{max}$ are the minimum and maximum of $\bar{\mathbf{x}}$, respectively.

B. Multi-Task Learning for Im Detection

As depicted in Fig. 2, feature vector is first passed through a shared DNN to obtain a hidden representation vector, $\bar{\mathbf{h}}$. This hidden representation is passed to both in-phase and quadrature prediction DNNs in order to get the predictions for the indices of the current received vector. This approach constitutes an implementation of multi-task learning (MTL). In the realm of DNNs, MTL is an ML technique where a single network is trained to perform multiple related tasks simultaneously, sharing some layers and parameters across tasks. MTL offers several advantages, including reduced computational load, and improved generalization performance.

The shared DNN consists of an input layer and two hidden layers. In a similar fashion, both of the index prediction DNNs are structured with two hidden layers and an output layer, which is subsequently followed by a softmax function to allocate probabilities to each potential index value. The sizing of input-output dimensions for these layers is adaptable and contingent upon the particular arrangement of transmit and receive antennas. As an example, FusionDet architecture for a (4×4) MIMO setup is given in Fig. 3. In this configuration, input is of size 40, where first 8 elements are from post-processed received signal vector and remaining 32 elements are from channel matrix. Input-output dimensions of hidden and output layers are set to 16, since there are 16 possible permutation matrices in this configuration.

C. Integrating FusionDet Into Detection

To utilize FusionDet predictions in detection process, output layers of both index prediction DNNs are passed through the softmax function separately to assign the posterior probabilities of index predictions, namely $P(i = \hat{i}_{FD}|\bar{\mathbf{x}})$ and $P(j = \hat{j}_{FD}|\bar{\mathbf{x}})$. The softmax function operates on an input vector $\mathbf{y}$ comprised of K real numbers, transforming it into a probability distribution with K probabilities that are directly related to the exponential values of the input numbers as given

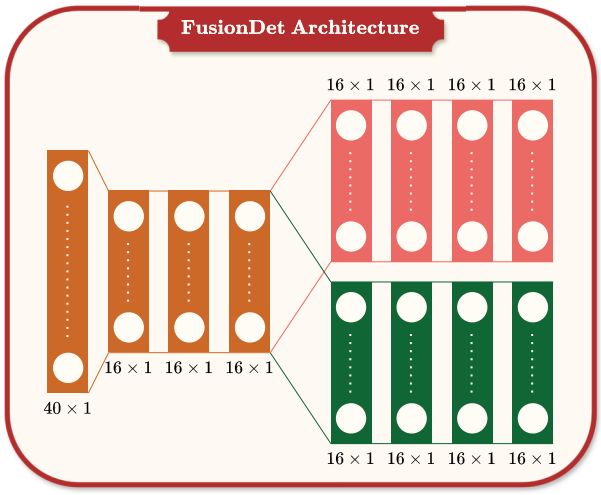


Fig. 3. The DNN architecture of FusionDet layer for (4×4) MIMO setup.

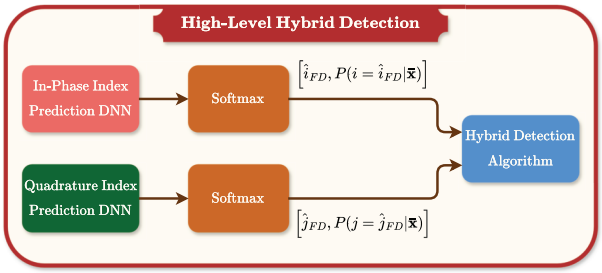


Fig. 4. The high-level integration of FusionDet outputs into the hybrid detection algorithm.

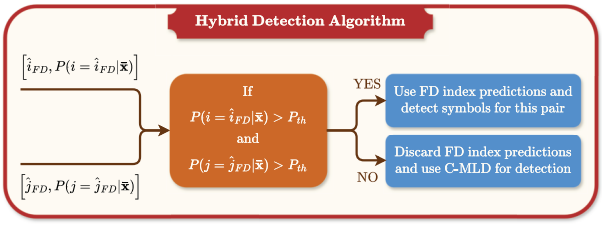


Fig. 5. The hybrid detection algorithm.

in Eq. 34, where y_i is the i^{th} element of vector $\mathbf{y}$. Then, as depicted in Fig. 4, the index predictions and their assigned probabilities are fed into a hybrid detection algorithm.

$$\text{softmax}(\mathbf{y})_i = \frac{e^{y_i}}{\sum_{j=1}^{K} e^{y_j}}, \quad i \in \{1, 2, \ldots, K\}. \tag{34}$$

Hybrid detection algorithm involves two main steps. As the first step of hybrid detection, $P(i = \hat{i}_{FD}|\bar{\mathbf{x}})$ and $P(j = \hat{j}_{FD}|\bar{\mathbf{x}})$ are compared against a probability threshold between 0 and 1, P_{th}. If both of them are above P_{th}, the second step will be symbol detection for $(\hat{i}_{FD}, \hat{j}_{FD})$ as described in C-MLD. In this case, the computational burden of C-MLD is significantly reduced since only a single index pair will be considered. If $P(i = \hat{i}_{FD}|\bar{\mathbf{x}})$ or $P(j = \hat{j}_{FD}|\bar{\mathbf{x}})$ is below P_{th}, then FusionDet predictions are discarded and C-MLD is used to detect both index bits and symbol bits. Hybrid detection algorithm steps are illustrated in Fig. 5.

Note that the purpose of using a threshold, P_{th}, for the posterior probabilities is to provide a flexible trade-off between BER performance and detection complexity. A high P_{th} will result in a BER performance closer to C-MLD. However, the percentage of discarded predictions will increase, resulting in high complexity. On the other hand, a low P_{th} will reduce detection complexity while slightly increasing the BER. Analysis of this trade-off will be discussed in more detail later.

VI. Theoretical Analysis

This section will formulate the theoretical bit error probability (BEP) upper bound of the proposed QPMM scheme in order to validate the computer simulations. Following that, the complexity of the QPMM and PMM detectors will be analyzed to demonstrate the superiority of the proposed C-MLD and FusionDet detectors over the Joint MLD.

A. Bit Error Probability

The theoretical BEP upper bound of the proposed QPMM scheme can be derived by applying the union bound technique employing joint MLD. The received signal vector after post-processing can be given as follows:

$$\tilde{\mathbf{y}} = \beta_I \mathbf{\Lambda} \mathbf{P}_I \mathbf{\Lambda} \mathbf{s}_I + j\beta_Q \mathbf{\Lambda} \mathbf{P}_Q \mathbf{\Lambda} \mathbf{s}_Q + \tilde{\mathbf{n}}. \tag{35}$$

Assuming that (i, j) combination and the transmitted data symbol vector $\mathbf{s}$, represented as $(i, j, \mathbf{s})$, are erroneously detected as $(\bar{i}, \bar{j}, \bar{\mathbf{s}})$, the theoretical BEP upper bound can be expressed as:

$$\begin{aligned} P_b \leq \frac{1}{N_{tot}} \sum_{i=1}^{N_P} \sum_{j=1}^{N_P} \sum_{\mathbf{s} \in \mathcal{S}} \sum_{\bar{i}=1}^{N_P} \sum_{\bar{j}=1}^{N_P} \sum_{\bar{\mathbf{s}} \in \mathcal{S}} \\ \times \frac{e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}})}{\log_2(N_{tot})} \\ \times \bar{P}_e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}}), \end{aligned} \tag{36}$$

where $N_{tot} = N_P^2 M^{N_D}$ is the number of all combinations of $(i, j, \mathbf{s})$, $e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}})$ is the number of bit errors for the corresponding pairwise error event, and $\bar{P}_e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}})$ is the pairwise error probability (PEP) of the erroneous detection. The conditional PEP (CPEP) depending on $\mathbf{H}$ can be given by:

$$\bar{P}_e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}}|\mathbf{H}) = Q\left(\sqrt{\frac{\zeta}{2N_0}}\right), \tag{37}$$

where $\zeta = ||\beta_I \mathbf{\Lambda} \mathbf{P}_I^i \mathbf{\Lambda} \mathbf{s}_I - \bar{\beta}_I \mathbf{\Lambda} \mathbf{P}_I^{\bar{i}} \mathbf{\Lambda} \bar{\mathbf{s}}_I + j\beta_Q \mathbf{\Lambda} \mathbf{P}_Q^i \mathbf{\Lambda} \mathbf{s}_Q - j\bar{\beta}_Q \mathbf{\Lambda} \mathbf{P}_Q^{\bar{i}} \mathbf{\Lambda} \bar{\mathbf{s}}_Q)||^2$. The unconditional PEP can be obtained by taking the expectation of CPEP over $\mathbf{H}$, which can be represented as follows:

$$\bar{P}_e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}}) = \mathbb{E}_{\mathbf{H}}\left[Q\left(\sqrt{\frac{\zeta}{2N_0}}\right)\right]. \tag{38}$$

Furthermore, the Craigs's formula alternative of the Q function can be used as below:

$$Q(x) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} e^{\left(-\frac{x^2}{2\sin^2\theta}\right)} d\theta. \tag{39}$$

Therefore, the expression of PEP can be given as follows, where $M_\zeta(s)$ is the moment generating function (MGF) of ζ:

$$\bar{P}_e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}}) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} M_\zeta\left(\frac{-1}{4N_0 \sin^2\theta}\right) d\theta. \tag{40}$$

TABLE II
COMPLEXITY OF THE DETECTORS FOR THE QPMM, PMM, SM, AND QSM SCHEMES BASED ON THE TOTAL NUMBER OF REQUIRED REAL MULTIPLICATIONS IN DETECTION

System	Detector	Complexity
QPMM	Joint MLD	$N_P^2 M^{N_D}(2N_R + 4N_D^3 + 2N_D^2 + 2N_D)$
QPMM	C-MLD	$N_P^2 N_D(6M + 6)$
PMM	Joint MLD	$N_P M^{N_D}(2N_R + 2N_D^3 + 2N_D^2 + 2N_D)$
PMM	C-MLD	$N_P N_D(4M + 4)$
PMM	LCD	$N_P(2N_D^3 + 2N_D^2 + 2N_D) + N_D^2 + 2MN_D$
SM	ZF	$(2M + N_R + 2)N_T + 8N_T^2 N_R$
SM	Joint MLD	$8MN_T N_R$
QSM	ZF	$(4N_T M + N_R + 2)N_T + 8N_T^2 N_R$
QSM	Joint MLD	$8MN_T^2 N_R$

Note that ζ for any $\bar{P}_e(i, j, \mathbf{s} \to \bar{i}, \bar{j}, \bar{\mathbf{s}})$ fits to the Gamma distribution. Thus, the MGF of ζ is in the following form where ν is the scale and k is the shape of the Gamma distribution. Shape and scale parameters can be calculated for each pair to obtain PEP.

$$M_\zeta(s) = (1 - \nu s)^{-k}, \quad \text{for } s < \frac{1}{\nu}. \tag{41}$$

B. Detector Complexity

In this paper, it is assumed that the complexity is the number of real multiplications, where division is also considered as a multiplication and complex multiplication corresponds to four real multiplications. The complexity of the QPMM and PMM detectors are given in Table II.

For joint MLD of QPMM, both $\beta_I \mathbf{\Lambda} \mathbf{P}_I^i \mathbf{\Lambda} \mathbf{s}_I$ and $\beta_Q \mathbf{\Lambda} \mathbf{P}_Q^j \mathbf{\Lambda} \mathbf{s}_Q$ requires $2N_D^3 + N_D^2 + N_D$ multiplications, $\|\epsilon\|^2 = \sum_{k=1}^{N_R} \left|\epsilon_I^k\right|^2 + \left|\epsilon_Q^k\right|^2$ requires $2N_R$ multiplications. Also, total number of combinations to evaluate is $N_p^2 \cdot M^{N_D}$. If C-MLD is used, $\epsilon_k^{i,j}$ requires $6M + 6$ multiplications, thus $\bar{\epsilon}^{i,j}$ requires $N_D(6M + 6)$ multiplications with a total number of combinations as N_P^2.

For joint MLD of PMM, $\beta \mathbf{\Lambda} \mathbf{P}_i \mathbf{\Lambda} \mathbf{s}$ requires $2N_D^3 + 2N_D^2 + 2N_D$ multiplications, $\|\epsilon\|^2 = \sum_{k=1}^{N_R} \left|\epsilon_I^k\right|^2 + \left|\epsilon_Q^k\right|^2$ requires $2N_R$ multiplications and total number of combinations to evaluate is $N_p \cdot M^{N_D}$. If the LCD is used, both $\mathbf{P}_i \mathbf{\Lambda} \mathbf{1}$ and $\mathbf{\Lambda}^{-1} \tilde{\mathbf{y}}$ requires $N_D^3 + N_D^2$ multiplications, $abs(\mathbf{\Lambda}^{-1} \tilde{\mathbf{y}})$ requires $2N_D$ multiplications. Having N_P permutation matrices, the complexity of the first stage is $N_p \left(2N_D^3 + 2N_D^2 + 2N_D\right)$. In the second stage, $\mathbf{P}_{\hat{i}}^T \tilde{\mathbf{y}}$ requires N_D^2 multiplications and s_k requires $2M$ multiplications resulting in a total of $N_D^2 + 2MN_D$ multiplications in second stage. However, when the proposed C-MLD is used for detection, $\hat{s}_k^i$ requires $4M$, ϵ_k^i requires 4 multiplications. Thus $\bar{\epsilon}^i$ requires $N_D(4M + 4)$ multiplications, with total number of combinations as N_P.

For zero-forcing (ZF) detection of SM/QSM, we first calculate the estimate of transmitted signal by multiplying the received vector with pseudo-inverse of channel matrix H. Due to this operation, $(8N_T + 1)N_T N_R$ multiplications are made. Then, for SM, the estimated transmit vector is compared to all possible transmit vectors and it is mapped to closest transmit vector. For this mapping, $2N_T$ multiplications come from absolute values of received vector. Also, $2N_T M$ multiplications are done to compare nonzero elements of possible transmit vectors. For QSM, $4N_T^2 M$ multiplications are needed to compare nonzero elements.

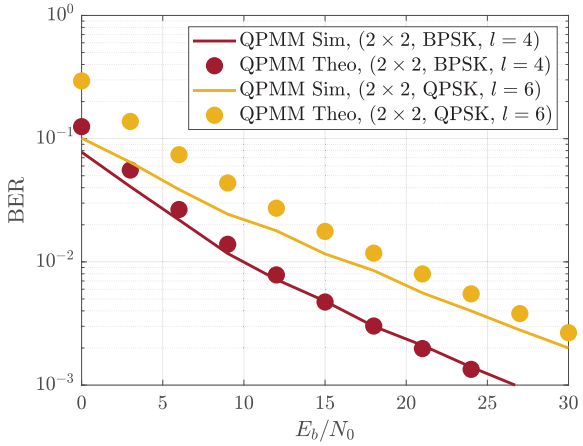


Fig. 6. The BER performance of QPMM for (2×2) setup with BPSK and QPSK modulations.

VII. SIMULATION RESULTS

A. QPMM Simulations

The BER performance of the QPMM scheme is examined and compared to that of the conventional PMM approach [4] in this section for various MIMO configurations and modulation levels. In addition, the achievable bit rate of the QPMM scheme is compared to PMM for quadrature PSK (QPSK) and the increasing number of APM symbols. The required modulation level of the PMM method to achieve the same bit rate with the QPMM scheme employing binary PSK (BPSK) for the increasing number of APM symbols is also investigated. Furthermore, the BER performance and complexity of the proposed C-MLD are compared to the optimal joint MLD and LCD. It is assumed that $P_{tot} = N_T$ for both QPMM and PMM. Also, the transmitted energy per bit is given as $E_b = 1/l$. Therefore, the signal-to-noise ratio (SNR) is defined as $E_b/N_0 = 1/lN_0$ in all simulations, where l represents the bit rate, defined as the number of bits transmitted per channel use (bpcu). It should be emphasized that the constellation of PSK symbols is rotated by $\pi/4$ to ensure that both of the IQ components are nonzero. It should be noted that the proposed C-MLD detector is used for both the QPMM and PMM schemes in Figs. 6–10. The simulation results provided in this section, and the codes for reproducing the figures are available at the online database.[1]

Fig. 6 verifies the computer simulations regarding the BER performance of QPMM with theoretical upper bounds for $N_T = N_R = N_D = 2$ with BPSK ($l = 4$) and QPSK ($l = 6$)

[1]https://github.com/burakozpoyraz/QPMM

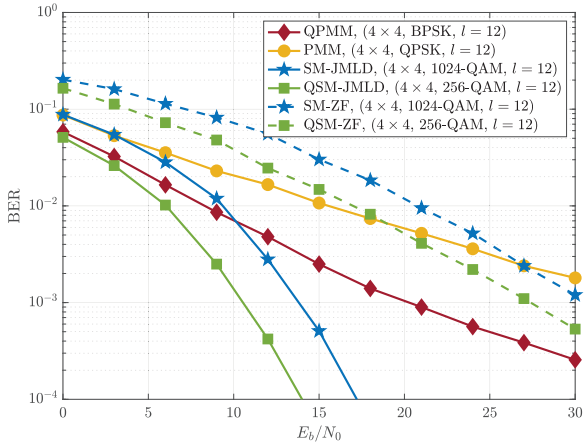


Fig. 7. Comparison of the BER performance of the QPMM, PMM, SM, and QSM schemes for $l = 12$.

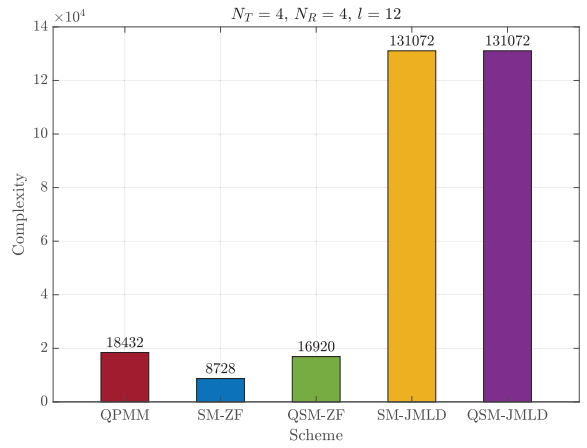


Fig. 8. Complexity comparison of the QPMM, SM, and QSM schemes for $N_T = N_R = 4$ and $l = 12$.

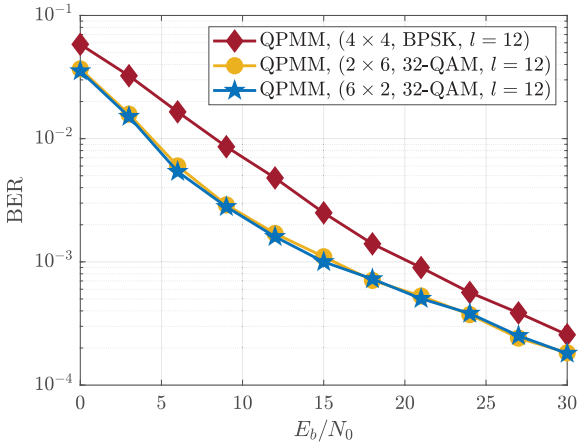


Fig. 9. The BER performance of QPMM for various MIMO setups and modulation levels, and $l = 12$.

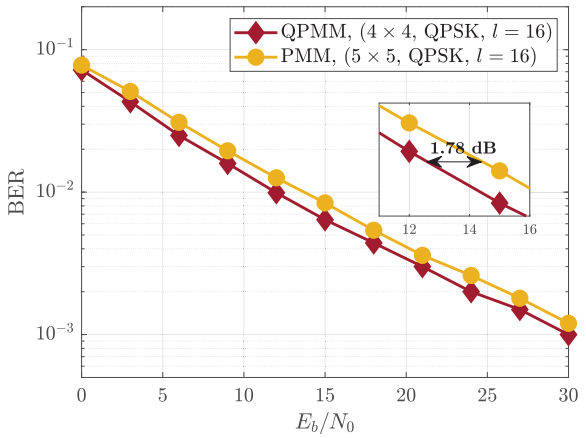


Fig. 10. Comparison of the BER performance of the QPMM and PMM schemes for different MIMO configurations, QPSK, and $l = 16$.

modulations. It can be observed that the simulation results approach the theoretical upper bound for lower bit rates.

In Fig. 7, the proposed QPMM technique is compared to the conventional PMM, SM, and QSM schemes for $N_T = N_R = N_D = 4$ and $l = 12$. Since the IQ extension of the permutation matrix indexing doubles the number of spatial bits, the PMM method has to increase the modulation level from BPSK to QPSK. As a result, the QPMM scheme outperforms its predecessor at all SNR regimes. Furthermore, the proposed QPMM scheme exhibits BER performance that is either slightly superior or comparable to that of the conventional SM and QSM schemes with the optimal joint MLD in lower SNR regimes, whereas SM and QSM yield better BER performance for higher SNR values. However, SM and QSM require significantly higher modulation levels compared to QPMM to achieve the same bit rate, which considerably increases the signal processing complexity at the transmitter. On the other hand, when SM and QSM employ the low-complexity ZF detector, their BER performance significantly deteriorates despite having detection complexities similar to the QPMM scheme with the C-MLD as demonstrated in Fig. 8.

Fig. 9 analyzes the BER performance of QPMM for various MIMO configurations and modulation levels. It can be seen that (2×6) and (6×2) MIMO setups achieve the same BER performance since the channel ranks are identical and $N_D = 2$ APM symbols are transmitted in both cases. On the other hand, increasing the channel rank and decreasing the modulation level for the same bit rate yields worse BER performance. Therefore, it is concluded that expanding the constellation instead of the channel rank to transmit more bits is better for the BER performance.

The influence of channel rank on BER performance can also be shown in Fig. 10. Employing two permutation matrices for the IQ components separately with a lower channel rank has more or less 1.78 dB gain over using a single permutation matrix with a higher channel rank. Therefore, it is advantageous to utilize QPMM over PMM under the same modulation level and bit rate.

Fig 11(a) illustrates the bit rate of QPMM and PMM for the increasing number of APM symbols and considering QPSK modulation. Since the QPMM technique doubles the number of spatial bits, the bit rate of QPMM is always higher than PMM for the identical modulation level. It can also be observed that the bit rate gap is expanding as N_D grows. Fig 11(b) shows the necessary modulated level of the PMM method for the identical bit rate with the QPMM scheme employing the BPSK modulation. As the difference between the spatial bits of QPMM and PMM enlarges as N_D increases, the required constellation size of PMM rises.

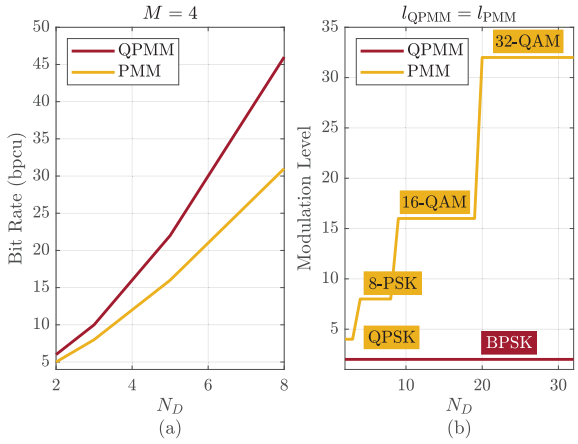


Fig. 11. Comparison of the (a) achievable bit rate of QPMM and PMM for increasing number of APM symbols, and QPSK. (b) required modulation level of PMM to attain the same bit rate with QPMM employing BPSK.

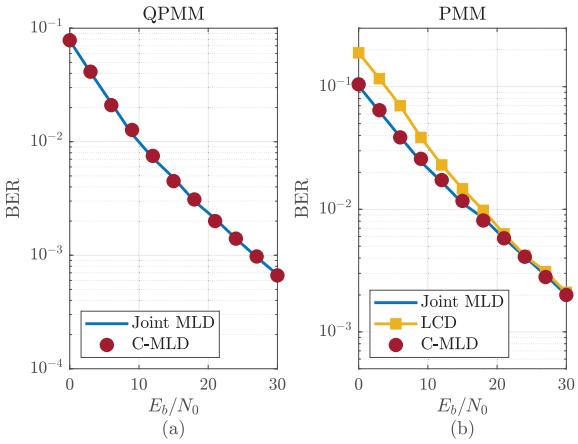


Fig. 12. Comparison of the BER performance of the (a) joint MLD and C-MLD in for the QPMM scheme. (b) joint MLD, LCD, and C-MLD for the PMM scheme.

The BER performance of the proposed C-MLD is compared to the joint MLD for the QPMM scheme with BPSK and (2×2) MIMO setup in Fig. 12(a). It can be seen that C-MLD achieves the same BER performance as the joint MLD. C-MLD is compared to the joint MLD and LCD in Fig. 12(b) for the PMM scheme with the same simulation parameters as Fig. 12(a). C-MLD attains the identical BER performance as the joint MLD, like in the QPMM scheme. Nevertheless, LCD has a worse BER performance at low SNR regimes while achieving the optimal performance at higher SNR regimes. This figure verifies that the proposed C-MLD is advantageous over LCD for the PMM scheme in BER performance.

Finally, Fig. 13 depicts detector complexities of the QPMM and PMM systems under various MIMO settings and modulation levels. The complexity values are obtained according to the formulas in Table II. Figs. 13(a), (b) and (c) show that the complexity of C-MLD is considerably lower than that of the joint MLD for all settings. In addition, increasing the number of APM symbols expands the complexity gap more than enlarging the constellation. Combining these results with the identical BER performance from Fig. 12(a), it can be concluded that the proposed C-MLD is advantageous over the joint MLD for the QPMM scheme. Likewise, C-MLD for the PMM method brings lower complexity than both the joint MLD and LCD in Figs. 13(d) and (f), while attaining almost the same complexity with LCD in Fig. 13(e). Since C-MLD reaches the same BER performance level as the joint MLD, C-MLD is the best detector also for the PMM scheme.

TABLE III
THE TRAINING PARAMETERS OF FUSIONDET

Parameter	Value
Learning rate	0.0005
Batch size	128
Number of Epochs	5000
Learning Rate Scheduler	Adam
Number of Epochs	Exponential LR w/ $\gamma = 0.9995$
Loss function	Cross Entropy Loss
Evaluation Metric	Bit error rate of indices

B. FusionDet Simulations

In this study, a (4×4) MIMO configuration is employed to examine the application of FusionDet. This setup involves a total of 16 potential values for both the i and j indices. Fig. 3 provides the dimensions for this specific configuration, employing BPSK modulation. To facilitate the experimentation, a large dataset is generated for training purposes, utilizing MATLAB simulations. Each dataset is tailored to different SNR values, comprising $1M$ input-output pairs. From these, $50K$ pairs are randomly chosen to construct the test dataset, while an additional $95K$ pairs are selected for validation. The remaining $855K$ pairs are allocated for the training of the FusionDet model.

The actual training of FusionDet takes place using the PyTorch library in Python, with the training parameters detailed in Table III. During the training process, the model's performance is assessed by evaluating losses for both indices based on a specified loss function. The cumulative loss for both indices is calculated and employed to update the model. Hyperparameter selection involves an extensive exploration of various values to determine the best-performing parameter set, validated against the performance on the validation dataset.

Throughout the training iterations, performance on the validation dataset is continuously monitored at each epoch. The model parameters demonstrating the highest validation performance are ultimately chosen to facilitate a comprehensive evaluation of the model's performance. This approach ensures a robust and comparative assessment of the model's effectiveness.

The FusionDet system's evaluation involves assessing its performance in various coverage scenarios using an independent dataset comprising $500K$ samples generated via MATLAB. Coverage, within this context, signifies the proportion of samples that FusionDet successfully identifies. As P_{th} increases, more samples are dismissed by the FusionDet system, subsequently reducing the coverage. Conversely, a decrease in P_{th} leads to fewer rejections of samples by FusionDet, thereby augmenting the coverage. The impact of

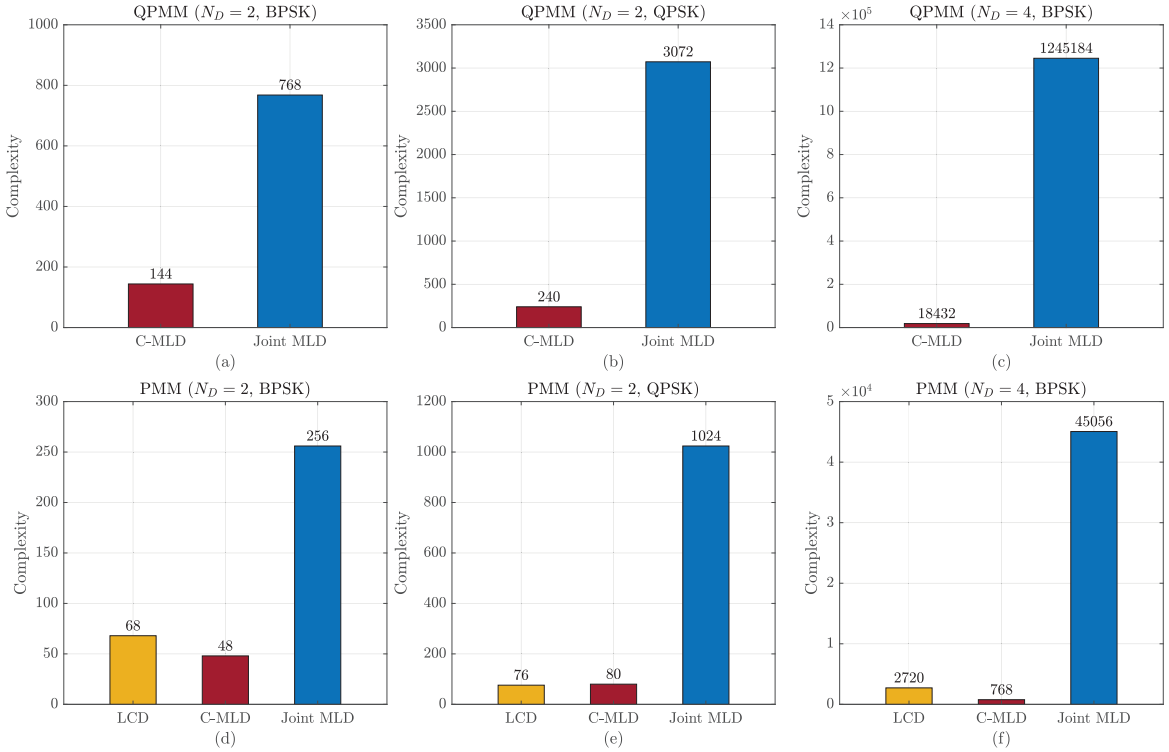


Fig. 13. Complexity comparison of the joint MLD and C-MLD for QPMM with (a) $N_D = 2$ and BPSK. (b) $N_D = 2$ and QPSK. (c) $N_D = 4$ and BPSK, and complexity comparison of the joint MLD, LCD, and C-MLD for PMM with (d) $N_D = 2$ and BPSK. (e) $N_D = 2$ and QPSK. (f) $N_D = 4$ and BPSK.

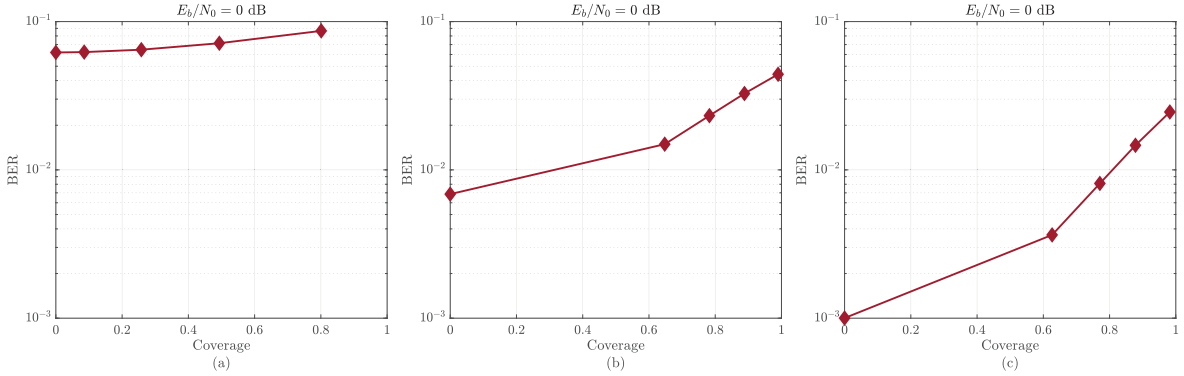


Fig. 14. The BER performance of QPMM for different FusionDet coverage scenarios considering (a) 0 dB. (b) 10 dB. (c) 20 dB.

coverage on BER for distinct SNR settings (0 dB, 10 dB, and 20 dB) is exhibited in Figs. 14(a), 14(b), and 14(c). Varying P_{th} values (0.50, 0.625, 0.75, 0.875, and 1) are considered within these simulations.

To elucidate the trade-off in BER performance, different coverage scenarios are compared across 0 dB, 10 dB, and 20 dB SNR, as depicted in Fig. 15. As anticipated, lowering the P_{th} value enhances the coverage of FusionDet, reducing its computational complexity but causing an increase in BER.

Additionally, a complexity comparison, outlined in Table IV, illustrates the computational load. Notably, the complexity is quantified as the average number of operations for a single detection. Specifically, the FusionDet system necessitates 2688 multiplications for identifying i and j indices. Moreover, $6 \times M$ multiplications are required for a single symbol detection. Hence, the overall detection complexity for this particular setup is computed as $2688 + 4 \times 6 \times 2 = 2736$. Therefore, if the coverage

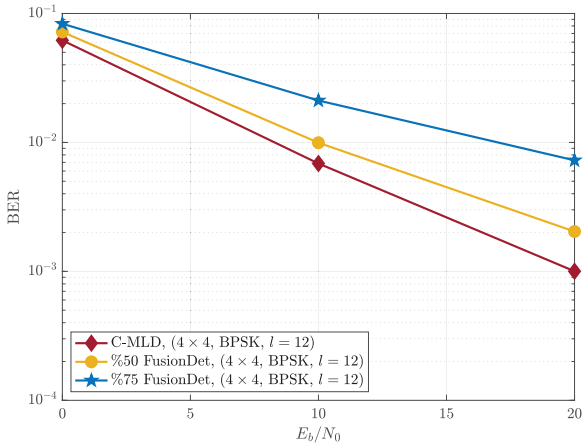


Fig. 15. The BER performance of QPMM considering different FusionDet coverage scenarios.

is %75, average of multiplications is calculated as given in Eq. 42. In Eq. 42, 2688 multiplications are required for

TABLE IV
COMPLEXITY EVALUATION FOR DIFFERENT COVERAGES

FD Coverage	Complexity
%0	18432
%50	11928
%75	7332

every sample to decide if the FD indices will be used. If FD indices are used, 48 multiplications are needed for symbol detection. Otherwise, C-MLD algorithm is utilized and 18432 multiplications are needed. Generalized complexity equation depending on the coverage is also given in Eq. 42.

$$\begin{aligned}\text{Complexity} &= \frac{100 \times 2688 + 75 \times 48 + 25 \times 18432}{100}, \\ &= 2688 + c \times 48 + (1-c) \times 18432, \\ &= 7332, \end{aligned} \tag{42}$$

where c is the FD coverage.

VIII. CONCLUSION AND FUTURE DIRECTIONS

QPMM, a novel IM-aided MIMO scheme, and a low complexity detector, C-MLD, is introduced in this paper. Similar to the PMM method [4], QPMM employs IM in the spatial domain, by modulating the spatial bits with indexing two permutation matrices for the IQ components of the transmit symbol vector. It is designed to both improve BER performance and spectral efficiency. It has been shown that the QPMM scheme achieves notably better BER performance than PMM [4] scheme. Also, PMM scheme is observed to require higher modulation levels for the same bit rate with QPMM. In addition to the performance improvements, it has been shown that the C-MLD receiver achieves optimal BER performance with much lower complexity compared to MLD. Moreover, a DL-based detector scheme, named FusionDet, is proposed. The computer simulations reveal that FusionDet provides a trade-off between detection complexity and BER performance. As an extension of this research, the proposed QPMM scheme can be applied in an MU system. Furthermore, the proposed DL-based FusionDet detector could be enhanced to enable joint detection of indices and symbols, while maintaining manageable detection complexity.

REFERENCES

[1] E. Basar, M. Wen, R. Mesleh, M. Di Renzo, Y. Xiao, and H. Haas, "Index modulation techniques for next-generation wireless networks," *IEEE Access*, vol. 5, pp. 16693–16746, 2017.

[2] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 168–175, Jul. 2016.

[3] T. Mao, Q. Wang, Z. Wang, and S. Chen, "Novel index modulation techniques: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 315–348, 1st Quart., 2019.

[4] R. F. Siregar, N. Rajatheva, and M. Latva-Aho, "Permutation matrix modulation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4947–4961, Jul. 2023.

[5] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[6] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.

[7] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. 44th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2010, pp. 1498–1502.

[8] L.-L. Yang, "Transmitter preprocessing aided spatial modulation for multiple-input multiple-output systems," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5.

[9] R. Zhang, L.-L. Yang, and L. Hanzo, "Generalised pre-coding aided spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5434–5443, Nov. 2013.

[10] R. Mesleh, S. S. Ikki, and H. M. Aggoune, "Quadrature spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2738–2742, Jun. 2015.

[11] C.-C. Cheng, H. Sari, S. Sezginer, and Y. T. Su, "Enhanced spatial modulation with multiple signal constellations," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2237–2248, Jun. 2015.

[12] J. Li, M. Wen, X. Cheng, Y. Yan, S. Song, and M. H. Lee, "Generalized precoding-aided quadrature spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1881–1886, Feb. 2017.

[13] C. Masouros and L. Hanzo, "Dual-layered MIMO transmission for increased bandwidth efficiency," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3139–3149, May 2016.

[14] C. Liu, L.-L. Yang, W. Wang, and F. Wang, "Joint transmitter–receiver spatial modulation," *IEEE Access*, vol. 6, pp. 6411–6423, 2018.

[15] N. S. Perovic, M. Di Renzo, and M. F. Flanagan, "Transmit-receive generalized spatial modulation based on dual-layered MIMO transmission," in *Proc. IEEE 30th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2019, pp. 1–7.

[16] Q. Si and M. Jin, "A grouped pre-coding aided spatial modulation for MIMO systems," *IEEE Access*, vol. 8, pp. 44643–44651, 2020.

[17] M.-T. O. El Astal, M. S. Muhareb, and A. M. Abu-Hudrouss, "Generalized space-time block coded quadrature spatial modulation," in *Proc. Int. Conf. Promising Electron. Technol. (ICPET)*, Dec. 2020, pp. 50–55.

[18] S. Gadhai and R. Budhiraja, "Joint-mapping-based variable active antenna spatial modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1668–1672, Oct. 2020.

[19] G. Kaddoum, M. F. A. Ahmed, and Y. Nijsure, "Code index modulation: A high data rate and energy efficient communication system," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 175–178, Feb. 2015.

[20] F. Cogen, E. Aydin, N. Kabaoglu, E. Basar, and H. Ilhan, "Generalized code index modulation and spatial modulation for high rate and energy-efficient MIMO systems on Rayleigh block-fading channel," *IEEE Syst. J.*, vol. 15, no. 1, pp. 538–545, Mar. 2021.

[21] C. Zhang, H. Lu, and J. Liu, "IMRecoNet: Learn to detect in index modulation aided MIMO systems with complex valued neural networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 2, pp. 1791–1805, Feb. 2023.

[22] L. Xiang, Y. Liu, T. Van Luong, R. G. Maunder, L.-L. Yang, and L. Hanzo, "Deep-Learning-Aided joint channel estimation and data detection for spatial modulation," *IEEE Access*, vol. 8, pp. 191910–191919, 2020.

[23] T. Wang, F. Yang, J. Song, and Z. Han, "Deep convolutional neural network-based detector for index modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1705–1709, Oct. 2020.

[24] J. Kim, H. Ro, and H. Park, "Deep learning-based detector for dual mode OFDM with index modulation," *IEEE Wireless Commun. Lett.*, vol. 10, no. 7, pp. 1562–1566, Jul. 2021.

[25] R. Jiang et al., "Deep learning-aided signal detection for two-stage index modulated universal filtered multi-carrier systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 1, pp. 136–154, Mar. 2022.

[26] S. Katla, L. Xiang, Y. Zhang, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Deep learning assisted detection for index modulation aided mmWave systems," *IEEE Access*, vol. 8, pp. 202738–202754, 2020.

[27] H. Esmaiel, H. A. Leftah, N. U. R. Junejo, and H. Sun, "Deep learning-based index modulation for underground communications," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 2122–2132, 2023.