EURECOM
Sophia Antipolis

Communication Systems Department

Novel Joint Estimation and Detection Metrics for Short-Data Transmission Systems

Authors :

Mody Sy & Raymond Knopp

Email:mody.sy@eurecom.fr

May 20, 2023

Contents

1 Introduction

It is expected that the 6G air-interface will build upon the 5G standard and address new pardigms for feedback-based cyber-physical systems combining communications and sensing. In particular, there will be a need for tight control loops using the air-interface to control 6G-enabled objects with high-reliability, perhaps even requiring lower latencies than those achieved by current 5G technology, for example sub-1ms uplink application-layer latency in microwave spectrum. Although 5G transmission formats can provide very short-packet transmission through the use of mini-slots, the ratio of tranining information to data is not necessarily adapted to extremely short data transmission. Moreover, the transmission formats are designed with conventional quasi-coherent receivers which can be quite sub-optimal in such scenarios where accurate channel estimation is impossible because of sporadic transmission of short packets. One such instance is because of stringent decoding latency constraints such as those emerging in so-called *Ultra-Reliable-Low-Latency Communication* (URLLC) industrial IoT applications. This would be similar for evolved channel state information (CSI) feedback control channels or future combined-sensing and communication paradigms requiring rapid sensory feedback to the network. The area of short block transmission has garnered significant attention in recent years, with extensive research conducted on various aspects, including the design of signal codes and the establishment of state-of-the-art converse and achievability bounds for both coherent and non-coherent communications [1–6].

In this work we investigate bit-interleaved coded modulation (BICM) and detection strategies for packets in the range of 20-100 bits for these envisaged beyong 5G/6G signaling scenarios.
Zehavi [7] proposed bit-interleaved coded modulation (BICM) as a pragmatic approach to coded modulation. Its basic principle is the ability of an interleaving permutation to separate an underlying binary code from an arbitrary higher-order modulation [8]. Per-bit log-likelihood ratios are used to convey soft metrics from the demodulator to the decoder in order to reduce information loss. Thus, this fundamental observation spurred interest in BICM. In [9], Caire et al., later, conducted a thorough analysis of BICM in terms of information rate and error probability including both coherent and non-coherent detection. Furthermore, BICM is seen as a standard coding approach for wireless communications channels and is the workhorse of today's high spectral efficiency systems as well as low spectral efficiency orthogonal modulation systems, including satellite communication, and broadband internet. Interestingly, one of the key benefits of BICM is that it allows the data rate to be increased without significantly increasing the error rate. This is because the error-correcting code helps to compensate for errors that occur during transmission, allowing the data to be transmitted at a higher rate without sacrificing reliability. It is particularly useful in situations where the communication channel is prone to errors, and a high level of reliability is required. 3GPP systems have made use of BICM since the 3G-era. Therefore, in order to improve their efficiency and enable high performance communication, schemes such as rate matching, scrambling and other processes inherent to modern wireless communication standards are de facto added to the reference BICM schematic. In addition, the underlying detection and decoding metrics must provide enhanced performance and low complexity trade-off in *Ultra-Reliable-Low-Latency Communication* (URLLC)

perpectives. One of the main applications is in the field of mission-critical communications, such as those used by emergency services or in industrial control systems. In these cases, it is essential that the communication is both ultra reliable (i.e. with a very low probability of failure) and has extremely low latencies (i.e. the time it takes for a message to be transmitted and received). Overall, in addition to mission-critical applications, URLLC is also expected to have a number of other important use cases, and will enable a wide range of new applications that require extremely high levels of reliability and low latencies. In this paper, we look into bit-interleaved coded modulation metrics exploiting joint detection and estimation which are amenable to situations where low-density DMRS are interleaved with coded data symbols. We are primarily interested in situations where accurate channel estimation is impossible. One such instance is because of stringent decoding latency constraints such as those emerging in URLLC systems. Another similar instance would be in the case extremely short and sporadic packets either for control channels or certain machine-type communication paradigms. We show that by using a properly conceived metric exploiting interleaved DMRS in the decoding metric computation, we can achieve performance approaching a receiver with perfect channel estimation and significant coding gains compared to a conventional 5G OFDM receiver. The scheme performs detection over contiguous groups of modulated symbols including those from the DMRS to provide soft metrics for the bits in each group to the channel decoder. We evaluate performance using a full 5G transceiver chain for both polar and LDPC coded formats. The schemes are applicable to both uplink and downlink transmission where packets are encoded into a small number of OFDM symbols with interleaved DMRS. Additionally, we investigate the impact of varying densities of reference signals on performance. The article is structured as follows. Section II lays out the system model, Problem reformulation and foundations of NR polar and LDPC-coded modulations, Section III focuses on the proposed BICM metrics, Section IV presents the results and performance analysis, and finally Section V concludes the paper.

Notation : Scalars are denoted by italic letters, vectors and matrices are denoted by bold-face lower-case and upper-case letters, respectively. For a complex-valued vector $\boldsymbol{x}$, $||\boldsymbol{x}||$ denotes its Euclidean norm, $|\cdot|$ denotes the absolute value. $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. Re $\cdot$ denotes the real part of a complex number. $\boldsymbol{I}_0(\cdot)$ the zero-th order Modified Bessel function of the first kind. $diag(\boldsymbol{A})$ denotes a diagonal matrix with each diagonal element being the corresponding element in $\boldsymbol{A}$. $\boldsymbol{I}$ is an identity matrix with appropriate dimensions. Galois field is denoted by $GF(2)$. m *s.t.* $e_j = b$ means the set of the constellation symbols *such that* bit $j-th$ bit e is equal to b. $\mathcal{L}(\cdot)$ denotes log likelihood function / ratio. The superscripts ^{T}, * and ^{H} denote the transpose, conjugate, and Hermitian (complex conjugate transpose).

2 System Model and Problem Formulation

2.1 System Model

Consider a discrete-time model in which the transmitted and received symbols are N-dimensional column vectors, and thus a system is designed in such a way

that the relationship between the transmitted and received signals is as follows:

$$\boldsymbol{y}_r = \text{diag}(\boldsymbol{h}_r)\boldsymbol{x_m} + \boldsymbol{z}_r, r = 0, 1, \cdots, N_R - 1, \tag{1}$$

where $\boldsymbol{y}_r$ represents an observed vector in N complex dimensions, $\boldsymbol{x}_m$ is and N-dimensional modulated vector transporting B channel bits, so that the message $m = 0, 1 \ldots, 2^B - 1$, z is additive white Gaussian noise whose real and imaginary components are independent and have variance σ^2 in each dimension. Various models for h will be used in this study and will be described along with the corresponding receiver structures. N_R represents the number of observations of the transmitted vector over a multi-antenna receiver.

The transmitted vector $\boldsymbol{x}_m$ is often composed of data independent components which are known to the receiver. These are so-called *pilot* or *demodulation reference signals (DMRS)* which are conceived in order to allow for resolving channel ambiguity in time, frequency and space. In practice, the reference signals are used for estimating the vector channels $\{\boldsymbol{h}_r\}$ and are commonly interleaved among the data-dependent components according the characteristics of the propagation channel. It is notably the case in current OFDM systems. In earlier CDMA systems, DMRS were sometimes superimposed on top of data-dependent signals. We denote the number of data dimensions by N_d and reference signal dimensions by N_p where $N_d + N_p = N$. In 3GPP standard, N is typically equal to $12KL$. This represents the number of complex dimensions or resource elements in the physical resource blocks. The number of physical resource blocks, K ranges from 1 to 16, while the number of symbols, L, ranges from 1 to 14, and can be increased if multiple slots are used for signaling the B bits. The assumption in this work is that the data-dependent components of $\boldsymbol{x}_m$ are generated from a binary code whose output is interleaved and mapped to an M-ary modulation symbol alphabet. We will assume that the binary code generates E bits and the interleaver mapping is one-to-one so that E bits are also fed to the modulator. The binary-code and interleaver combination can thus be seen as a (E, B) binary block code. Denote the E coded bits as $e_k, k = 0, 1, \cdots, E - 1$. Adjacent $\log_2 M$ bit-tuples are used to select the N_E modulated symbols in the symbol alphabet. Typically, we will assume that a Gray mapping is used in the case of non-binary modulation.

2.2 Problem Formulation

We denote the likelihood function for the observed vector on a particular receiver branch with respect to a given transmitted signal as

$$q\left(\boldsymbol{x}(m), \{\boldsymbol{y}\}\right) = P\left(\{\boldsymbol{y}\} \mid \boldsymbol{x}(m)\right), \tag{2}$$

and the resulting maximum-likelihood receiver for detecting m is

$$\hat{m} = \underset{m=0,1,\cdots,2^B-1}{\text{argmax}} q\left(\boldsymbol{x}(m), \{\boldsymbol{y}\}\right) \tag{3}$$

The likelihood of coded bit $e_j \in \{0, 1\}$ is

$$q_{j,b}\left(\{\boldsymbol{y}\}\right) \sum_{m \ s.t. \ e_j = b} q\left(\boldsymbol{x}(m), \{\boldsymbol{y}\}\right). \tag{4}$$

e_j is the $j - th$ bit of symbol x, $m \ s.t. \ e_j = b$ is the set of the constellation symbols such that bit $b = \{0, 1\}$. As is common in the case of BICM-based systems, the soft input to the binary channel decoder is given as the log-likelihood

ratio (LLR) for coded bit j.

$$\mathcal{L}_j\left(\{\boldsymbol{y}\}\right)\log\left[\frac{q_{j,0}\left(\{\boldsymbol{y}\}\right)}{q_{j,1}\left(\{\boldsymbol{y}\}\right)}\right] \tag{5}$$

Note that in the above expressions we do not limit the dimensionality of the observations when computing likelihoods of particular bits. In the original work of Caire *et al* [9] the authors assume an ideal interleaving model which allows limiting the observation interval of a particular coded bit to the symbol in which it is conveyed. For long blocks this assumption is realistic for arbitrary modulation signal sets and is sufficient for BPSK and QPSK irrespective of the block length when the channel is known perfectly. Nevertheless, practical systems usually apply single symbol likelihood functions for short blocks and high-order modulations. For the primary case of interest here, namely transmission without channel state information, single symbol detection is impossible. At the very least, the observation of one reference symbol must be used to generate likelihoods of the coded bits of a data symbol, thus warranting the study of block detection.

2.3 General Framework

2.3.1 Bit-Interleaved Polar-coded Modulation (BIPCM)

Bit Interleaved Polar Coded Modulation is referred to as BIPCM. In this instance we are dealing with the CRC-Aided Polar (CA-Polar) coding scheme, one of the basic code construction techniques established by the 3GPP Standard [10]. Using polar codes as a channel coding scheme for 5G control channels has demonstrated the significance of Arikan's invention [11], and its applicability in commercial systems has been proven. This new coding family achieves capacity rather than merely approaching it as it is based on the idea of channel polarization. Polar codes can be used for any code rate and for any code length shorter than the maximum code length due to their adaptability.

They are the first kind of forward error correction codes that achieve symmetric capacity for any binary-input discrete memoryless channel under low-complexity encoding and low-complexity successive cancelation (SC) decoding with order of $O(N\log N)$ for infinite length codes. In NR, the polar code is used to encode broadcast channel as well as DCI and uplink control information (UCI). 3GPP NR uses a variant of the polar code called distributed CRC (D-CRC) polar code, that is, a combination of *CRC-assisted and PC polar codes*, which interleaves a CRC- concatenated block and relocates some of the PC bits into the middle positions of this block prior to performing the conventional polar encoding [12]. This allows a decoder to early terminate the decoding process as soon as any parity check is not successful.

Assume that the input message(*UL/DL Control Information*) before CRC attachment is $a(0), a(1), \ldots, a(A-1)$, where A is input sequence, parity bits are
$p(0), p(1), \ldots, p(L-1)$, L is the number of parity bits. For the downlink, a CRC of length $L = 24$ bits is used, and for the uplink, depending on the quantity of A, CRCs of length $L = 6$ and $L = 11$ bits are used. The message bits after attaching CRC are $b(0), b(1), \ldots, b(B-1)$, where B is the size of Control information with CRC bits: $B = A + L$.
The input bit sequence to the code block segmentation is denoted

$a(0), a(1), \ldots, a(A-1)$,where value of A is no larger than 1706. Assume that the maximum code block size is A' and C the number of existing code blocks, the sequence
$c_r(0), c_r(1), \ldots, c_r(A'/C - 1)$ is used to calculate the CRC parity bits
$p_r(0), p_r(1), \ldots, p_r(L-1)$.The sequence of bits resulting after attaching a CRC to the r-th code block is denoted by $c_r(0), c_(1), \ldots, c_r(K_r - 1)$, where K_r is the number of bits in the $r - th$ Code block to be fed to the channel encoder. Then, the coded bit are denoted by $d(0), d(1), \ldots, d(N_r - 1)$ where $N_r = 2^n$. Denote by E_r the rate matching output sequence length of the $r - th$ coded-block: **if** $E_r \leq (9/8) \cdot 2^{(\lceil \log_2 E_r \rceil - 1)}$ and $K/E_r < 9/16$ $n_1 = \lceil \log_2 E_r \rceil - 1$ **else** $n_1 = \lceil \log_2 E_r \rceil$. And then, $R_{\min} = 1/8$; $n_2 = \lceil \log_2 (K/R_{\min}) \rceil$; $n = \max \{ \min \{ n_1, n_2, n_{\max} \}, n_{\min} \}$
where $n_{\min}$ and $n_{\max}$ provide a lower and an upper bound on the code length, respectively. In particular, and $n_{\min} = 5$ and $n_{\max} = 9$ for the downlink control channel, whereas $n_{\max} = 10$ for the uplink control channel. The polar encoding is based on the following procedure [10]:

- The sequence $c(0), c(1), \ldots, c(K_r - 1)$ is interleaved into bit sequence $c'(0), c'(1), \ldots, c'(K_r - 1)$ via a definite interleaving pattern [10].

- The interleaved vector $\boldsymbol{c'}$ is assigned to the information set along with the PC bits, while the remaining bits in the N-bit vector $\boldsymbol{u}$ are frozen. Hence, $\boldsymbol{u} = u(0), u(1), \ldots, u(N_r - 1)$ is generated according to the clause 5.3.1.2 [10].

- Denote $\boldsymbol{G}_{N_r} = (\boldsymbol{G}_2)^{\otimes n}$ as the $n-th$ Kronecker power of matrix $\boldsymbol{G}_2$, where $\boldsymbol{G}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, the output after encoding $\boldsymbol{d} = d(0), d(1), \ldots, d(N_r - 1)$ is obtained by $\boldsymbol{d} = \boldsymbol{u} \boldsymbol{G}_{N_r}$, where encoding is performed in $GF(2)$.

As a result, a rate matching process is perfomred per coded block and consists of *sub-block interleaving*, bit collection, and bit interleaving. The input bit. Thus, the output bit sequence after rate matching are the sequences f_{rk} , for $r = 0, \ldots, C - 1$ and $k = 0, \ldots, E_r - 1$, where E_r is the number of rate matched bits for the $r - th$ code block. At the rate matching stage, perforation, shortening, or repetition $(E_r \geq N_r)$ might be applied. The aim is to convert a vector of N_r bits into a vector of E_r bits. Afterwards, we perform code block concatenation to recast all code block messages into a sequence of transport block messages. The output bit sequence after the code block concatenation is $g(0), g(1), \ldots, g(E-1)$. The receiving chain is therefore the transmitting chain in the reverse flow. Furthermore, currently, the main polar code decoding algorithms are the SC algorithm [11] , the SCL algorithm [13, 14] , the CA-SCL algorithm [15] , the BP algorithm [16] , the SCAN algorithm [17] , and various simplified versions of the algorithm can be found in the literature [13, 15–23]. The SC algorithm was originally proposed by Arikan, but its performance is poor in the case of finite length codes. SCL is an improved version of the SC algorithm with better performance, and the principle is to provide multiple paths over SC, while CA-SCL is a cyclic redundancy check on the message bits over SCL [24]. It is an improvement on the SCL decoder where a high-rate CRC code is added to the polar code to help choose the right codeword from the final list of pathways. Every time a SCL-decoder fails, it has been seen that the

right codeword is included in the list. In terms of performance, these algorithms are ranked as follows:*CA-SCL> state of the art SCL>BP=SCAN>SC.* In this instance, the channel decoder technique must utilize the CRC-Aided Successive Cancellation List decoding for downlink (DCI or BCH) or uplink (UCI) messages going forward to benefit from the performance improvements it provides. Additionally, the well acknowledged potential of CA-SCL decoding to perform better than Turbo or LDPC codes was one of the primary reasons the 3GPP adopted polar codes.

The overall representation of the BIPCM/BILCM process, from the MAC to the physical layer processing, is presented in Figure 1. This figure depicts the transmit-end procedure for uplink channels, which encompasses the addition of a transport block CRC, code block segmentation with additional CRC attachment, channel encoding, rate matching, code block concatenation, and modulation. It is important to note that the receiving chain is simply the reverse flow of the transmitting chain.

Figure 1: Bit-Interleaved Polar/LDPC coded Modulation (BIPCM/BILPCM) : Transmitter end

2.3.2 Bit-Interleaved LDPC-coded Modulation (BILCM)

Bit-Interleaved LDPC-Coded Modulation is referred to as BILCM. Originally proposed by Gallager in early 1960s [25], low-density parity-check (LDPC) coding is currently adopted in 5G NR for both uplink and downlink shared transport channels. LDPC codes are appropriate for 5G NR shared channels due to its high throughput, low latency, efficient decoding complexity and rate compatibility. The performance of LDPC codes in 5G NR demonstrates an error floor at or below the 10^5 block error rate (BLER), which is a significant advantage over traditional coding techniques. The QC-LDPC family serves as the foundation for 5G NR LDPC codes. NR LDPC code is constructed from a Base Graph Matrix (BG) of dimension $M \times N$, designated as $\boldsymbol{H}_{BG}$. The selection of the $\boldsymbol{H}_{BG}$ matrices in the 5G NR coding process is based on the coding rate and the length of the transport block or code block. There are two base graphs, $BG1$ with dimensions of $N = 68$ and $M = 46$ optimized for large information block sizes of $K \leq 8448$ and high coding rates between $1/3 \leq R \leq 8/9$, and $BG2$ with dimensions of $N = 52$ and $M = 42$ optimized for small information block sizes of $K \leq 3840$ and lower coding rates between $1/5 \leq R \leq 2/3$. These codes are suitable for high reliability scenarios due to their ability to achieve additional coding gain at low-code rates. The maximum number of information bits for $BG1$ is $K = 22Z_c$ and for $BG2$ is $K = 20Z_c$, where Z_c is the lifting size. There are 51 lifting sizes ranging from 2 to 384 for each base graph. The parity check matrix, denoted as **H**, is obtained by replacing each element of H_{BG} with a cyclic permutation identity matrix, $I(P_{ij})$. In other words, each element of H_{BG} is replaced by the corresponding Cyclic Permutation Matrix (CPM). The size of the matrix **H** is $m \times n$, with $m = M \times Z_c$, $n = N \times Z_c$, and $k = n - m = (N - M) \times Z_c$. Both $BG1$ and $BG2$ have similar structures. There are various effective LDPC encoding techniques because of the structure and

features of base graphs. As in [26], a brand-new effective encoding technique that acts as a high throughput, low complexity coding architecture has been proposed.

The transmission process involves first attaching CRC to the transport block. CRC is an error detection code used to measure BLER after decoding. Hence, the entire transport block is used to calculate CRC parity bits. Assume that the transport message before CRC attachment is $a(0), a(1), \ldots, a(A-1)$, where A is the size of the transport block message. Parity bits are $p(0), p(1), \ldots, p(L-1)$, where L is the number of parity bits. if $A > 3824$, L is 24, otherwise L is 16 is used. he message bits after attaching CRC are $b(1), b(2), \ldots, b(B)$, B is the size of transport block information with CRC bits and $B = A + L$. LDPC base graph is selected based on the size of transport block message A and transport block coding rate R. If $A \leq 292$,or if $A \leq 3824$ and $R \leq 0.67$, or if $R \leq 0.25$, LDPC BG_2 is used. Otherwise, LDPC BG_1 is used [10]. The output of code block segmentation + CRC attachment is $c_r(1), c_r(2), \ldots, c_r(K_r)$, assuming $K_r = K'_r + L$, where K' is the number of bits in $r - th$ code block , L is attached CRC on the $r - th$ Code block. Thus, each code block message is encoded independently. In 3GPP NR, the input bit sequence is represented as $\boldsymbol{c} = [c(0), c(1), \ldots, c(K_r - 1)]^T$, where K_r is the number of information bits in a code block, the the redundant bits are called parity bits denoted by $\boldsymbol{w} = [w(0), w(1), \ldots, w(N + 2Z_c - K_r + 1)]^T$ .The LDPC-coded bits are denoted by $d(0), d(1), \ldots, d(N_r - 1)$. A code block is encoded by the LDPC encoder based on the following procedure [10]:

1. Find the set with index iLS which contains Z_c in [10].

2. Set $d_{k-2Z_c} = c_k, \forall k = 2Z_c, \ldots, K_r - 1$

3. Generate $N_r + 2Z_c - K_r$ parity bits $\boldsymbol{w} = [w(0), w(1), \ldots, w(N + 2Z_c - K_r + 1)]^T$ such that $\boldsymbol{H} \times [\boldsymbol{c}\ \boldsymbol{w}]^T = \boldsymbol{0}$

4. The encoding is performed in $GF(2)$.

5. Set $d_{k-2Z_c} = w_{k-K_r}, \forall k = K_r, \ldots, N_r + 2Z_c - 1$

Thus, rate matching and code block concatenation processes proceed. The receiving chain serves as the inverse counterpart of the transmitting chain. The decoding of Low-Density Parity-Check (LDPC) codes is carried out on each code block individually, and a range of decoding techniques can be implemented. Among these, belief propagation (BP) methods, which rely on iterative exchange of messages between bit nodes and check nodes, are the most commonly utilized for LDPC decoding. Although the BP method presents a considerable computational complexity, it offers near-optimal decoding performance [27]. To achieve a better trade-off between performance and complexity, several effective decoding algorithms have been proposed in scientific literature. Among these algorithms, layered message passing decoding appears to be a promising approach for ultra-reliable low-latency communication (URLLC) due to its ability to speed up convergence time. Therefore, it is deemed suitable for investigation in the current study. The subsequent principle may be used to illustrate the layered decoding principle as described in the work of WANG [28] :

- Each layer independently processes variable node operations and checks node operations.

- Current layer's input LLR is the prior layer's output LLR.

- The output LLR of the decoding algorithm, which is the output LLR of the last layer, will be used to make the decision.

- The input LLR of the current layer can be updated according to the equation $\mathcal{L}_{k+1,i} = \mathcal{L}_{k,i} - \mathcal{L}_{k+1,i'}$ where $\mathcal{L}_{k+1,i}$ is the updated input LLR of layer $k+1$, $\mathcal{L}_{k,i}$ is the output LLR of previous layer $\mathcal{L}_{k+1,i'}$ and is the old input LLR of layer $k+1$.

2.3.3 Modulation and Resource Mapping

In both scenarios, the encoded payload undergoes rate-matching and block concatenation prior to being fed to a QPSK modulator. This process yields a set of complex-valued modulation symbols, represented as $x(0), x(1), \ldots, x\left(N_d/2 - 1\right)$. Subsequently, the resource allocation process is executed, wherein one or multiple OFDM symbols are utilized to allocate the modulated symbols to resource blocks and insert the DMRS resources. The insertion of DMRS resources serves the purpose of providing non-coherent detection and channel estimation. The number of resource blocks is governed by the payload size and coding settings. When the payload size is small, fewer resource blocks are required, thus maintaining a constant effective coding rate. As illustrated in Figure 2, the resource mapping here is embedded in the same spirit as in a 3GPP PUCCH2 transmission.
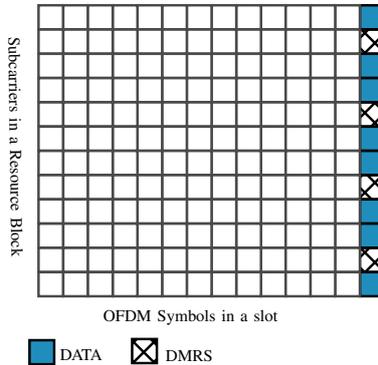
Figure 2: General resource mapping: 1 OFDM symbol

3   BICM Receivers

3.1   Perfect Channel State Information

In the instance of perfect channel state information, the set of observed random vectors needs to be augmented by the set of channel vectors $\{\boldsymbol{h}_r\}$ in equations (2-5). Thus, the perfect LOS channel is defined as $\boldsymbol{h}_r = A_r e^{j\theta_r}\boldsymbol{I}$ where θ_r and $A_r > 0$ denote a known phase and gain respectively. The likelihood function is

$$q\left(\boldsymbol{x_m}, \{\boldsymbol{y_r}, \boldsymbol{h_r}\}\right) = P\left(\{\boldsymbol{y_r}, \boldsymbol{h_r}\} \mid \boldsymbol{x_m}\right) = \\ P\left(\{\boldsymbol{y_r}\} \mid \boldsymbol{x_m}, \boldsymbol{h_r}\right) P\left(\{\boldsymbol{h_r}\} \mid \boldsymbol{x_m}\right) \tag{6}$$

If the transmitted signal $\boldsymbol{x}_m$ is independent of the channel realization $\{\boldsymbol{h}_r\}$, the term $P\left(\{\boldsymbol{h}_r\} \mid \boldsymbol{x_m}\right)$ in (6) can be dropped since it will disappear in (5). The likelihood function is equivalent to

$$q\left(\boldsymbol{x_m}, \{\boldsymbol{y}, \boldsymbol{h}\}\right) = \prod_{r=0}^{N_R-1} \frac{1}{(\pi N_0)^N} \exp\left(-\frac{||\boldsymbol{y}_r - \boldsymbol{h}_r \boldsymbol{x}_m||^2}{N_0}\right) \qquad (7)$$

Lemma 1 : Given that,

$$||\boldsymbol{y}_r - \boldsymbol{h}_r \boldsymbol{x}_m||^2 = ||\boldsymbol{y}_r||^2 + ||\boldsymbol{h}_r \boldsymbol{x}_m||^2 - 2\mathrm{Re}\left(\boldsymbol{y}_r \boldsymbol{h}_r^* \boldsymbol{x}_m^*\right)$$

Corollary 1 : Then the likelihood function after removing multiplicative terms independent of m is

$$q\left(\boldsymbol{x_m}, \{\boldsymbol{y}, \boldsymbol{h}\}\right) \propto \prod_{r=0}^{N_R-1} \exp\left(\frac{2\,\mathrm{Re}\left(\boldsymbol{y}_r \boldsymbol{h}_r^* \boldsymbol{x}_m^*\right) - ||\boldsymbol{h}_r \boldsymbol{x}_m||^2}{N_0}\right) \qquad (8)$$

Corollary 2 : The likelihood and log-likelihood ratio (LLR) of coded bit $e_j \in \{0, 1\}$ are respectively derived using (4, 5).
We typically simplify (5) via a *max-log approximation* $\log\left\{\sum_i \exp\left(\lambda_i\right)\right\} \sim \max_i \left\{\lambda_i\right\}$ letting (5) to be simplified as

$$\begin{aligned}
\mathcal{L}_j\left(\{\boldsymbol{y}\}\right) = \max_{m \, s.t. \, e_j=0} \frac{1}{N_0} \sum_{r=0}^{N_R-1} 2\,\mathrm{Re}\left(\boldsymbol{y}_r \boldsymbol{h}_r^* \boldsymbol{x}_m^*\right) - ||\boldsymbol{h}_r \boldsymbol{x}_m||^2 \\
- \max_{m \, s.t. \, e_j=1} \frac{1}{N_0} \sum_{r=0}^{N_R-1} 2\,\mathrm{Re}\left(\boldsymbol{y}_r \boldsymbol{h}_r^* \boldsymbol{x}_m^*\right) - ||\boldsymbol{h}_r \boldsymbol{x}_m||^2
\end{aligned} \qquad (9)$$

Remark 1 : This is considered as our ideal performance metric when comparing with those described in the subsequent section. They are also typically used in conventional receivers by replacing $\boldsymbol{h}_r$ with a least-squares estimate $\hat{\boldsymbol{h}}_r$.

3.2   Fading Channels

We describe BICM metrics for a general non-coherent fading channel with unknown phase on the line-of-sight components and fully unknown diffuse components, but known average gain. The overall unknown channel gain is given by $\boldsymbol{h}_r = A_r\left(\sqrt{\alpha}e^{j\theta_r} + \sqrt{1-\alpha}h_{r,f}\right)\boldsymbol{I}$ where θ_r is assumed to be i.i.d. uniform random variables on $[0, 2\pi)$ and $A_r > 0$ denotes average gain, $h_{r,f}$ is a zero-mean unit-variance circularly-symmetric complex Gaussian random variable and α is the relative strength of the LOS component. The amplitude $|h_{r,i}|$ on each receiver is thus Ricean distributed. It is worth noting that the i.i.d. assumption for the θ_r is somewhat unrealistic for a modern array receiver with accurate calibration. The phase differences would be more appropriately characterized by two random-phases, one originating from the time-delay between transmitter and receiver and the other from the angle of arrival of the incoming wave. The phase differences of individual antenna elements for a given carrier frequency could then be determined from the angle of arrival and the particular geometry of the array. To avoid assuming a particular array geometry, the i.i.d. uniform model provides a simpler and universal means to derive a receiver metric.

Proposition 1 : The corresponding likelihood function after neglecting multiplicative terms independent of the transmitted message, can be shown to be

$$q\left(\boldsymbol{x_m}, \{\boldsymbol{y}\}\right) = \prod_{r=0}^{N_R-1} \frac{1}{L_m} \exp\left(-\frac{\alpha A_r^2 \|\boldsymbol{x}_m\|^2}{\boldsymbol{L}_m} + \beta_m \left|\boldsymbol{x}_m^{\mathrm{H}} \boldsymbol{y}_r\right|^2 \right) \times \mathrm{I}_0\left(\frac{2\sqrt{\alpha}\boldsymbol{A}_r}{\boldsymbol{L}_m} \left|\boldsymbol{x}_m^H \boldsymbol{y}_r\right| \right),$$

(10)

where $\boldsymbol{L}_m = N_0 + 2(1-\alpha)A_r^2 \|\boldsymbol{x}_m\|^2$, $\boldsymbol{\beta}_m = \frac{2(1-\alpha)A_r^2}{N_0(N_0+2(1-\alpha)A_r^2\|\boldsymbol{x}_m\|^2)}$ and $I_0(x)$ is the zero-order modified Bessel function. The likelihood and log-likelihood ratio (LLR) of coded bit $e_j \in \{0, 1\}$ are respectively derived using (4, 5).
Proof : See Appendix section A.
Corollary 2 : Metric calculations based on (10) are computationally complex from an implementation perspective and are typically simplified. As is the case for the known channel, we can apply the *max-log approximation* after first using an exponential approximation $I_0(z) \sim \frac{e^z}{\sqrt{2\pi z}} \sim e^z$ yielding the approximated log-likelihood ratio (LLR) for coded bit j given

$$\mathcal{L}_j(\{\boldsymbol{y}\}) = \max_{m\ s.t.\ e_j=0}\left(\sum_{r=0}^{N_R-1} -\frac{\alpha A_r^2 \|\boldsymbol{x}_m\|^2}{\boldsymbol{L}_m} + \beta_m \left|\boldsymbol{x}_m^{\mathrm{H}} \boldsymbol{y}_r\right|^2 + \frac{2\sqrt{\alpha}\boldsymbol{A}_r}{\boldsymbol{L}_m} \left|\boldsymbol{x}_m^H \boldsymbol{y}_r\right| \right) - \sum_{m\ s.t.\ e_j=0} \sum_{r=0}^{N_R-1} \log(\boldsymbol{L}_m)$$
$$- \max_{m\ s.t.\ e_j=1}\left(\sum_{r=0}^{N_R-1} -\frac{\alpha A_r^2 \|\boldsymbol{x}_m\|^2}{\boldsymbol{L}_m} + \beta_m \left|\boldsymbol{x}_m^{\mathrm{H}} \boldsymbol{y}_r\right|^2 + \frac{2\sqrt{\alpha}\boldsymbol{A}_r}{\boldsymbol{L}_m} \left|\boldsymbol{x}_m^H \boldsymbol{y}_r\right| \right) - \sum_{m\ s.t.\ e_j=1} \sum_{r=0}^{N_R-1} \log(\boldsymbol{L}_m)$$

(11)

Remark 2 : Note that in (10), many of the terms can be dropped when $|\boldsymbol{x}_m|$ is constant for all m, as would be the case for BPSK or QPSK modulation for instance. Strong LOS channels can also neglect the quadratic terms in (10). when $\alpha = 1$, corresponding to a pure LOS channel, the likelihood simplifies to the case of the classical non-coherent channel [29][Sec. 5-4-2].

3.3   Joint Estimation and Detection Principle

For the case of polar or LDPC-coded data, we are motivated for complexity reasons to segment the coded streams into small bloks for detection. Under an ideal interleaving assumption [9] with known channels, detection can be performed individual modulated symbols. With joint detection and estimation and interleaved DMRS and data symbols, we will consider short blocks comprising both data and DMRS over which to compute the above metrics. This consists in subdividing the N-dimensional vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ into smaller segments of N_b-dimensional blocks and applying the bit LLR metric on each of the underlying segments.
Proposition 2 : Noticing the form of the metrics and the fact that the data and DMRS symbols do not overlap, we can easily see that an estimated channel impulse response (CIR) is part of the metrics. By writing $\boldsymbol{x}_m^{(i)} = \boldsymbol{x}_{m,p}^{(i)} + \boldsymbol{x}_{m,d}^{(i)}$ where d, p and i are subscripts representing data, DMRS components, i^{th} segment of

block respectively, we can reveal $\hat{h}_r^{\text{LS}}$ in the metrics:

$$\left| \boldsymbol{x}_m^{(i)\text{H}} \boldsymbol{y}_r^{(i)} \right| = \left| \underbrace{\boldsymbol{x}_{m,p}^{(i)\text{H}} \boldsymbol{y}_{r,p}^{(i)}}_{\text{channel estimate}} + \boldsymbol{x}_{m,d}^{(i)\text{H}} \boldsymbol{y}_{r,d}^{(i)} \right|$$

$$= \left| \hat{\boldsymbol{h}}_r^{\text{LS}} + \boldsymbol{y}_{r,d}^{(i)} \boldsymbol{x}_{m,d}^{(i)\text{H}} \right| \tag{12}$$

where $\hat{\boldsymbol{h}}_r^{\text{LS}}$ is the CIR after performing Least-squares (LS) channel estimation using averaging or smoothing over an appropriate number of dimensions exhibiting channel coherence. In the process of short-block detection, we can make use of such a channel estimate that In general, the channel estimation procedure will work as usual and the resulting estimates are fed into the metrics considered here.
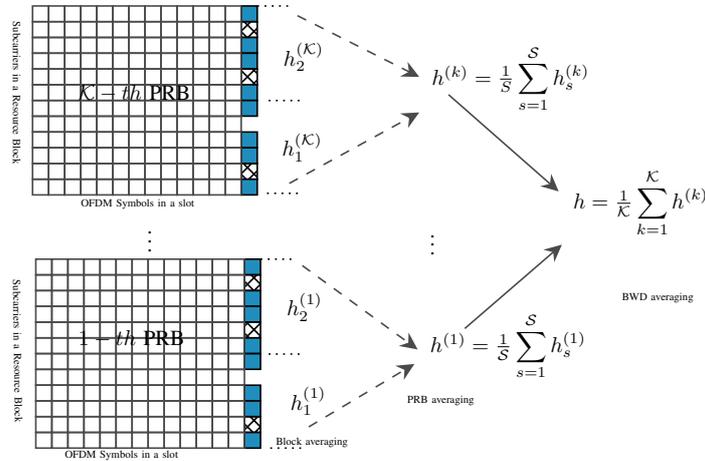
Figure 3: Conceptual illustration of the Joint Estimation and Detection Principle

4 Numerial Results

4.1 Metric Performance Analysis

The simulations are based on NR POLAR and NR LDPC coding schemes paired with QPSK and 16 QAM modulations. The transmission process involves a transport block length of 48 bits. The resource population process is conducted using a single OFDM symbol with 4 PRBs and 48 REs (32 REs for data components and 16 REs for DMRS components), wherein the DMRS sequences occupy 4 REs per PRB. This transmission structure, where the reference and data components are transmitted concurrently within common OFDM symbols, is commonly employed in the PUCCH or PUSCH, as well as in some downlink control channels. Hence, from the perspective of the MAC and PHY layers, the underlying transmission utilizing BICPM can be considered as a PUCCH2 transmission for medium to long block lengths (> 11 bits), while that utilizing BILCM can be viewed as a PUSCH transmission. The results illustrated in Figure 4 show the performance of the Bit Interleaved Coded Modulation (BICM) for joint estimation and detection over a gaussian channel, specifically

when $\alpha = 1$ is assessed to understand the performance discrepancy between the *Perfect CSI* and *No CSI* situations in extreme coverage scenarios characterized by low signal-to-noise ratio.
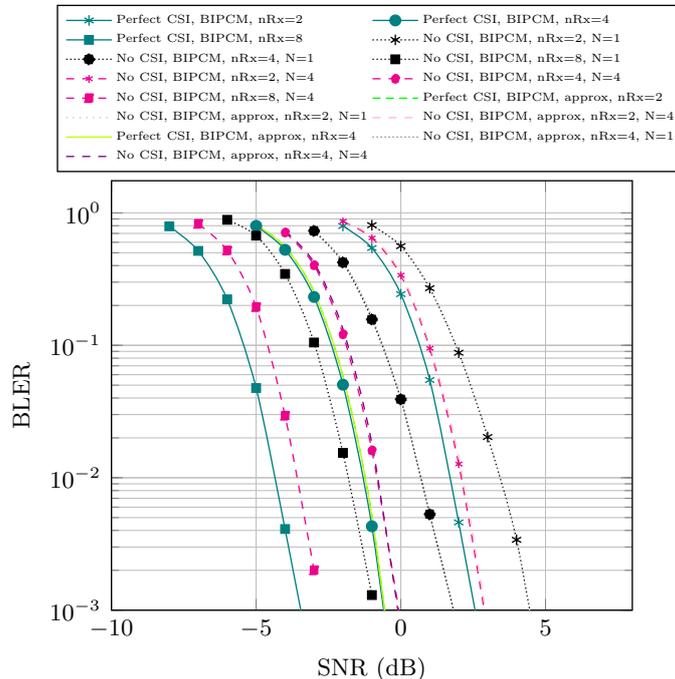
Figure 4: Block Error Rate, 48 bits(TBs+CRC), NRPOLAR BICM,R=48/64, (CRC)-aided successive-cancellation list decoder(List length=8), QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs (32 data, 16 dmrs), nRx = 4 classical non-coherent channel

Note that the $N = 1$ case also corresponds to the conventional receiver involving channel estimation. The joint estimation/detection approach yields a perfomance gain of 1.25 dB, 1.5 dB and 1.75 dB using $N = 4$ over 2, 4 and 8 receive antennas respectively. From this insight, it is apparent that when the number of antennas increases, the performance gap between the *Perfect CSI* and the *No CSI* situations (e.g., $N = 4$) expands. Furthermore, the graphs indicate that the max-log metric performs nearly as well as the accurate metric (e.g., at nRx = 2). This leads to the conclusion that when Gray-mapped constellations are employed, the max-log metric is known to have a minimal impact on receiver performance. However, as the modulation order increases, the difference in performance between optimal and suboptimal techniques for generating LLRs becomes significant as discussed in [30] [31]. The logarithmic calculations tied to the precise metric add an extra layer of complexity when incorporating the requisite multiplicative and additive operations during LLR processing. Given this, it is deemed more reasonable to employ the max-log approximation as a means of mitigating the underlying computational complexity. Similarly, the results in Figure 4 using BIPCM are congruent with those presented in Figure 5 that employs BILCM, in both single and multiple antenna configurations. Although the code rates and transmission parameters are identical, BIPCM offers significantly better performance gains than BILCM. This is potentially due to the fact that the 3GPP polar code has been optimized for very short block lengths, while the 3GPP LDPC code targets much longer

transport block lengths.
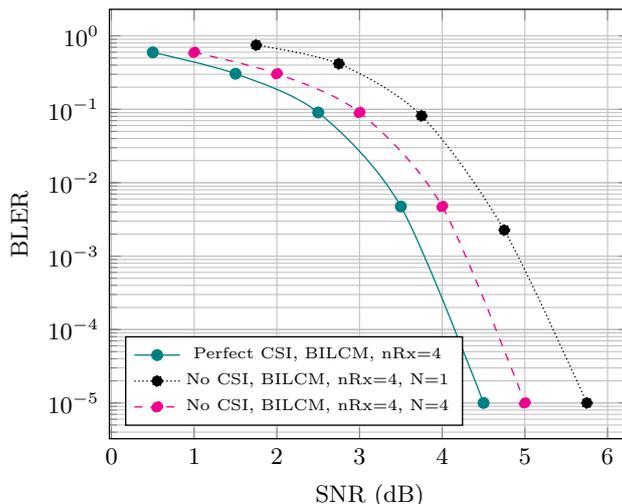
Figure 5: Block Error Rate, 48 bits(TBs+CRC), NRLDPC BICM, R=48/64, Layered belief propagation decoder, iteration=30, QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs (32 data, 16 dmrs), nRx = 4, classical non-coherent channel

In addition, to determine the potential benefit obtained from the joint estimation and detection principle, we perform evaluations on realistic frequency selective channels, particularly over the TDL-C NLOS wireless channel. TDL-C is a 3GPP reference channel model with a long delay-spread and are especially used to emphasize non-MIMO assessments [32]. The simulations consider a delay spread of 300 ns and a 4-antenna configuration with independent and identically distributed realizations at each antenna port. The metric utilized is given by (10) with $\alpha = 1$. For this purpose, we deal with low(QPSK) and high(16 QAM) order modulations for a comprehensive analysis. In fact, Quadrature Amplitude Modulations are used in typical wireless digital communications. The main difference between M-QAM and QPSK is that their spectral width is narrower than that of QPSK. The findings reveal that employing block detection provides an advantage across all modulation orders, with discernible gains in performance. A gain of 2.25 dB, and 2 dB is observed at $N = 4$ for QPSK and 16 QAM configurations, respectively. In addition, it should be noted that the BLER of 16-QAM is still much higher than that of QPSK.

Furthermore, the outcomes of the system's performance in line-of-sight (LOS) channels with varying levels of α are showcased in Figure 7.

When $\alpha = 1$, the channel exhibits Rayleigh fading, while for $\alpha = 0$, it behaves as a classical non-coherent Gaussian channel. Interestingly, when α is less than or equal to 0.5, the system's performance closely resembles that of a pure Rayleigh channel, implying that the fading component dominates the Gaussian component due to the quadratic term in the metric expression. Empirical findings for $\alpha = 0.75$ indicate a moderate level of performance, falling between the results associated with $\alpha = 0$ and $\alpha = 1$. It is worth noting that the precise value of alpha remains undisclosed to the receiver, necessitating the need for estimation. Nevertheless, regardless of the estimated value of α, it seems to have no discernible impact at the receiver level when $\alpha \leq 0.5$ at the transmitter level.

14
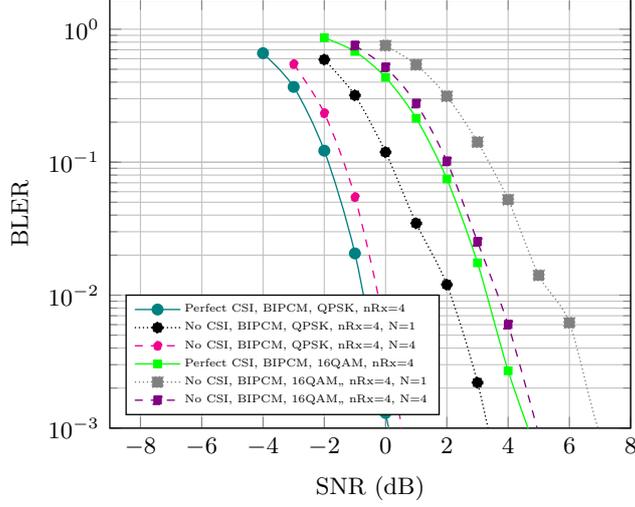
Figure 6: Block Error Rate, 48 bits(TBs+CRC), NRPOLAR BICM, R=48/64, (CRC)-aided successive-cancellation list decoder,(4 PRBs, 48 REs (32 data, 16 dmrs)), 1 OFDM symbol, QPSK, 16 QAM modulations, nRx = 4, TDL-C channel
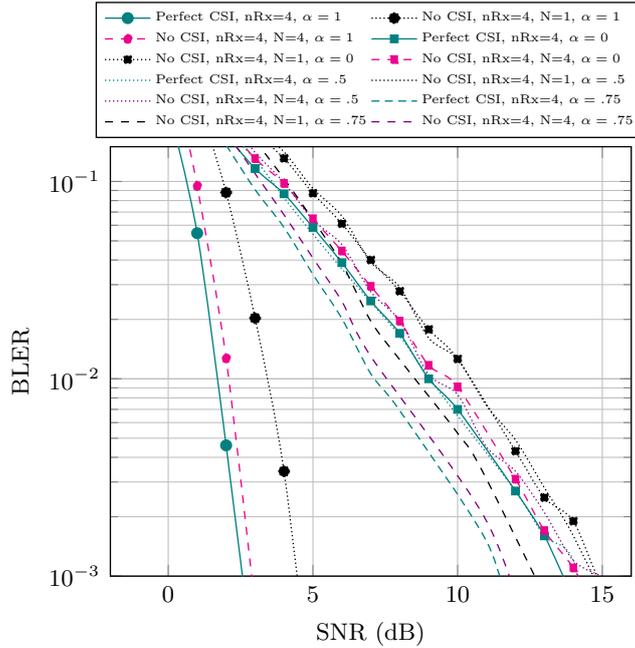
Figure 7: Block Error Rate, 48 bits(TBs+CRC), NRPOLAR BICM,R=48/64, (CRC)-aided successive-cancellation list decoder(List length=8), QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs (32 data, 16 dmrs), nRx = 2, Fading Channels, $\alpha = \{0, \ 0.5, \ 0.75, 1\}$

Finally, we can assess the above results with respect to the finite block length bounds that have been established in the scientific literature [3] [33] [34]. For a more comprehensive understanding of the bounds utilized in Figure 8, interested readers are encouraged to refer to the works of authors [1] and [5]. For this purpose, we consider the *metaconverse* (MC) bound for a thorough

15

comparative analysis. It can be observed that when the block error rate (BLER)
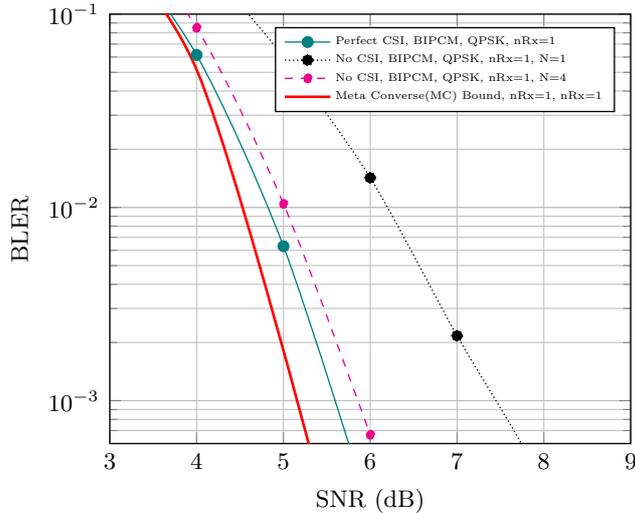
Figure 8: Block Error Rate, 48 bits(TBs+CRC), NR POLAR BICM ,R=48/64, (CRC)-aided successive-cancellation list decoder(List length=8), QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs (32 data, 16 dmrs), nRx = 1(SISO), classical non coherent Channel, vs Metaconverse bounds

reaches a threshold of 1%, the performance difference between the *MC Bound* and the *No-CSI*$(N = 4)$ is 0.5 dB, compared to 1.75 dB for the *No-CSI*$(N = 1)$, in a scenario with a single receive antenna.

4.2   Impact of DMRS density

In instances where the reference and data symbols are jointly conveyed in common OFDM symbols, we can look into the impact of dmrs density on performance. Figure 9 illustrates the resources mapping process according to the density of dmrs per PRB ranging from sparse to dense (i.e. 2, 3, 4 and 6. Fig-
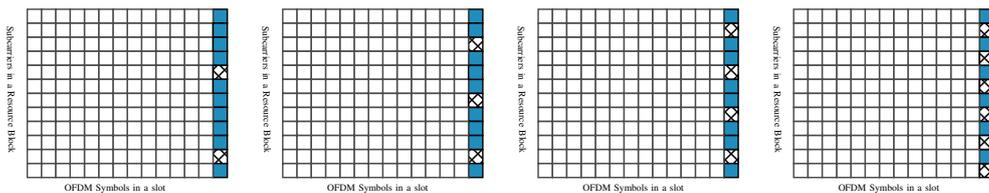
Figure 9: Block Error Rate, 24 bits(TBs), NR POLAR BICM, (CRC)-aided successive-cancellation list decoder, QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs, set of ($\{24, 32, 36, 40\}$ data, $\{24, 16, 12, 8\}$ dmrs), nRx = 4, classical non-coherent channel

ure 10 depicts the performance on the awgn channel in the situations *Perfect CSI* and *No CSI (N = 1, N = 4)* depending on the density of dmrs per PRB.

To achieve a comprehensive understanding of the results, Table 1 presents the disparity between Perfect and Imperfect Channel State Information (CSI) for $N = 1$ and $N = 4$ in correlation with the distribution of DMRS per Physical Resource Block (PRB). In essence, fewer DMRS has merit of additional coding

16
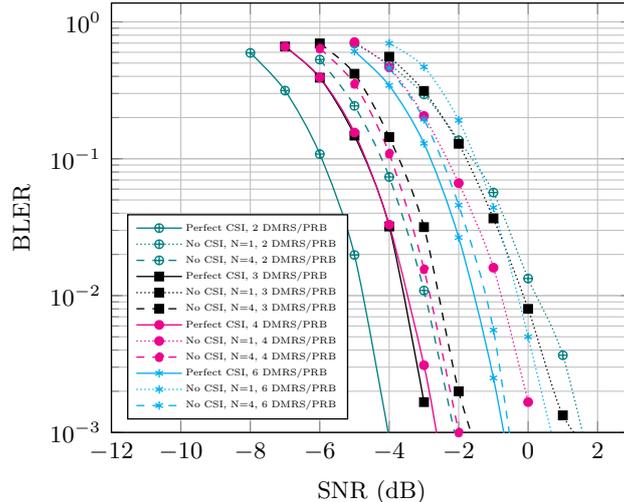
Figure 10: Block Error Rate, 24 bits(TBs), NR POLAR BICM, (CRC)-aided successive-cancellation list decoder, QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs, set of ($\{24, 32, 36, 40\}$ data, $\{24, 16, 12, 8\}$ dmrs), nRx = 4, classical non-coherent channel

Table 1: Performance gap between Perfect CSI (PCSI) vs No CSI (N=4 and N=1) cases over nRx=4

dmrs per PRB	Gap(N=4) [dB]	Gap(N=1) [dB]
2	1.75	4.75
3	1	3.5
4	0.625	2.625
6	0.375	1.375

rates. Therefore, performance improves as DMRS density decreases. However, it should be noted that even with $N = 4$, a low dmrs density setup expands the performance gap between *Perfect CSI* and *No CSI*. It may be advantageous in some instances to maintain the density of DMRSs in a certain sweet spot or simply to rely on sparse or even low DMRS density while increasing their power via an adaptive adjustment. More specifically, a precise approach is to identify the configuration with the minimum number of DMRSs which allows the transmitter to slightly increase the power of the underlying signals. However, choosing a low dmrs density has a detrimental effect on channel estimate quality. Even if the receiver with block detection ($N = 4$) seems to be less sensitive to it with respect to the conventional receiver ($N = 1$). There appears to be a sweet spot in terms of dmrs density per PRB, as evidenced by the results presented in Figure 10. Therefore, the ideal DMRS distribution setup is obtained by incorporating four DMRSs per physical resource block (PRB) compared to those employing two, three or six DMRSs per PRB , using the block detection principle (N=1, N=4). In practice, transmission with a low density of dmrs appears to be more valuable and should be favoured in future communication standards in order to convey more data symbols than reference signals. Consequently, it is advisable to consider configurations with either one or two DMRs per PRB. However in order to reap from the low dmrs density, it is important to carry out some sort of *adaptive dmrs/data power adjustment*

that would enhance the channel estimate accuracy, leading to an improvement in performance from a holistic perspective. For this purpose, the system model can be reconceived as $\boldsymbol{y}_r = \left(\boldsymbol{x}_m^{(d)} + \beta \boldsymbol{x}_m^{(p)} \right) \boldsymbol{h}_r + \boldsymbol{z}_r$. The adaptive power adjustment procedure is contingent on the values of β. The dmrs Power is to be slightly increased in a judicious fashion since β must be perfectly calibrated to ensure compliance with potential radio frequency constraints. As depicted in Figure 11, the performance improvement can be observed as a function of varying values of β. The optimal performance enhancement is achieved when
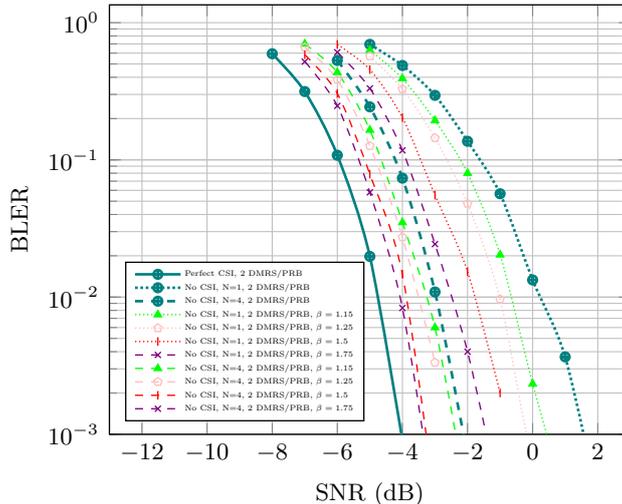
Figure 11: Block Error Rate, 24 bits(TBs), NR POLAR BICM, (CRC)-aided successive-cancellation list decoder, QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs, (40REs= data, 8REs= dmrs), nRx = 4, over a classical non-coherent channel, adaptive power adjustment via β with N = 4 and N=1

β is set to 1.75. It is noted that by selecting $\beta = 1.5$, a gain of 1 dB and 2 dB can be attained when N is equal to 4 and 1, respectively. Overall, the implications of varying DMRS density within the 3GPP standard are significant. Specifically, it is feasible to reduce the number of DMRS per PRB to one or two, while allowing the User Equipment (UE) to adjust the power allocation between the DMRS and data transmission. This flexibility in DMRS density and power allocation is transparent to the receiver.

5 Conclusions

This paper presented novel bit-interleaved coded modulation metrics for joint estimation detection using a training or reference signal transmission strategy for medium to long block length channels. We showed that it is possible to enhance the performance and sensitivity of advanced receivers, especially when channel state information is unknown and the density of training dimensions is low. The proposed techniques take advantage of joint estimation/detection. The performance analysis made use of a full 5G transmitter and receiver chain for both Polar and LDPC coded transmissions paired with M-ary PSK/QAM modulation schemes. We considered transmissions where reference signals are interleaved with data and both are transmitted over a small number of OFDM symbols so that near-perfect channel estimation cannot be achieved. This is

particularly adapted to mini-slot transmissions for ultra-reliable low-latency communications or short-packet random-access use-cases. We characterized the performance for up to eight receiving antenna configurations in order to determine the performance gain offered by the proposed BICM detection in realistic basestation receiver scenarios. Our findings demonstrate that BICM metrics combined with the joint estimation/detection principle can be used to achieve detection performance that is close to that of a coherent receiver with perfect channel state information for both polar and LDPC coded configurations. Furthermore, we show that for transmissions with low DMRS density, a good trade-off can be achieved in terms of additional coding gain and improved channel estimation quality by adaptive DMRS power adjustment. Future work will include investigations of higher-spectral efficiency joint detection-estimation bit-metrics for MIMO-QAM.

Appendix

$$P\left(\{\boldsymbol{y}_r\} \mid \boldsymbol{x_m}\right) = \frac{1}{2\pi \det \Phi} \exp\left(-\frac{1}{2}\left(\boldsymbol{y}_r - \mu\left(\boldsymbol{x}_m, \theta_r\right)\right)^{\mathrm{H}} \Phi^{-1}\left(\boldsymbol{y}_r - \mu\left(\boldsymbol{x}_m, \theta_r\right)\right)\right)$$
$$= \frac{1}{2\pi \det \Phi} \exp\left(-\frac{1}{2}\left(\boldsymbol{y}_r - \sqrt{\alpha}A_r e^{j\theta_r}\boldsymbol{x}_m\right)^{\mathrm{H}} \Phi^{-1}\left(\boldsymbol{y}_r - \sqrt{\alpha}A_r e^{j\theta_r}\boldsymbol{x}_m\right)\right)$$

Covariance Matrix

knowing that

$$\boldsymbol{y}_r - \sqrt{\alpha}A_r e^{j\theta_r}\boldsymbol{x}_m = \sqrt{1-\alpha}A_r\boldsymbol{h}_{r,f}\boldsymbol{x}_m + \boldsymbol{z}_r \tag{13}$$

then

$$\Phi \triangleq \frac{1}{2}\mathbb{E}\left[\left(\sqrt{1-\alpha}A_r\boldsymbol{h}_{r,f}\boldsymbol{x}_m + \boldsymbol{z}_r\right)\left(\sqrt{1-\alpha}A_r\boldsymbol{h}_{r,f}\boldsymbol{x}_m + \boldsymbol{z}_r\right)^{\mathrm{H}}\right]$$
$$\triangleq (1-\alpha)A_r^2\boldsymbol{x}_m\boldsymbol{x}_m^{\mathrm{H}}\sigma_h^2 + \sigma_z^2\boldsymbol{I}_N, \text{ where } \sigma_h^2 = 1 \tag{14}$$
$$\triangleq (1-\alpha)A_r^2\boldsymbol{x}_m\boldsymbol{x}_m^{\mathrm{H}} + \frac{N_0}{2}I_N$$

Determinant

$$\det \Phi = \det\left((1-\alpha)A_r^2\boldsymbol{x}_m\boldsymbol{x}_m^{\mathrm{H}} + \sigma_z^2\boldsymbol{I}\right)$$
$$= \det\left(\sigma_z^2\boldsymbol{I} + k\boldsymbol{x}_m\boldsymbol{x}_m^{\mathrm{H}}\right), \text{ where, } k = (1-\alpha)A_r^2 \tag{15}$$

Applying the following mathematical properties:
(i) $det(\boldsymbol{AB}) = det(\boldsymbol{BA})$,
(ii) $det(\boldsymbol{I} + \boldsymbol{AB}) = det(\boldsymbol{I} + \boldsymbol{BA})$, Cf. Sylvester's determinant theorem [35].

$$\det \Phi = \frac{1}{2}\left(N_0 + 2(1-\alpha)A_r^2 \|\boldsymbol{x}_m\|^2\right) \tag{16}$$

Inverse of Φ (Φ^{-1})

Φ involves the sum of two matrices. Thus, some matrix inversion lemmas must be introduced. The most common one is the Woodbury Matrix identity (i.e. matrix inversion lemma, Sherman–Morrison–Woodbury formula, or just

Woodbury formula).

*Lemma :*The Woodbury matrix identity [36]

$$(\boldsymbol{A} + \boldsymbol{UCV})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{C}^{-1} + \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}\boldsymbol{A}^{-1}, \qquad (17)$$

where $\boldsymbol{A}$, $\boldsymbol{U}$, $\boldsymbol{C}$ and $\boldsymbol{V}$ are conformable matrices: A is $n \times n$, $\boldsymbol{C}$ is $k \times k$, $\boldsymbol{U}$ is $n \times k$, and $\boldsymbol{V}$ is $k \times n$.

Here we have the special case where $\boldsymbol{V}, \boldsymbol{U}$ are vectors, as in *Sherman–Morrison formula* where $\boldsymbol{U} = \boldsymbol{V} = \boldsymbol{I}$.

let say :

$$\begin{cases} \boldsymbol{A} = \sigma_z^2 \boldsymbol{I} \\ \boldsymbol{C} = (1-\alpha)A_r^2\boldsymbol{I} = k\boldsymbol{I}, \text{ where, } k = (1-\alpha)A_r^2 \\ \boldsymbol{U} = \boldsymbol{x}_m \\ \boldsymbol{V} = \boldsymbol{x}_m^{\mathrm{H}} \end{cases} \qquad (18)$$

$$\begin{aligned} \Phi^{-1} &= (\boldsymbol{A} + \boldsymbol{UCV})^{-1} \\ &= \frac{2}{N_0} - \frac{2}{N_0}\boldsymbol{x}_m \left[\frac{2k}{N_0 + 2k\|\boldsymbol{x}_m\|^2}\right]\boldsymbol{x}_m^{\mathrm{H}} \\ &= \frac{2}{N_0} - \frac{2}{N_0}\boldsymbol{x}_m \left(\frac{2(1-\alpha)A_r^2}{N_0 + 2(1-\alpha)A_r^2\|\boldsymbol{x}_m\|^2}\right)\boldsymbol{x}_m^{\mathrm{H}} \end{aligned} \qquad (19)$$

let say $\beta_m = \frac{2(1-\alpha)A_r^2}{N_0(N_0 + 2(1-\alpha)A_r^2\|\boldsymbol{x}_m\|^2)}$

then

$$\Phi^{-1} = \frac{2}{N_0} - 2\boldsymbol{x}_m\beta_m\boldsymbol{x}_m^{\mathrm{H}} \qquad (20)$$

Likelihood function

let's say $\boldsymbol{\mu} = \sqrt{\alpha}A_r e^{j\theta_r}\boldsymbol{x}_m$

$$\begin{aligned} q(\boldsymbol{x_m}, \{\boldsymbol{y}\}) &= \prod_{r=0}^{N_R-1} \frac{1}{2\pi\det\Phi}\exp\left(-\frac{1}{2}(\boldsymbol{y}_r - \mu)^{\mathrm{H}}\left(\frac{2}{N_0} - 2\boldsymbol{x}_m\beta_m\boldsymbol{x}_m^{\mathrm{H}}\right)(\boldsymbol{y}_r - \mu)\right) \\ &= \prod_{r=0}^{N_R-1} \frac{1}{2\pi\det\Phi}\exp\left(-\frac{1}{N_0}|\boldsymbol{y}_r - \mu|^2 + \beta_m\left|(\boldsymbol{y}_r - \mu)^{\mathrm{H}}\boldsymbol{x}_m\right|^2\right) \end{aligned}$$

By extending the terms into the exponential,
and ignoring those that are independent of m, we obtain

$$\alpha A_r^2\|\boldsymbol{x}_m\|^2\left(\frac{1}{N_0} - \beta_m|\boldsymbol{x}_m|^2\right) + \beta_m\left|\boldsymbol{x}_m^{\mathrm{H}}\boldsymbol{y}_r\right|^2$$

$$+2\sqrt{\alpha}\boldsymbol{A}_r\left(\frac{1}{N_0} - \beta_m|\boldsymbol{x}_m|^2\right)|\boldsymbol{x}_m^H\boldsymbol{y}_r|\cos(\phi_r + \theta_r)$$

since θ is unknown, the likelihood function is equivalent to

$$q(\boldsymbol{x_m}, \{\boldsymbol{y}\}) = \prod_{r=0}^{N_R-1} \frac{1}{2\pi\det\Phi}\int_0^{2\pi}\exp\left(-\frac{1}{N_0}|\boldsymbol{y}_r - \mu|^2 + \beta_m\left|\left(\boldsymbol{y}_r^{\mathrm{H}} - \mu^{\mathrm{H}}\right)\boldsymbol{x}_m\right|^2\right)\mathrm{d}\theta_r$$

$$q\left(\boldsymbol{x_m}, \{\boldsymbol{y}\}\right) = \prod_{r=0}^{N_R-1} \frac{1}{2\pi \det \Phi} \exp\left(-\alpha A_r^2 \left\|\boldsymbol{x}_m\right\|^2 \left(\frac{1}{N_0} - \beta_m \left|\boldsymbol{x}_m\right|^2\right) \beta_m \left|\boldsymbol{x}_m^{\mathrm{H}} \boldsymbol{y}_r\right|^2\right)$$
$$\times \int_0^{2\pi} \exp\left(2\sqrt{\alpha}\boldsymbol{A}_r \left(\frac{1}{N_0} - \beta_m \left|\boldsymbol{x}_m\right|^2\right) |\boldsymbol{x}_m^H \boldsymbol{y}_r| \cos\left(\phi_r + \theta_r\right)\right) \mathrm{d}\theta_r$$

knowing that $\frac{1}{\pi} \int_{\varphi=0}^{\pi} \exp(z cos(\varphi)) \mathrm{d}\varphi = \mathrm{I}_0(z)$ [37], where $I_0(\cdot)$ is the zero-th order Modified Bessel function of the first kind.

$$q\left(\boldsymbol{x_m}, \{\boldsymbol{y}\}\right) = \prod_{r=0}^{N_R-1} \frac{2}{N_0 + 2(1-\alpha)A_r^2 \left\|\boldsymbol{x}_m\right\|^2} \exp\left(-\alpha A_r^2\right.$$
$$\left\|\boldsymbol{x}_m\right\|^2 \left(\frac{1}{N_0} - \beta_m \left|\boldsymbol{x}_m\right|^2\right) + \beta_m \left|\boldsymbol{x}_m^{\mathrm{H}} \boldsymbol{y}_r\right|^2\right) \times \tag{21}$$
$$\mathrm{I}_0\left(2\sqrt{\alpha}\boldsymbol{A}_r \left(\frac{1}{N_0} - \beta_m \left|\boldsymbol{x}_m\right|^2\right) \left|\boldsymbol{x}_m^H \boldsymbol{y}_r\right|\right)$$

Let Say $\boldsymbol{L}_m = N_0 + 2(1-\alpha)A_r^2 \left\|\boldsymbol{x}_m\right\|^2$, and then after ignoring multiplicative term that are independent of m

it comes

$$q\left(\boldsymbol{x_m}, \{\boldsymbol{y}\}\right) \propto \prod_{r=0}^{N_R-1} \frac{1}{\boldsymbol{L}_m} \exp\left(-\alpha A_r^2 \left\|\boldsymbol{x}_m\right\|^2 \left(\frac{1}{N_0} - \beta_m \left\|\boldsymbol{x}_m\right\|^2\right)\right.$$
$$\left. + \beta_m \left|\boldsymbol{x}_m^{\mathrm{H}} \boldsymbol{y}_r\right|^2\right) \times \mathrm{I}_0\left(2\sqrt{\alpha}\boldsymbol{A}_r \left(\frac{1}{N_0} - \beta_m \left\|\boldsymbol{x}_m\right\|^2\right) \left|\boldsymbol{x}_m^H \boldsymbol{y}_r\right|\right) \tag{22}$$

Expressing $\boldsymbol{\beta}_m$ as a function of $\boldsymbol{L}_m$, we have the relation

$$\beta_m = \frac{1}{\left\|\boldsymbol{x}_m\right\|^2 N_0} - \frac{1}{\left\|\boldsymbol{x}_m\right\|^2 \boldsymbol{L}_m} \tag{23}$$

Bibliography

[1] M. Xhemrishi, M. C. Coşkun, G. Liva, J. Östmanl and G. Durisi, "List Decoding of Short Codes for Communication over Unknown Fading Channels," 2019 53rd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2019, pp. 810-814

[2] P. Yuan, M. C. Coşkun and G. Kramer, "Polar-Coded Non-Coherent Communication," in IEEE Communications Letters, vol. 25, no. 6, pp. 1786-1790, June 2021.

[3] Y. Polyanskiy, H. V. Poor and S. Verdu, "Channel Coding Rate in the Finite Blocklength Regime," in IEEE Transactions on Information Theory, vol. 56, no. 5, pp. 2307-2359, May 2010

[4] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy and W. Yang, "Short-Packet Communications Over Multiple-Antenna Rayleigh-Fading Channels," in IEEE Transactions on Communications, vol. 64, no. 2, pp. 618-629, Feb. 2016.

[5] J. Östman, G. Durisi, E. G. Ström, M. C. Coşkun and G. Liva, "Short Packets Over Block-Memoryless Fading Channels: Pilot-Assisted or Noncoherent Transmission?," in IEEE Transactions on Communications, vol. 67, no. 2, pp. 1521-1536, Feb. 2019.

[6] A. Lancho, J. Östman, G. Durisi, T. Koch and G. Vazquez-Vilar, "Saddlepoint Approximations for Short-Packet Wireless Communications," in IEEE Transactions on Wireless Communications, vol. 19, no. 7, pp. 4831-4846, July 2020.

[7] E. Zehavi, "8-PSK trellis codes for a Rayleigh fading channel",IEEE Transactions on Communication, vol.40, pp. 873–883, 1992, May 1992.

[8] A. G. i Fàbregas, A Martinez and G. Caire, "Bit-Interleaved Coded Modulation", Foundations and Trends® in Communications and Information Theory Vol. 5: No. 1–2, pp 1-153, November 2008.

[9] G. Caire, G. Taricco and E. Biglieri, " Bit- Interleaved Coded Modulation", IEEE Transactions on Information Theory, vol. 44, pp. 927–946, May 1998.

[10] 3GPP TS 38.212 V16.2.0, " Technical Specification Group Radio Access Network, Multiplexing and channel coding", July 2020.

[11] E. Arikan, "Channel Polarization: A Method for constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels", IEEE Transactions on Information Theory, vol. 55, No. 7, pp. 3051-3073, July 2009.

[12] S. Ahmadi, " 5G NR, Architecture, Technology, Implementation, and Operation of 3GPP New Radio Standard",London, United Kingdom : Academic Press, an imprint of Elsevier, 2019.

[13] A. Balatsoukas-Stimming, M. B. Parizi and A. Burg, "LLR-Based Successive Cancellation List Decoding of Polar Codes," in IEEE Transactions on Signal Processing, vol. 63, no. 19, pp. 5165-5179, Oct.1, 2015, doi: 10.1109/TSP.2015.2439211.

[14] I. Tal and A. Vardy, "List Decoding of Polar Codes," in IEEE Transactions on Information Theory, vol. 61, no. 5, pp. 2213-2226, May 2015.

[15] Q. Zhang, A. Liu, X. Pan and K. Pan, "CRC Code Design for List Decoding of Polar Codes," in IEEE Communications Letters, vol. 21, no. 6, pp. 1229-1232, June 2017, doi: 10.1109/LCOMM.2017.2672539.

[16] E. Arikan, "A performance comparison of polar codes and Reed-Muller codes," in IEEE Communications Letters, vol. 12, no. 6, pp. 447-449, June 2008, doi: 10.1109/LCOMM.2008.080017.

[17] U. U. Fayyaz and J. R. Barry, "Polar codes for partial response channels," 2013 IEEE International Conference on Communications Giard, P., Balatsoukas-Stimming, A., Sarkis, G. et al., "Fast Low-Complexity Decoders for Low-Rate Polar Codes", Journal of Signal Processing Systems, vol. 90, pp. 675–685, August 2016.

[18] J. Lin, J. Sha, L. Li, C. Xiong, Z. Yan and Z. Wang, "A high throughput belief propagation decoder architecture for polar codes," 2016 IEEE International Symposium on Circuits and Systems (ISCAS), 2016, pp. 153-156, doi: 10.1109/ISCAS.2016.7527193.

[19] J. Lin, Z. Yan and Z. Wang, "Efficient Soft Cancelation Decoder Architectures for Polar Codes," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 25, no. 1, pp. 87-99, Jan. 2017, doi: 10.1109/TVLSI.2016.2577883.

[20] J. Lin, C. Xiong and Z. Yan, "Reduced complexity belief propagation decoders for polar codes," 2015 IEEE Workshop on Signal Processing Systems (SiPS), 2015, pp. 1-6, doi: 10.1109/SiPS.2015.7344984.

[21] A. Alamdar-Yazdi and F. R. Kschischang, "A Simplified Successive-Cancellation Decoder for Polar Codes," in IEEE Communications Letters, vol. 15, no. 12, pp. 1378-1380, December 2011, doi: 10.1109/LCOMM.2011.101811.111480.

[22] B. Yuan and K. K. Parhi, "Early Stopping Criteria for Energy-Efficient Low-Latency Belief-Propagation Polar Code Decoders," in IEEE Transactions on Signal Processing, vol. 62, no. 24, pp. 6496-6506, Dec.15, 2014, doi: 10.1109/TSP.2014.2366712.

[23] U. U. Fayyaz and J. R. Barry, "Low-Complexity Soft-Output Decoding of Polar Codes," in IEEE Journal on Selected Areas in Communications, vol. 32, no. 5, pp. 958-966, May 2014, doi: 10.1109/JSAC.2014.140515.

[24] K. Niu and K. Chen, "CRC-Aided Decoding of Polar Codes", IEEE Communications Letters, vol. 16, No. 10, pp.1668-167, October 2015.

[25] R. G. Gallager, "Low Density Parity Check Codes", Cambridge, USA: MIT Press, July 1963.

[26] T. Nguyen, T. N. Tan, and H. Lee, "Efficient QC-LDPC Encoder for 5G New Radio", Electronics, vol. 8, p. 668, June 2019

[27] K. Sun and M. Jiang, "A Hybrid Decoding Algorithm for Low-Rate LDPC codes in 5G," 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), 2018, pp. 1-5, doi: 10.1109/WCSP.2018.8555597.

[28] LIFANG WANG "Implementation of Low-Density Parity-Check codes for 5G NR shared channels",master Thesis, August 30, 2021/

[29] J.G. Proakis, "Digital Communications", McGraw Hill, 4th Edition, 2000.

[30] B. Classon, K. Blankenship and V. Desai, "Channel codng for 4G systems with adaptive modulation and codng," in IEEE Wireless Communications, vol. 9, no. 2, pp. 8-13, April 2002.

[31] L. Szczecinski, A. Alvarado and R. Feick, "Closed-form approximation of Coded BER in QAM-based BICM Faded Transmission," 2008 IEEE Sarnoff Symposium, Princeton, NJ, USA, pp. 1-5, 2008.

[32] 3GPP TR 38 901 v16.1.0," Technical Report, Study on channel model for frequencies from 0.5 to 100 GHz", November 2020.

[33] A. Martinez and A. G. i Fàbregas, "Saddlepoint approximation of random-coding bounds," 2011 Information Theory and Applications Workshop, La Jolla, CA, USA, 2011, pp. 1-6

[34] T. Erseghe, "Coding in the Finite-Blocklength Regime: Bounds Based on Laplace Integrals and Their Asymptotic Approximations," in IEEE Transactions on Information Theory, vol. 62, no. 12, pp. 6854-6883, Dec. 2016

[35] Sylvester, James J., "On the relation between the minor determinants of linearly equivalent quadratic functions". Philosophical Magazine 1, pp. 295–305, 1851.

[36] Woodbury, Max A., "Inverting modified matrices", Statistical Research Group, Memo. Rep. no. 42, Princeton University, Princeton, N. J., 1950. pp.4

[37] I. S. Gradshteyn and I. M. Ryzhik, "Table of Integrals, Series and Products" Academic Press, Seventh edition, 2007.