

Enhanced BICM Receivers for Ultra-Reliable Low-Latency Short Packet Communications

Mody Sy , *Member IEEE* and Raymond Knopp , *Fellow IEEE*

Abstract—This paper presents enhanced receiver metrics for joint estimation-detection in short blocklength transmissions, addressing scenarios with unknown channel state information and low or sparse training resource density. We show that it is possible to enhance the performance and sensitivity through block-wise joint estimation-detection compared to standard receivers. The performance analysis makes use of a full 5G transmitter and receiver chains for both Polar and LDPC coded transmissions paired with QPSK modulation scheme. We consider transmissions where reference signals are interleaved with coded data and both are transmitted over a small number of OFDM symbols so that near-perfect channel estimation cannot be achieved. Unlike conventional symbol-by-symbol detection in BICM systems, where the observation for a given coded bit is confined to the symbol in which it is conveyed, the proposed method performs block-wise joint detection over a sliding window of adjacent symbols to fundamentally leverages their statistical dependencies. Accordingly, the LLR for a particular coded bit incorporates information from all symbols within the detection window, rather than being constrained to its host symbol alone. Performance evaluation spans SIMO and SU-MIMO configurations, emphasizing the efficacy of the estimation-detection strategy in realistic base station receiver scenarios. Our findings demonstrate that when the detection windows used in the metric units are on the order of four modulated symbols, the proposed receivers remarkably outperform the conventional ones and can be used to achieve detection performance that is close to that of coherent receivers with perfect CSI.

Index Terms—Coded Modulation, 5G NR Polar code, 5G NR LDPC Code, 5G NR Physical Uplink Channels, Short Packet Communications, Unknown Channel State Information, Joint Estimation and Detection.

I. INTRODUCTION

IT IS EXPECTED that the 6G air-interface will build upon the 5G standard and address new paradigms for feedback-based cyber-physical systems combining communications and sensing. In particular, the deployment of 6G-enabled devices will necessitate ultra-tight control loops over the air interface, demanding unprecedented levels of reliability and latency potentially exceeding the sub-millisecond uplink application-layer latency targets in the microwave spectrum currently set by 5G. This critical requirement for ultra-reliable and low-latency communications fundamentally drives ongoing advancements in the physical layer, with a particular emphasis on the design of highly sophisticated receivers. Notably, reliability emerges as a cornerstone metric, underpinning receiver architectures designed to satisfy the stringent key performance

indicators (KPIs) imposed by beyond-5G (B5G) and 6G wireless communication standards.

However, although 5G transmission formats can provide very short-packet transmission through the use of mini-slots, the ratio of training information to data is not necessarily adapted to extremely short data transmission. Moreover, these formats are primarily designed under the assumption of conventional quasi-coherent receivers, which can become highly suboptimal in scenarios where accurate channel estimation is infeasible either due to the sporadic nature of such transmissions or to stringent constraints on decoding latency.

This challenge is particularly 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 for future joint-sensing and communication paradigms requiring rapid sensory feedback to the network. 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 that require extremely high levels of reliability and low latency. In this work, we investigate *bit-interleaved coded modulation* receiver design for short data transmission. Specifically, we are interested in designing joint estimation-detection based receivers compliant with polar and *low-density parity-check* (LDPC) coded modulation transmission, targeting short packets in the range of 20–100 bits for the envisaged beyond 5G/6G signaling scenarios. Indeed, BICM remains a widely adopted coded-modulation technique for error-prone wireless channels. On the receiver side, its efficacy hinges on the underlying detection and decoding metrics, underscoring the trade-off between enhanced performance and low complexity. Noteworthy is the historical integration of BICM into 3GPP systems, a practice dating back to the 3G-era.

Furthermore, there is a wealth of literature on BICM receivers from various perspectives [1] [2] demonstrating their potential impact and importance. Among the pioneers who sparked interest in BICM was the seminal work conducted by Caire *et al.* [2], wherein they provided a comprehensive analysis in terms of information rate and error probability. Afterwards, numerous research inquiries have been directed towards the design of reliable low-complexity receivers for MIMO (*multiple input multiple output*) and Non-MIMO BICM systems, but mostly restricted to coherent communication scenarios. Staring closely at the primary focus of this investigation, namely the transmission of short packets, it is evident that this area has garnered significant scholarly attention in recent years. Extensive research has addressed various aspects, including the design of signal codes and receiver algorithms [3]–[12], as well as the establishment of state-of-the-art converse and achievability bounds [13]–[17].

Manuscript prepared on October 06, 2025. The authors are with the Department of Communication Systems, EURECOM, BIOT, 06410, France (e-mails: mody.sy@eurecom.fr).

This article was in part presented at the 2023 IEEE Conference on Standards for Communications and Networking (Munich, Germany, November 2023) and the 2024 European Conference on Networks and Communications & 6G Summit (Antwerp, Belgium, June 2024).

This work distinguishes itself from prior literature by introducing novel enhanced receiver designs tailored to scenarios with imperfect *channel state information* (CSI) over various channel conditions. As previously mentioned, the proposed receivers are specifically designed to effectively support the reception of short data packets in beyond 5G/6G signaling scenarios by evaluating their performance over 5G short block channels using both Polar and LDPC coded modulation formats. (a) We look into receiver metrics exploiting *joint estimation and detection* (JED), which are particularly amenable to configurations where low-density of *demodulation reference signals* (DMRS) are interleaved with coded data symbols. (b) We specifically address situations where accurate channel estimation is impossible, demonstrating that a well-conceived joint estimation–detection receiver, leveraging interleaved DMRS within the detection metric, can achieve remarkably significant performance gains over conventional 5G *orthogonal frequency division multiplexing* (OFDM) receivers, and can potentially approach the performance of a perfect CSI receiver, applicable to both uplink and downlink transmission scenarios.

Explicitly, our proposal consists of designing and utilizing novel soft-likelihood metrics that directly integrate channel estimation performed via joint least squares followed by averaging or smoothing across DMRS dimensions for bit-level LLR generation. Moreover, we apply to the underlying soft-likelihood metrics an advanced block-wise joint detection scheme defined over a detection window of four modulated symbols ($M = 4$). Unlike conventional symbol-by-symbol detection in BICM systems, where the observation for a given coded bit is confined to the symbol in which it is conveyed, the proposed block-wise detection approach fundamentally leverages the statistical dependencies between adjacent symbols; that is, the LLR for a given coded bit incorporates information from all symbols within the detection window, rather than being constrained to its host symbol alone, thereby enhancing detection reliability. Hence, our contributions span the following principal avenues. Initially, we introduce a BICM receiver metric specifically tailored for non-coherent fading channels in *single-input multiple-output* (SIMO) transmissions. These metrics effectively address challenges arising from both line-of-sight (LOS) and *non-line-of-sight* (NLOS) fading channels. Secondly, we extend the BICM receiver metric design to *single-user* MIMO systems, specifically addressing block fading channel conditions.

The article is structured as follows. Section II lays out the system model and the foundations of 5G polar and LDPC coded modulations, Section III focuses on the proposed BICM receiver metrics, Section IV presents the numerical 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 $\mathbf{x}$, $\|\mathbf{x}\|$ denotes its Euclidean norm, $|\cdot|$ denotes the absolute value. $\|\cdot\|_F$ is the Frobenius norm of matrix. $\text{tr}\{\cdot\}$ denotes the trace of matrix. $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. $\text{Re}(\cdot)$ denotes the real part of a complex number. $I_0(\cdot)$ is the zero-th order

modified Bessel function of the first kind. $\mathbf{I}$ is an identity matrix with appropriate dimensions. $\mathbf{x} \in \mathcal{X}_b^j = \{\mathbf{x} : e_j = b\}$ is the subset of symbols $\mathbf{x}$ for which the j -th bit of the label e is equal to $b = \{0, 1\}$, $\mathcal{X}$ is the modulation alphabet (e.g., QPSK, 16-QAM, ...). The number of bits required to a symbol is denoted by $m := \log_2 |\mathcal{X}|$, where $|\mathcal{X}|$ is the cardinality of $\mathcal{X}$. $\Lambda^j(\cdot)$ denotes log likelihood ratio, with $j = 1, 2, \dots, m$. M is the number of symbols making up a block on which joint detection is performed. The superscripts $\top$ and † denote the transpose and the complex conjugate transpose or Hermitian. The operator $\uplus$ denotes a *disjoint union*: $\mathbf{d} = \uplus_{b=1}^B \mathcal{D}_b$ means that the index sets $\mathcal{D}_b$ are pairwise disjoint and their union exactly covers $\mathbf{d}$, i.e., $\mathcal{D}_b \cap \mathcal{D}_{b'} = \emptyset$ for all $b \neq b'$, and $\bigcup_{b=1}^B \mathcal{D}_b = \mathbf{d}$.

II. GENERAL FRAMEWORK

A. Bit-Interleaved Polar-coded Modulation (BIPCM)

Bit interleaved polar coded modulation is referred to as BIPCM. In this instance, we are dealing with the *cyclic redundancy check* (CRC)-aided polar coding scheme, one of the basic code construction techniques established by the 3GPP Standard [18]. Using polar codes as a channel coding scheme for 5G control channels has demonstrated the significance of Arikan's invention [19], 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. In 5G new radio, the polar codes are employed to encode *broadcast channel* (BCH) as well as *downlink control information* (DCI) and *uplink control information* (UCI). Furthermore, the transmission process is straightforward and complies with the 3GPP standard specifications [18].

With respect to the decoding process, several main polar code decoding algorithms are currently used, including the SC algorithm [19], the *successive cancellation list* (SCL) algorithm [20] [21], the *CRC-aided SCL* (CA-SCL) algorithm [22] [23], the *belief propagation* (BP) algorithm [24], and the *successive cancellation with adaptive node* (SCAN) algorithm [25]. The SCL algorithm improves upon the SC algorithm by providing multiple paths and outperforms it in terms of performance. The CA-SCL algorithm incorporates a high-rate CRC code to assist in selecting the correct codeword from the final list of paths in the SCL decoder, effectively enhancing its reliability. It has been observed that the right codeword is usually included in the list every time the SCL decoder fails. The performance ranking of the decoding algorithms is as follows: CA-SCL > state-of-the-art SCL > BP = SCAN > SC. Therefore, for improved performance, the channel decoder technique should utilize CA-SCL decoding for downlink (DCI or BCH) or uplink (UCI) messages. The adoption of polar codes by 3GPP was partly due to the well-acknowledged potential of CA-SCL decoding to outperform Turbo or LDPC codes.

B. Bit-Interleaved LDPC-coded Modulation (BILCM)

Bit-Interleaved LDPC-Coded Modulation is referred to as BILCM. First proposed by Gallager in the early 1960s [26], LDPC coding has proven to be highly suitable for 5G NR due to its advantages such as high throughput, low latency, efficient decoding complexity, and rate compatibility. The performance of LDPC codes in 5G NR is impressive, exhibiting an error floor at or below the 10^{-5} block error rate (BLER), a significant improvement over traditional coding techniques.

Furthermore, the BILCM transmission procedure is almost identical to that described with BIPCM.

A code block is encoded by the 5G LDPC encoder according to the procedure defined in the 3GPP standard [18]. At the receiver, the LDPC decoding is performed on each code block individually. For LDPC decoding, various techniques can be implemented, with belief propagation (BP) methods being the most commonly used. BP methods rely on iterative message exchange between bit nodes and check nodes, offering near-optimal decoding performance at the cost of computational complexity. However, to strike a better balance between performance and complexity, several simplified and effective decoding algorithms have been proposed in the scientific literature. One such decoding algorithm is layered message passing [27], which stands out as a promising approach for URLLC due to its ability to speed up convergence times [28], [29], making it a suitable candidate for short packet transmissions.

The foundations of polar and LDPC coding and decoding are beyond the scope of this paper, but interested readers may wish to refer to one of our prior correspondences [30].

C. Modulation and Resource Mapping

In both scenarios, the encoded payload undergoes rate-matching and code block concatenation prior to being fed into a QPSK modulator. This process yields a set of complex-valued modulation symbols. Subsequently, the resource mapping process is executed, where one or multiple OFDM symbols are used to allocate the modulated symbols to resource blocks and insert the DMRS resources. The number of resource blocks is determined by the payload size and coding settings. When the payload size is small, fewer resource blocks are required, thereby maintaining a constant effective coding rate.

Furthermore, the transmitted signal $\mathbf{x}$ typically consists of data-dependent $\mathbf{x}^{(d)}$ and data-independent $\mathbf{x}^{(p)}$ components, known as pilot or reference signals. The reference signals are used to resolve channel ambiguity across time, frequency, and/or spatial domains. Specifically, they are employed to estimate the channel. In practice, the reference signals are commonly interleaved among the data-dependent components. It is notably the case in current OFDM systems. In earlier CDMA systems, reference signals were sometimes superimposed on top of data-dependent signals. The number of data dimensions is denoted by N_d , and the number of reference signal dimensions is denoted 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 (REs) in the physical resource blocks (PRBs). The number of PRBs, K ranges from 1 to 16, while the number

of OFDM symbols, L , ranges from 1 to 14, and can be increased if multiple slots are used for signaling the channel bits. Resource elements $\mathbf{x} = \{x_n : n = 1, 2, \dots, N\}$ are mapped onto N subcarriers such that $\forall n \in \mathcal{N}_p \cup \mathcal{N}_d$ with $\mathcal{N}_p$ the set of subcarriers for DMRS and $\mathcal{N}_d$ the set of subcarriers for data. The assumption in this work is that the data-dependent components $\mathbf{x}^{(d)}$ are generated from a binary code whose output is interleaved and subsequently mapped onto an $\mathcal{M}$ -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. We denote the E interleaved coded bits as $e_k, k = 0, 1, \dots, E-1$. Adjacent $\log_2 |\mathcal{X}|$ bit-tuples are used to select the modulated symbols in the symbol alphabet. Unless otherwise stated, we assume that Gray mapping is applied when using non-binary modulation schemes. In instance of the quadrature phase-shift keying (QPSK) modulation, pairs of bits $e[2n]$ and $e[2n+1]$, are mapped to complex-valued modulation symbols as follows:

$$\mathbf{x}^{(d)}[n] = \begin{cases} \frac{1}{\sqrt{2}}(1+i), & \text{if } (e[2n], e[2n+1]) = (0, 0) \\ \frac{1}{\sqrt{2}}(-1+i), & \text{if } (e[2n], e[2n+1]) = (0, 1) \\ \frac{1}{\sqrt{2}}(-1-i), & \text{if } (e[2n], e[2n+1]) = (1, 1) \\ \frac{1}{\sqrt{2}}(1-i), & \text{if } (e[2n], e[2n+1]) = (1, 0) \end{cases} \quad \forall n \in \mathcal{N}_d. \quad (1)$$

Likewise, it is worth examining how DMRS sequences making up $\mathbf{x}^{(p)}$ are constructed. These sequences are derived from Zadoff-Chu (ZC) sequences, leveraging their desirable properties in terms of ideal auto-correlation, low PAPR (Peak-to-Average Power Ratio), constant amplitude, and their proven standardization adoption in LTE and 5G NR. In that respect, the DMRS sequence is defined as:

$$\mathbf{x}^{(p)}[n] = \begin{cases} \exp\left(-i\pi \cdot \frac{un^2}{N_p}\right) & \text{if } N_p \text{ even} \\ \exp\left(-i\pi \cdot \frac{un(n+1)}{N_p}\right) & \text{if } N_p \text{ odd} \end{cases}, \quad \forall n \in \mathcal{N}_p, \quad (2)$$

where u is the root index, which must be coprime with N_p (i.e., $\gcd(u, N_p) = 1$). The power of each DMRS sequence is typically normalized to unity. Furthermore, in spatially multiplexed MIMO systems, the resource mapping procedure must account for pilot symbol allocation not only in time and frequency but also across the spatial dimension. Specifically, training symbols must be transmitted in a way that avoids interference to ensure accurate CSI estimation. This requires careful consideration of pilot allocation strategies, which may exploit frequency orthogonality, time orthogonality, or signal orthogonality. In this study, signal orthogonality is adopted for the MIMO configuration. Indeed, in a MIMO transmission system with ($N_T \geq 1$), these antennas often share the same time-frequency resources for DMRS. To prevent inter-antenna interference, Zadoff-Chu sequences with the same root are orthogonalized across antennas using cyclic shifts (i.e., linear phase rotations). Thus we define the DMRS sequence $\mathbf{x}_t^{(p)}[n]$ transmitted via antenna port $t = 1, 2, \dots, N_T$ by

$$\mathbf{x}_t^{(p)}[n] = \mathbf{x}^{(p)}[n] \cdot \exp\left(i \cdot \frac{2\pi(t-1)n}{N_T}\right), \quad \forall n \in \mathcal{N}_p. \quad (3)$$

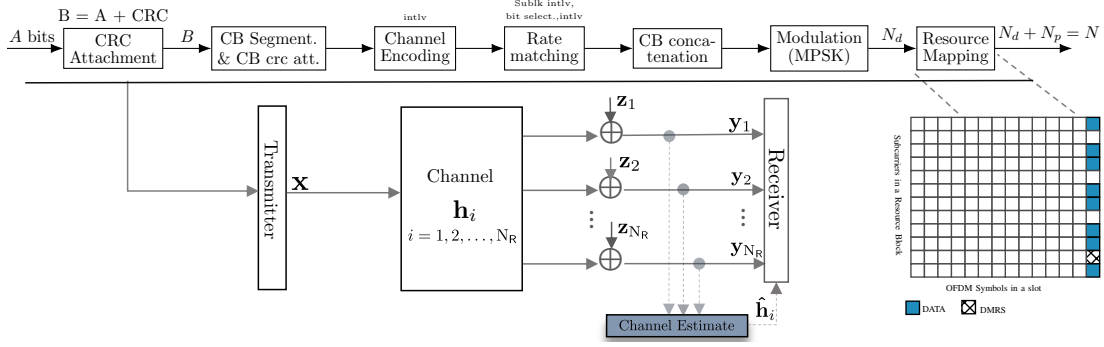


Fig. 1. Bit-Interleaved Polar/LDPC coded Modulation (SIMO BIPCM/ SIMO BILCM) : Transmitter end.

The second term $e^{i\phi_t(n)}$, where $\phi_t(n) = \frac{2\pi(t-1)n}{N_T}$, applies a linear cyclic phase shift to orthogonalize the DMRS sequences across different transmit antennas. To ensure this orthogonality, as required by the 5G standard, an additional condition must be satisfied. Indeed, the number of DMRS positions N_p must be at least equal to the number of antenna ports N_T , which means: $N_T \leq N_p$.

Figure 1 presents an overview of the BIPCM/BILCM process short block uplink channels. The transmit-end procedure includes several steps such as adding a transport block CRC, segmenting code blocks with additional CRC attachment, channel encoding, rate matching, code block concatenation, and modulation. It is important to emphasize that the receiving chain simply follows the reverse flow of the transmitter-end.

III. RECEIVER DESIGN

A. $(1 \times N_R)$ SIMO Non-Coherent Fading Channel

Considering a SIMO OFDM BICM system with a single antenna element on the transmit array ($N_T = 1$) and multiple element receive arrays (N_R). The transmitted and received signals 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:

$$\mathbf{y}_r = \mathbf{h}_r \mathbf{x} + \mathbf{z}_r, \quad r = 0, 1, \dots, N_R - 1, \quad (4)$$

where $\mathbf{z}_r$ is additive white Gaussian noise whose real and imaginary components are independent and have variance and $\mathbf{h}_r$ represents the channel vector.

1) Perfect Channel State Information

In the instance of perfect channel state information, the likelihood function is shown to be:

$$q(\mathbf{x}, \{\mathbf{y}_r, \mathbf{h}_r\}) = p(\mathbf{y}_r, \mathbf{h}_r | \mathbf{x}) = p(\mathbf{y}_r | \mathbf{x}, \mathbf{h}_r) p(\mathbf{h}_r | \mathbf{x}). \quad (5)$$

If the transmitted signal $\mathbf{x}$ is independent of the channel realization $\mathbf{h}_r$, the term $p(\mathbf{h}_r | \mathbf{x})$ in (5) can be dropped since it will disappear in (6). The likelihood function is commonly equivalent to :

$$q(\mathbf{x}, \{\mathbf{y}_r, \mathbf{h}_r\}) \propto \exp\left(-\frac{1}{N_0} \|\mathbf{y}_r - \mathbf{h}_r \mathbf{x}\|^2\right). \quad (6)$$

Using the norm extension property, ignoring terms that are independent of $\mathbf{x}$, the likelihood function then simply becomes:

$$q(\mathbf{x}, \{\mathbf{y}_r, \mathbf{h}_r\}) \propto \exp\left(\frac{2}{N_0} \text{Re}(\mathbf{y}_r \mathbf{h}_r^\dagger \mathbf{x}^\dagger) - \frac{1}{N_0} \|\mathbf{h}_r \mathbf{x}\|^2\right). \quad (7)$$

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

$$q(e_j(\mathbf{x}) = b, \{\mathbf{y}_r, \mathbf{h}_r\}) = \sum_{\mathbf{x} \in \mathcal{X}_b^j} q(\mathbf{x}, \{\mathbf{y}_r, \mathbf{h}_r\}). \quad (8)$$

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 the j -th coded bit, such that :

$$\Lambda^j(\mathbf{y}_r) = \log \frac{q(e_j(\mathbf{x}) = 0, \{\mathbf{y}_r, \mathbf{h}_r\})}{q(e_j(\mathbf{x}) = 1, \{\mathbf{y}_r, \mathbf{h}_r\})}. \quad (9)$$

We simplify (9) using a *max-log approximation*: $\log\{\sum_i \exp(\lambda_i)\} \sim \max_i \{\lambda_i\}$, resulting in (10).

$$\Lambda^j(\mathbf{y}) = \sum_{b \in \{0, 1\}} (-1)^b \left[\max_{\mathbf{x} \in \mathcal{X}_b^j} \frac{1}{N_0} \left(\sum_{r=0}^{N_R-1} 2 \text{Re}(\mathbf{y}_r \mathbf{h}_r^\dagger \mathbf{x}^\dagger) - \|\mathbf{h}_r \mathbf{x}\|^2 \right) \right]. \quad (10)$$

This metric is typically used in *Perfect CSI* based receivers as well as in conventional quasi-coherent receivers which employs least-squares channel estimation followed by linear interpolation, thereby replacing $\mathbf{h}$ with $\hat{\mathbf{h}}$. We consider these receivers as a benchmark for comparison with the subsequent proposed receivers.

Moreover, within the framework of a conventional receiver, it is presupposed that, at the very least, the observation of a single reference signal spans a PRB in order to generate the coded bits corresponding to each data symbol in that PRB. In addition, this conventional metric applies a so-called *symbol-by-symbol* detection, that's to say that each symbol is detected independently of the other symbols. We will note this as *No CSI Conv*.

2) Unknown Channel State information

Non-coherent fading channels refer to communication channels in which the fading coefficients are not known a priori and must be estimated, thereby requiring detection or estimation techniques that do not rely on explicit CSI. In what follows, we describe BICM metrics for a general non-coherent fading channel with unknown phase on the line-of-sight (LOS) components and fully unknown diffuse (Non-LOS) components. The overall unknown channel gain is given by $\mathbf{h}_r = (\sqrt{\alpha} e^{j\theta_r} + \sqrt{1-\alpha} \mathbf{h}_r^{(f)}) \mathbf{I}$, where θ_r is assumed to be i.i.d. uniform random variables on $[0, 2\pi)$, $\mathbf{h}_r^{(f)}$ is a zero-mean, unit-variance, circularly-symmetric complex Gaussian random variable and α is the relative strength of the LOS component. The amplitude $|\mathbf{h}_r|$ on each receive branch 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. [A novel soft-likelihood metric for SIMO.]

Neglecting multiplicative terms independent of the transmitted message, the likelihood function can be expressed as follows:

$$q(\mathbf{x}, \mathbf{y}) = \prod_{r=0}^{N_R-1} \frac{1}{\mathbf{L}_x} \exp \left(-\frac{\alpha \|\mathbf{x}\|^2}{\mathbf{L}_x} + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 \right) \times I_0 \left(\frac{2\sqrt{\alpha}}{\mathbf{L}_x} |\mathbf{x}^\dagger \mathbf{y}_r| \right), \quad (11)$$

where $\mathbf{L}_x = N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2$, $\beta_x = \frac{2(1-\alpha)}{N_0(N_0+2(1-\alpha)\|\mathbf{x}\|^2)}$ and $I_0(\cdot)$ is the zero-order modified Bessel function.

Proof. see Appendix Section A. ■

We then apply equations (8)–(9) to succinctly generate the LLR of the j -th coded bit.

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. [2], 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. Furthermore, 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.

Furthermore, the LLR metric calculations based on (11) in the logarithmic domain can be hard to implement and computationally prohibitive. A common simplification is the *max-log approximation*. First, an exponential approximation is applied to the modified Bessel function of the first kind $I_0(z)$, which results in $I_0(z) \sim \frac{e^z}{\sqrt{2\pi z}} \sim e^z$.

Thus the maxlog-likelihood ratio for the j -th coded bit is given by (12).

$$\Lambda^j(\mathbf{y}) = \sum_{b \in \{0,1\}} (-1)^b \left[\max_{\mathbf{x} \in \mathcal{X}_b^j} \left(\sum_{r=0}^{N_R-1} \psi(\mathbf{x}, \mathbf{y}_r) \right) - \sum_{\mathbf{x} \in \mathcal{X}_b^j} N_R \log \mathbf{L}_x \right], \quad (12)$$

with the kernel function :

$$\psi(\mathbf{x}, \mathbf{y}_r) := -\frac{\alpha \|\mathbf{x}\|^2}{\mathbf{L}_x} + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 + \frac{2\sqrt{\alpha}}{\mathbf{L}_x} |\mathbf{x}^\dagger \mathbf{y}_r|.$$

It's also worth noting that the max-log metric performs nearly as well as the accurate log-based metric since Gray-mapped constellations are in use. The max-log metric appears to have a somewhat minimal impact on receiver performance when operating with low modulation orders.

Remark 1. In equation (12), many terms can be omitted when the magnitude of $\mathbf{x}$ remains constant, as is the case for BPSK or QPSK modulations. Additionally, in the presence of strong line-of-sight (LOS) channels, the quadratic terms in equation (12) can also be skipped.

Corollary 1. Observing the structure of the metric and considering the absence of overlap between the data and the DMRS symbols, we can easily see that the channel estimate is part of the metric. By modeling $\mathbf{x}$ as a composite signal comprising both data and DMRS in an interleaved fashion within a common OFDM symbol, this yields: $\mathbf{x} = \mathbf{x}^{(p)} + \mathbf{x}^{(d)}$ where d and p are subscripts representing data, DMRS components, respectively. Hence, we can highlight $\hat{\mathbf{h}}^{\text{LS}}$ through the metric expression:

$$|\mathbf{x}^\dagger \mathbf{y}_r| = \left| \underbrace{\mathbf{x}^{(p)\dagger} \mathbf{y}_r^{(p)}}_{\text{channel estimate}} + \underbrace{\mathbf{x}^{(p)\dagger} \mathbf{y}_r^{(d)}}_0 + \underbrace{\mathbf{x}^{(d)\dagger} \mathbf{y}_r^{(p)}}_0 + \mathbf{x}^{(d)\dagger} \mathbf{y}_r^{(d)} \right| \quad (13)$$

$$= \left| \underbrace{\mathbf{x}^{(p)\dagger} \mathbf{y}_r^{(p)}}_{\text{channel estimate}} + \mathbf{x}^{(d)\dagger} \mathbf{y}_r^{(d)} \right| = \left| N_p \hat{\mathbf{h}}_r^{\text{LS}} + \mathbf{x}^{(d)\dagger} \mathbf{y}_r^{(d)} \right|,$$

where $\mathbf{x}^{(p)\dagger} \mathbf{y}_r^{(p)} = (\mathbf{x}^{(p)\dagger} \mathbf{x}^{(p)}) \hat{\mathbf{h}}_r^{\text{LS}} = \|\mathbf{x}^{(p)}\|^2 \hat{\mathbf{h}}_r^{\text{LS}} = N_p \rho \hat{\mathbf{h}}_r^{\text{LS}}$, ρ is the reference signal power and is typically normalized to unity. The channel estimate $\hat{\mathbf{h}}_r^{\text{LS}}$ is obtained via a joint *least-squares* (LS) channel estimation using averaging or smoothing over the number of reference signals within the channel coherence (namely here, the coherence bandwidth). In general, the channel estimation procedure will work as usual, and the resulting channel estimate is fed into the newly derived soft LLR metric considered in proposition 1.

3) Joint Estimation and Detection

For the case of polar or LDPC coded data, there is a convincing motivation to divide the coded streams into smaller blocks for detection for some complexity reasons. Assuming an ideal interleaving scenario with perfect CSI, detection can be performed on individual modulated symbols. However, in the presence of the unknown CSI, and considering joint estimation and detection, where interleaved DMRS and data symbols are considered, we need to deal with short blocks that contain both data or modulated symbols and DMRS symbols. Hence, the received and transmitted signal sets are subdivided into smaller block segments, forming detection windows. Phrased directly, the newly proposed metric is applied to each group of symbols, or the so-called detection window. Noting that, in this proposed detection approach, we do not restrict the observation interval of a particular coded bit to the symbol in which it is conveyed.

Proposition 2. [Block-wise joint estimation - detection.]

The proposal consists of applying an advanced joint estimation-detection per block of $M = 4$ modulated symbols forming a detection window to generate bit-level LLRs. Unlike conventional symbol-by-symbol detection in BICM systems where the observation for a particular coded bit is limited to the symbol in which it is conveyed, the proposed block-wise detection approach fundamentally leverages the statistical dependencies between adjacent symbols. In other respect, the LLR for a given coded bit incorporates information from all symbols within the detection window, rather than being constrained to its host symbol alone, thereby improving detection reliability.

To illustrate this statement, let $\mathbf{x}^{(d)} = \{\mathbf{x}_n^{(d)}\}_{n=1}^M \in \mathcal{X}^M$ be the candidate symbol vector corresponding to a block of M symbols (with $M = 4$ in our case). Let the received frame be denoted by $\mathbf{Y} \in \mathbb{C}^{N \times N_R}$, which contains both data and reference signals. We define the data-carrying positions (d) are divided into B non-overlapping blocks of size M , such that: $d = \bigcup_{b=1}^B \mathcal{D}_b$, with $|\mathcal{D}_b| = M$. For each block $b \in \{1, \dots, B\}$, we extract the corresponding *block data observation matrix* from the received signal: $\mathbf{Y}_b^{(d)} \in \mathbb{C}^{M \times N_R}$, which gathers all received data symbols at the M subcarriers in $\mathcal{D}_b$ across all N_R antennas. Each block $\mathbf{Y}_b^{(d)}$ is processed independently to compute the bit-wise likelihoods over the candidate symbol vectors. More precisely, unlike symbol-wise detection schemes, the proposed block-wise detection computes the LLRs of coded bits by evaluating joint symbol hypotheses over a detection window of size M . This involves assessing all possible candidate vectors $\mathbf{x}^{(d)} \in \mathcal{X}^M$, and partitioning them into two subsets (i.e., grouping vectors where the bit is 0 vs. 1) based on the value of the j -th coded bit. Consequently, each LLR $\Lambda^j(\cdot)$ is derived by marginalizing the block-wise likelihood function.

Considering QPSK modulation and detection windows of $M = 4$, each jointly detected symbol block has $|\mathcal{X}|^M = 256$ possible candidate symbol vectors in the search space, where each candidate $\mathbf{x}^{(d)} \in \mathcal{X}^M$ represents a distinct symbol vector. In other terms, the detector exhaustively searches over all 256 possible symbol vector hypotheses to compute the LLRs.

The proposal is conceptually detailed in Algorithm 1.

B. ($N_T \times N_R$) MIMO Rayleigh Block Fading Channel

Consider a SU-MIMO transmission model with N_T transmit and N_R receive antennas operating over a quasi-static, flat-fading channel. We assume no *inter-symbol interference* (ISI) and consider a time-invariant configuration, making it feasible to use the standard baseband complex-valued representation. Let $H_{t,r}$ represent the complex-valued path gain, serving as the fading coefficient from transmit antenna t to receive antenna r . At any given time instance, the complex-valued signal $\mathbf{x} = \{x_1, x_2, \dots, x_{N_T} \mid \mathbf{x} \in \mathbb{C}^{1 \times N_T}\}$ are transmitted through the N_T respective antennas. The received signal $\mathbf{y}$ belongs to the complex vector space $\mathbb{C}^{1 \times N_R}$. $\mathbf{z} \sim \mathbb{C}^{1 \times N_R}$ is the additive white Gaussian noise with independent real and imaginary components, each having a variance of σ^2 in every dimension,

Algorithm 1: Block-wise Joint Estimation-Detection (SIMO)

```

1 Input: Received frame:  $\mathbf{Y} \in \mathbb{C}^{N \times N_R}$ , Data and DMRS indices:
   (d), (p), Candidate list size:  $L = |\mathcal{X}|^M$ , Candidate symbol set
   tensor:  $\mathbf{X}^{(d)} \in \mathbb{X}^{M \times L}$ , Adjustable parameter:  $\alpha$ , noise power:  $N_0$ 
2 Output: LLR vector  $\Lambda(\mathbf{y}^{(d)})$ 
3 for each block  $b = 1, \dots, B$  do
4   Extract data observations over blocks:
5   Let  $\text{idx}_b = [(b-1)M + 1 : bM]$  then,
6   Set:  $\mathbf{Y}_b^{(d)} = \mathbf{Y}^{(d)}(\text{idx}_b, :)$   $\in \mathbb{C}^{M \times N_R}$ .
7   Set:  $\mathbf{y}_{b,r}^{(d)} := \mathbf{Y}_b^{(d)}(:, r) \in \mathbb{C}^{M \times 1}$ .
8   Step 1: Joint Channel Estimation
9   for each antenna  $r = 1, \dots, N_R$  do
10    Extract DMRS observations:  $\mathbf{y}_r^{(p)} = \mathbf{Y}^{(p)}(:, r) \in \mathbb{C}^{N_p \times 1}$ 
11    Compute the direct channel estimate:
12     $N_p \hat{\mathbf{h}}_r^{\text{LS}} = \mathbf{x}^{(p)\dagger} \mathbf{y}_r^{(p)}$ 
13   Step 2: Likelihood Function Processing
14   for each candidate index  $c = 1, \dots, L$  do
15    Set candidate symbol vector:  $\mathbf{x}_c^{(d)} = \mathbf{X}^{(d)}(:, c)$ 
16    Compute aggregated received signal energy:
17     $\mathcal{S}_c = \sum_{r=1}^{N_R} |N_p \hat{\mathbf{h}}_r^{\text{LS}} + \mathbf{x}_c^{(d)\dagger} \mathbf{y}_{b,r}^{(d)}|^2$ 
18    Set:  $\mathbf{L}_{x_c} = N_0 + 2(1-\alpha) \|\mathbf{x}_c^{(d)}\|^2$ ,
19    Set:  $\beta_{x_c} = \frac{2(1-\alpha)}{N_0 \cdot \mathbf{L}_{x_c}}$ 
20    Evaluate the likelihood function:  $q(\mathbf{x}_c^{(d)}, \mathbf{y}_b^{(d)}) =$ 
21     $\frac{1}{\mathbf{L}_{x_c}} \exp\left(-\frac{\alpha \|\mathbf{x}_c^{(d)}\|^2}{\mathbf{L}_{x_c}} + \beta_{x_c} \cdot \mathcal{S}_c\right) I_0\left(\frac{2\sqrt{\alpha \cdot \mathcal{S}_c}}{\mathbf{L}_{x_c}}\right)$ 
22   Step 3: Joint Detection
23   for each bit index  $j = 1, \dots, M \cdot \log_2 |\mathcal{X}|$  in the block do
24    Compute the LLR of coded bit  $e_j \in \{0, 1\}$ :
25     $\Lambda_j(\mathbf{y}_b^{(d)}) = \log \frac{\sum_{\mathbf{x}_c^{(d)} \in \mathcal{X}_0^{(d)}} q(\mathbf{x}_c^{(d)}, \mathbf{y}_b^{(d)})}{\sum_{\mathbf{x}_c^{(d)} \in \mathcal{X}_1^{(d)}} q(\mathbf{x}_c^{(d)}, \mathbf{y}_b^{(d)})}$ 
26   Step 4: Concatenate the block-wise LLR vectors
27 return  $\Lambda(\mathbf{y}^{(d)}) = [\Lambda(\mathbf{y}_1^{(d)}), \Lambda(\mathbf{y}_2^{(d)}), \dots, \Lambda(\mathbf{y}_B^{(d)})]$ 

```

and the channel matrix is denoted by $\mathbf{H} \in \mathbb{C}^{N_T \times N_R}$. The MIMO channel model is succinctly expressed as

$$\mathbf{y} = \mathbf{x}\mathbf{H} + \mathbf{z}. \quad (14)$$

However, the system model given in (14), which describes transmission during a single symbol interval, can be extended to multiple consecutive transmissions. Consider N consecutive transmitted vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, where N denotes the total number of symbol intervals.

For notational convenience, we organize these vectors into matrices as follows: $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^\top \in \mathbb{C}^{N \times N_T}$, $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^\top \in \mathbb{C}^{N \times N_R}$, $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N]^\top \in \mathbb{C}^{N \times N_R}$. The MIMO channel model can then be compactly expressed as

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \mathbf{Z}, \quad (15)$$

where $\mathbf{H}$ is assumed to be constant over the N symbol intervals or the N -symbol block and changes independently across blocks. We adopt a wide assumption regarding $\mathbf{H} \sim \mathcal{CN}(0, \mathbf{I})$, which is that its entries, $H_{t,r}$, are statistically independent for the sake of simplicity. Accordingly, the complex-valued fading coefficients $H_{t,r}$ can be treated as independent zero-mean complex Gaussian random variables with unit variance. Therefore, the MIMO channel model can be referred as

the *identically and independently distributed Rayleigh fading channel*, or more precisely, the Rayleigh block-fading. Both the fading coefficients and the noise follow complex Gaussian distributions. Thus, conditioned on the transmitted signal $\mathbf{X}$, the received signals are jointly complex Gaussian. In other terms, the received signal is zero mean $\mathbb{E}\{\mathbf{Y}|\mathbf{X}\} = 0$, circularly symmetric complex Gaussian with a $N \times N$ covariance matrix Φ_Y , concretely.

Hence, the likelihood function or conditional probability density is simply given by :

$$p(\mathbf{Y}|\mathbf{X}) := \frac{\exp(-\text{tr}\{\mathbf{Y}\Phi_Y^{-1}\mathbf{Y}^\dagger\})}{\pi^{N \times N_R} \det^{N_R}(\Phi_Y)}. \quad (16)$$

We will proceed by following the steps below to derive the detection metric. Consequently, to determine the formulation of the covariance matrix, Φ_Y , we shall invoke the subsequent theorem [31, Sec. 2, Th. 2], stating that for any $\mathbf{A}$ an arbitrary $M \times N$ complex matrix, and let $\mathbf{R} = \mathbf{A}\mathbf{W}$, where $\mathbf{W} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$, meaning that $W_1, \dots, W_M$ are independent and identically distributed with independent real and imaginary parts, then: $\Phi = \mathbb{E}[\mathbf{A}\mathbf{W}\mathbf{W}^\dagger\mathbf{A}^\dagger] = \mathbf{A}\mathbf{A}^\dagger$. Therefore, $\mathbf{R} \sim \mathcal{CN}(\mathbf{0}, \mathbf{A}\mathbf{A}^\dagger)$.

Stated directly, the covariance matrix can then be expressed as follows:

$$\begin{aligned} \Phi_Y &:= \mathbb{E}\{\mathbf{Y}\mathbf{Y}^\dagger\}, \\ &= \mathbb{E}\{(\mathbf{X}\mathbf{H} + \mathbf{Z})(\mathbf{X}\mathbf{H} + \mathbf{Z})^\dagger\} = \mathbb{E}\{\mathbf{X}\mathbf{H}\mathbf{H}^\dagger\mathbf{X}^\dagger\} + \mathbb{E}\{\mathbf{Z}\mathbf{Z}^\dagger\}, \quad (17) \\ &= \mathbf{X} \mathbb{E}\{\mathbf{H}\mathbf{H}^\dagger\} \mathbf{X}^\dagger + \mathbb{E}\{\mathbf{Z}\mathbf{Z}^\dagger\} = \mathbf{X}\mathbf{X}^\dagger + N_0\mathbf{I}. \end{aligned}$$

This expression of the covariance matrix is commonly encountered in the literature [32] [33].

Next, the determinant of Φ_Y is shown to be:

$$\det \Phi_Y = \det (N_0\mathbf{I} + \mathbf{X}\mathbf{X}^\dagger) = \prod_{n=1}^N (N_0 + \lambda_n), \quad (18)$$

where λ_n are the eigenvalues of $\mathbf{X}\mathbf{X}^\dagger$. Furthermore, the covariance matrix Φ involves the addition of two matrices, making it amenable to the application of matrix inversion lemmas such as the Sherman–Morrison–Woodbury formula, or more generally, the Woodbury matrix identity [34], to compute the inverse of the covariance matrix $\Phi_Y^{-1} = (N_0\mathbf{I} + \mathbf{X}\mathbf{X}^\dagger)^{-1}$.

$$(\mathbf{A} + \mathbf{UCV})^{-1} := \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}, \quad (19)$$

where $\mathbf{A}$, $\mathbf{U}$, $\mathbf{C}$, and $\mathbf{V}$ are matrices with comfortable dimensions: $\mathbf{A}$ is a $n \times n$ matrix, $\mathbf{C}$ is a $k \times k$ matrix, $\mathbf{U}$ is a $n \times k$ matrix, and $\mathbf{V}$ is a $k \times n$ matrix.

Saying $\mathbf{A} = N_0\mathbf{I}$, $\mathbf{C} = \mathbf{I}$, $\mathbf{U} = \mathbf{X}$, $\mathbf{V} = \mathbf{X}^\dagger$, then,

$$\begin{aligned} \Phi^{-1} &:= (\mathbf{A} + \mathbf{UCV})^{-1}, \\ &= N_0^{-1}\mathbf{I} - N_0^{-1}\mathbf{X} [N_0\mathbf{I} + \mathbf{X}^\dagger\mathbf{X}]^{-1}\mathbf{X}^\dagger, \quad (20) \\ &= N_0^{-1}\mathbf{I} - N_0^{-1}\mathbf{X}\mathbf{D}\mathbf{X}^\dagger, \quad \text{where } \mathbf{D} = [N_0\mathbf{I} + \mathbf{X}^\dagger\mathbf{X}]^{-1}. \end{aligned}$$

The likelihood function $q(\mathbf{X}, \mathbf{Y}) = p(\mathbf{Y}|\mathbf{X})$ can be stated as follows:

$$\begin{aligned} q(\mathbf{X}, \mathbf{Y}) &= \frac{1}{L_X} \exp(-\text{tr}\{\mathbf{Y}^\dagger\Phi_Y^{-1}\mathbf{Y}\}) \\ &= \frac{1}{L_X} \exp\left(-\text{tr}\left\{\mathbf{Y}^\dagger\left(\frac{1}{N_0}\mathbf{I} - \frac{1}{N_0}\mathbf{X}\mathbf{D}\mathbf{X}^\dagger\right)\mathbf{Y}\right\}\right), \quad (21) \end{aligned}$$

where $L_X = \pi^{N \times N_R} \det^{N_R} (N_0\mathbf{I} + \mathbf{X}\mathbf{X}^\dagger)$.

Ignoring the multiplicative terms independent of $\mathbf{X}$, (21) reduces to:

$$q(\mathbf{X}, \mathbf{Y}) \propto \frac{1}{L_X} \exp\left(\frac{1}{N_0} \text{tr}\left\{(\mathbf{X}^\dagger\mathbf{Y})^\dagger \mathbf{D} (\mathbf{X}^\dagger\mathbf{Y})\right\}\right). \quad (22)$$

As described in (13), we can incorporate the channel estimate into the metric to take the full merit of the JED principle. For this purpose, we simply rewrite $\mathbf{X} = \mathbf{X}^{(d)} + \mathbf{X}^{(p)}$. Then, we can reveal $\hat{\mathbf{H}}_{LS}$ in the metrics:

$$\mathbf{X}^\dagger\mathbf{Y} = \underbrace{\mathbf{X}^{(p)\dagger}\mathbf{Y}^{(p)}}_{\text{channel estimate}} + \mathbf{X}^{(d)\dagger}\mathbf{Y}^{(d)} = \mathbf{C}_p\hat{\mathbf{H}}_{LS} + \mathbf{X}^{(d)\dagger}\mathbf{Y}^{(d)}, \quad (23)$$

where $\mathbf{C}_p = \mathbf{X}^{(p)\dagger}\mathbf{X}^{(p)}$ given that $\hat{\mathbf{H}}_{LS} = \frac{\mathbf{X}^{(p)\dagger}\mathbf{Y}^{(p)}}{\mathbf{X}^{(p)\dagger}\mathbf{X}^{(p)}}$. This channel estimate is obtained via a joint least-squares (LS) channel estimation using averaging or smoothing over the number of dimensions exhibiting channel coherence.

Proposition 3. [A novel soft-likelihood metric for MIMO.]

Consistent with the above derivation steps, the proposed likelihood function for joint estimation - detection over MIMO Rayleigh block fading channel can be subsequently formulated as follows :

$$q(\mathbf{X}, \mathbf{Y}) = \frac{1}{L_X} \exp\left(\frac{1}{N_0} \text{tr}\left\{\left(\mathbf{C}_p\hat{\mathbf{H}}_{LS} + \mathbf{X}^{(d)\dagger}\mathbf{Y}^{(d)}\right)^\dagger \mathbf{D} \left(\mathbf{C}_p\hat{\mathbf{H}}_{LS} + \mathbf{X}^{(d)\dagger}\mathbf{Y}^{(d)}\right)\right\}\right). \quad (24)$$

Then, the likelihood of the coded bit e_j s.t. $b \in \{0, 1\}$ is given by

$$q(e_j(\mathbf{X}) = b, \mathbf{Y}) = \sum_{\mathbf{X} \in \mathcal{X}_b^j} q(\mathbf{X}, \mathbf{Y}). \quad (25)$$

The LLR bit metric for the j -th bit in BICM receiver is

$$\Lambda^j(\mathbf{Y}) = \log \frac{q(e_j(\mathbf{X}) = 0, \mathbf{Y})}{q(e_j(\mathbf{X}) = 1, \mathbf{Y})}. \quad (26)$$

To ease the process of implementing such a LLR bit metric in (26), one may use its *max-log approximation* version given in (27).

$$\Lambda^j(\mathbf{Y}) = \sum_{b \in \{0,1\}} (-1)^b \left[\max_{\mathbf{X} \in \mathcal{X}_b^j} \Gamma(\mathbf{X}, \mathbf{Y}) - \sum_{\mathbf{X} \in \mathcal{X}_b^j} \log L_X \right]. \quad (27)$$

with the score function or max-log decision metric

$$\Gamma(\mathbf{X}, \mathbf{Y}) := \frac{1}{N_0} \text{tr}\left\{\left(\mathbf{C}_p\hat{\mathbf{H}}_{LS} + \mathbf{X}^{(d)\dagger}\mathbf{Y}^{(d)}\right)^\dagger \mathbf{D} \left(\mathbf{C}_p\hat{\mathbf{H}}_{LS} + \mathbf{X}^{(d)\dagger}\mathbf{Y}^{(d)}\right)\right\}.$$

Remark 2. For a MIMO system with N_T transmit antennas employing block-wise detection (block size M) and QPSK modulation ($|\mathcal{X}| = 4$), each antenna's symbol vector is selected from $|\mathcal{X}|^M$ possible candidates. The joint search space across all N_T antennas then expands to $(|\mathcal{X}|^M)^{N_T}$ possible combinations - computationally expensive but optimal (ML joint detection). However, under the assumption of no spatial correlation, the search space can be processed independently for each of the N_T streams. In this case, each transmit stream is detected independently without inter-antenna joint processing. This stream-wise implementation reduces the computational complexity from exponential to linear scaling with respect to N_T , requiring only $N_T \times |\mathcal{X}|^M$ total candidates. This

per-stream block-wise joint detection approach demonstrates somewhat significantly reduced complexity compared to full joint detection (i.e., joint-stream block-wise detection).

Algorithm 2 outlines this stream-wise implementation of block-wise JED for spatially multiplexed MIMO systems.

Algorithm 2: Stream-wise Block-wise Joint Detection (MIMO)

1 **Input:** Received frame: $\mathbf{Y} \in \mathbb{C}^{N \times N_R}$, Data and DMRS indices: $(d), (p)$, Candidate list size per stream: $L = |\mathcal{X}|^M$, Candidate symbol tensor: $\mathbf{X}^{(d)} \in \mathcal{X}^{L \times M \times N_T}$, where each candidate block $\mathbf{X}^{(d)}(c, :, t) \in \mathcal{X}^M$, Noise power: N_0

2 **Output:** LLR vector for all coded bits: $\Lambda(\mathbf{Y}^{(d)})$

3 **begin**

4 *Step 1: Joint Channel Estimation (LS)*

5 Extract DMRS-based LS estimate for each transmit antenna $t = 1, \dots, N_T$:

6 $\hat{\mathbf{H}}_{LS}(t, :) = \frac{[\mathbf{X}^{(p)}(:, t)]^\dagger \mathbf{Y}^{(p)}}{[\mathbf{X}^{(p)}(:, t)]^\dagger [\mathbf{X}^{(p)}(:, t)]}, \quad \hat{\mathbf{H}}_{CP}(t, :) = [\mathbf{X}^{(p)}(:, t)]^\dagger \mathbf{Y}^{(p)}$

7 Compute MMSE equalizer once:

8 $\mathbf{W} = (\hat{\mathbf{H}}_{LS}^\dagger \hat{\mathbf{H}}_{LS} + N_0 \mathbf{I}_{N_T})^{-1} \hat{\mathbf{H}}_{LS}^\dagger$

9 *Step 2: Stream-wise Block Processing*

10 **for each transmit stream** $t = 1, \dots, N_T$ **do**

11 Extract equalizer vector $\mathbf{w}_t = [\mathbf{W}(t, :)]^\dagger$,

12 and compute effective noise: $N_{0, \text{eff}}^{(t)} = N_0 \|\mathbf{w}_t\|^2$.

13 *For each block* $b = 1, \dots, B$:

14 Compute block-index range and extract data block:

15 Set $\text{idx}_b = [(b-1)M + 1 : bM]$, then $\mathbf{Y}_b^{(d)} = \mathbf{Y}^{(d)}(\text{idx}_b, :)$.

16 Equalize data block: $\tilde{\mathbf{X}}_b^{(d)} = \mathbf{Y}_b^{(d)} \mathbf{W}$ (size: $M \times N_T$).

17 Extract stream- t component and form intermediate observation: $\mathbf{Y}_{b,t}^{(d)} = \tilde{\mathbf{X}}_b^{(d)}(:, t) [\hat{\mathbf{H}}_{LS}(t, :)] \in \mathbb{C}^{M \times N_R}$.

18 *Step 2.1: Likelihood Function Evaluation*

19 **for each candidate index** $c = 1, \dots, L$ **do**

20 Let $\mathbf{X}_c^{(d)}(:, t) \triangleq \mathbf{X}^{(d)}(c, :, t) \in \mathcal{X}^{M \times 1}$.

21 Compute: $\mathbf{A}(t, :) = \hat{\mathbf{H}}_{CP}(t, :) + [\mathbf{X}_c^{(d)}(:, t)]^\dagger \mathbf{Y}_{b,t}^{(d)}$.

22 Compute $\mathbf{D}(t, t) = (N_{0, \text{eff}}^{(t)} + [\mathbf{X}_c^{(d)}(:, t)]^\dagger [\mathbf{X}_c^{(d)}(:, t)])^{-1}$.

23 Compute the metric core function and Stabilize:

24 $\Gamma(\mathbf{X}_c^{(d)}(:, t), \mathbf{Y}_{b,t}^{(d)}) = \frac{1}{N_0} \sum_{r=1}^{N_R} \mathbf{A}^\dagger(t, r) \mathbf{D}(t, t) \mathbf{A}(t, r)$.

25 Evaluate likelihood metric:

26 $q(\mathbf{X}_c^{(d)}(:, t), \mathbf{Y}_{b,t}^{(d)}) = \frac{\exp(\Gamma(\mathbf{X}_c^{(d)}(:, t), \mathbf{Y}_{b,t}^{(d)}))}{\pi^{MN_R} \det(N_{0, \text{eff}}^{(t)} \mathbf{I}_M + [\mathbf{X}_c^{(d)}(:, t)] [\mathbf{X}_c^{(d)}(:, t)]^\dagger)^{N_R}}$.

27 *Step 2.2: LLR Computation*

28 **for each bit index** $j = 1, \dots, M \log_2 |\mathcal{X}|$ **in block** b **do**

29 Compute the LLR of coded bit $e_j \in \{0, 1\}$:

30 $\Lambda_j^{(t,b)} = \log \frac{\sum_{\mathbf{X}_c^{(d)}(:, t) \in \mathcal{X}_0^j} q(\mathbf{X}_c^{(d)}(:, t), \mathbf{Y}_{b,t}^{(d)})}{\sum_{\mathbf{X}_c^{(d)}(:, t) \in \mathcal{X}_1^j} q(\mathbf{X}_c^{(d)}(:, t), \mathbf{Y}_{b,t}^{(d)})}$.

31 Store $\{\Lambda_j^{(t,b)}\}$ as the LLR vector for stream t , block b .

32 *Step 3: Concatenate All LLRs*

33 $\Lambda(\mathbf{Y}^{(d)}) = \underbrace{[\Lambda^{(1,1)}, \dots, \Lambda^{(1,B)}]}_{\Lambda^{(1)}}, \underbrace{[\Lambda^{(2,1)}, \dots, \Lambda^{(2,B)}]}_{\Lambda^{(2)}}, \dots, \underbrace{[\Lambda^{(N_T,1)}, \dots, \Lambda^{(N_T,B)}]}_{\Lambda^{(N_T)}}$

34 **return** final LLR vector $\Lambda(\mathbf{Y}^{(d)})$

IV. NUMERICAL RESULTS

For illustrative purposes, we consider two distinct configurations: (1×4) SIMO BICM, and (4×4) MIMO BICM. The subsequent figures illustrate the performance of BIPCM/BILCM with joint estimation and detection using the proposed block-wise joint estimation-detection metrics under various channel conditions. The evaluation highlights performance differences across three scenarios: *Perfect CSI* or ideal or coherent receiver, *No CSI* JED using the block-wise ($M = 4$) joint detection approach, and *No CSI* Conv, using conventional symbol-by-symbol detection.

The simulations employ NR POLAR and NR LDPC coding schemes paired with QPSK modulation. The transmission involves a transport block of 48 bits. The resource allocation procedure uses a single OFDM symbol with 48 resource elements spread over 4 PRBs (comprising 32 REs for data component and fewer REs for DMRS component).

A. Performance Analysis

1) SIMO Non Coherent Fading channel

The results illustrated in Figure 2a Figure 2c delineate the performance of the (1×4) SIMO BICM systems, with joint estimation and detection, over LOS channel ($\alpha = 1$) and rayleigh fading channel ($\alpha = 0$), respectively. This evaluation is aiming to discern the performance disparity between the *Perfect CSI* and *No CSI* conditions, particularly in coverage scenarios characterized by low signal-to-noise ratios. Moreover, we explore the most challenging situation where the DMRS density is very low, in order to show how robust the proposed receiver metrics can be compared to the conventional ones. To do this, we consider a single DMRS symbol across the entire coherence bandwidth or, at the very least, across a PRB. In this way, two structures of the conventional receiver are considered: (*No CSI Conv.*, *LS CE + Interpol.*), where least squares channel estimation is followed by linear interpolation, and (*No CSI Conv.*, *LS CE + Avg.*), where channel estimation is performed by averaging over the DMRS symbols for a fairer comparison with the proposed JED receiver.

Figure 2a illustrates the performance over a pure LOS channel ($\alpha = 1$) under 5G polar coded modulation. The JED-based receiver achieves a performance gain of 3.75 dB and 0.75 dB compared to the conventional receivers at a BLER threshold of 1%. However, a gap of approximately 0.5 dB remains between the JED receiver (with $M = 4$) and the perfect CSI receiver.

Similarly, Figure 2b shows the performance under 5G LDPC coded modulation. The JED-based receiver provides a performance gain of 3.5 dB over the conventional receiver (*No CSI Conv.*, *LS CE + Interpolation*) at the same BLER threshold. To further highlight the advantage of the JED receiver over the conventional receiver, Figure 2b depicts performance over a pure Rayleigh fading channel ($\alpha = 0$) under 5G polar coded modulation. Here, the JED-based receiver ($M = 4$) achieves a gain of 2.75 dB with respect to the conventional receiver (*No CSI Conv.*, *LS CE + Interpolation*).

In addition, the Figures also includes finite blocklength bounds, integrating both converse and achievability results for a comprehensive comparative analysis. References are made to the scholarly works [14], [17] detailing the *metaconverse* (MC) and *Random Coding Union* (RCU) bounds employed. Indeed, The RCU provides an achievability bound and gives a practical estimate of how well a communication system can perform at finite blocklength. It estimates the error probability of a random code at a given blocklength, rate, and SNR. Conversely, the meta-converse bound establishes a lower bound on the minimum achievable error probability for any coding scheme. In other words, it complements the finite blocklength achievability bound by setting a benchmark for the best possible performance, ensuring that practical systems do

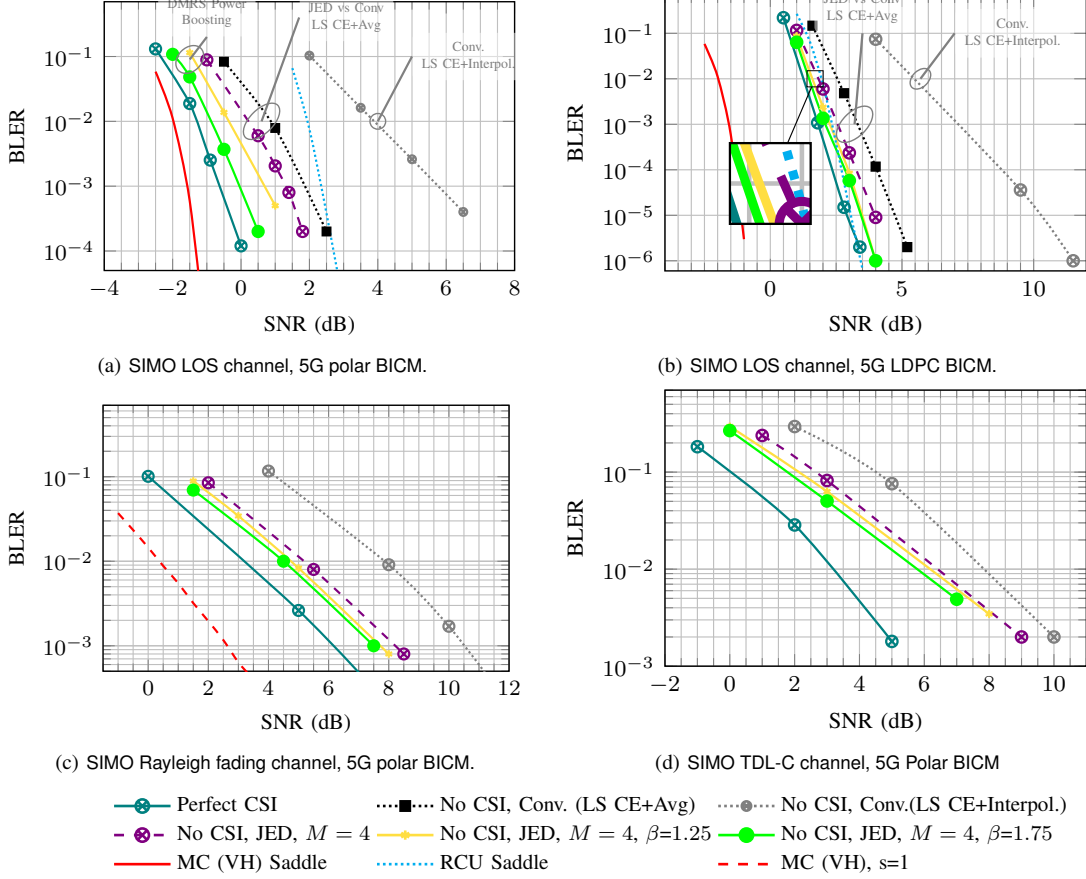


Fig. 2. Block Error Rate, 48 bits(TBs+CRC), NR POLAR BICM (CRC-aided successive-cancellation list decoder, List length=8), NR LDPC BICM (belief propagation decoder, iteration=30) QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs (32 data, 4 DMRS), (1×4) SIMO, vs outer (MC) and inner (RCUs) bounds, DMRS power Boosting via a scaling factor β .

not overestimate their capabilities, which justifies its inclusion in the graphs.

As shown in Figure 2a, the Gallager parameter s is numerically optimized (i.e., $s \rightarrow s^*$) to maximize the RCU bound and minimize the meta-converse Verdú-Han bound (i.e., the tightest bound), where s^* corresponds to the saddle point. Conversely, in Figure 2c, no optimization of the Gallager parameter s is performed; instead, a fixed value of $s = 1$ is chosen, resulting in a more relaxed constraint.

Next, Figure 2d compares the performances under a TDL-C Non-Line-of-Sight (Long Delay Spread = 300 ns, urban macro, Sampling rate $f_s = 30.72$ MHz) channel configuration with 4 receive antennas. The antenna ports were subjected to independent and identically distributed realizations, with no correlation modeling applied. Indeed, the TDL-C, is a 3GPP reference channel model characterized by its long delay-spread and emphasis for non-MIMO evaluation scenarios [35]. It models a multipath-rich environment where no dominant direct path (LOS) is explicitly present. Specifically, the TDL channel, being purely NLOS, behaves like a multi-tap Rayleigh channel where all taps are statistically distributed as independent and random Rayleigh fading. Consequently, we implement the proposed metric by setting $\alpha = 0$ in our simulations. Since the TDL-C channel is frequency selective, channel estimation using the least squares method followed by interpolation is typically required to better track channel fluctuations. How-

ever, it is important to recall that for the JED receiver under consideration, no interpolation is performed. Instead, a least-squares (LS) estimation is conducted, followed by averaging over the dimension of DMRS symbols, before incorporating the resulting channel estimate into the soft detection metric.

Furthermore, Figure 3 displays LLR distribution characteristics shown via histograms at SNR = 0 dB for the *No CSI* Conv., the proposed *No CSI* JED ($M = 4$), and the *Perfect CSI* receivers. It should be noted that LLR values represent the confidence level in binary decisions (bit = 0 or 1) after demodulation. An LLR value ≈ 0 indicates significant uncertainty or ambiguity in the binary decision, whereas values $\gg 0$ or $\ll 0$ correspond to high-confidence decisions (bit = 1 or 0, respectively). The histogram shapes thus reflect the quality of information provided by each receiver. The *No CSI* JED ($M = 4$) receiver exhibits a bimodal histogram with peaks near -3 and $+3$, showing well-separated LLR values that correspond to more confident and accurate decisions. This strongly suggests effective receiver operation without explicit CSI, further supported by the close similarity between its histogram and that of the ideal receiver. The discrepancy between the ideal receiver and the proposed one can be explained by the fact that the simulations are conducted under low DMRS density scenarios (i.e., 1 DMRS per PRB). In other respect, only fewer resource elements are allocated to DMRSs. This helps to some extent reduce the additional transmission

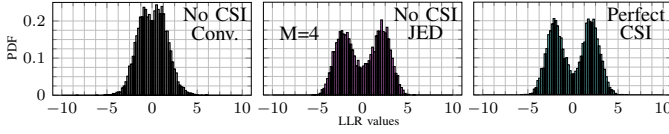


Fig. 3. LLR distribution characteristics via histograms at SNR = 0 dB.

overhead inherent in dense DMRS-assisted schemes.

Lastly, To fully exploit the performance potential of the JED receiver, we explore extra enhancement strategies. There are two possibilities: (a) DMRS densification, and (b) DMRS power boosting. If a certain sweet spot in terms of DMRS density per PRB cannot be found, increasing DMRS density inherently introduces additional transmission overhead. Therefore, the ideal approach is to prioritize transmission with low DMRS density just to bootstrap the JED receiver, followed by DMRS power boosting. Indeed, DMRS power boosting was extensively discussed in our prior correspondence [5], particularly in scenarios where reference and data symbols are jointly conveyed in common OFDM symbols. Conceptually, envision the signal as comprising a data component and a data-independent component, or pilots, in a frequency-interleaved fashion. To enhance the power of pilot signals within an interleaved set, scaling the power of DMRSs while keeping the data signals unchanged or constant is crucial. Put simply, the boosted transmitted signal, denoted as $\mathbf{x}_{\text{boosted}}$, is then defined as $\mathbf{x}_{\text{boosted}} = \mathbf{x}^{(d)} + \beta \mathbf{x}^{(p)}$. The adaptive power adjustment procedure is contingent on β values and aims to increase the power or strength of the pilot signals within the composite signal. Care should be taken to select an appropriate value for β to achieve the desired power augmentation without introducing distortion or signal saturation. To comply with potential radio frequency constraints, β must be perfectly calibrated. Consequently, optimal performance enhancement is achieved when β is set to 1.75 (corresponding to a 75% increase in DMRS power w.r.t. its initial value). For instance, Figure 2a shows that the proposed JED-based receiver delivers an additional gain of approximately 1 dB, approaching the performance bound of the ideal receiver. Overall, the implications of slightly adjusting the DMRS power within the 3GPP standard are significant. Specifically, it is feasible to allow the *user equipment* (UE) to adjust the power allocation between the DMRS and data transmission. This flexibility in adaptive DMRS power adjustment is somewhat transparent to the receiver.

2) MIMO Rayleigh Block Fading Channel

In previous simulations, we considered a low-density DMRS scenario with 1 DMRS per PRB. However, for MIMO cases, we increased the DMRS density to 4 per PRB. Indeed, channel estimation quality depends on the total number of DMRS resources per antenna stream not just on having orthogonal sequences. For a (4×4) MIMO configuration, 1 DMRS per antenna (stream) is insufficient to properly estimate all four channels. Multiple DMRS (e.g., 4) per antenna are required to ensure stable channel estimation. While a single orthogonal DMRS sequence can uniquely identify each antenna's channel, it does not provide reliable estimation, as it remains sensitive to channel variations, noise, and interference

Figure 5a illustrates the performance over a (4×4) MIMO Rayleigh block fading channel under 5G polar-coded modulation. The JED-based receiver achieves a performance gain of approximately 2.75 dB compared to the conventional receiver at a BLER threshold of 0.1%. However, there remains a performance gap of about 1.75 dB between the JED receiver ($M = 4$) and the ideal or *Perfect CSI* receiver

Similarly, Figure 5b shows the performance under 5G LDPC-coded modulation. The JED-based receiver provides a gain of approximately 2 dB over the conventional receiver, and the gap between the JED receiver ($M = 4$) and the perfect CSI receiver is around 1.75 dB.

Therefore, it can be remarkably asserted that the advanced JED-based receiver outperforms the conventional counterpart and demonstrates greater resilience under imperfect channel estimation.

B. Complexity Analysis

The complexity of the detection metrics is analysed using *Monte Carlo* simulation. The execution time highlights the time elapsed between the input and output of the demodulator, concisely, until the LLRs are generated. It is relevant to pinpoint the block size range wherein complexity is relatively low compared to conventional metrics in order to establish a better trade-off between performance and complexity. Analytically, within a very short block regime (i.e, small input symbol size), the conventional and JED receiver metrics will almost close. Considering the JED Metrics proposed in (11), we observe a quasi-linear complexity as a function of the input symbol size applied at the demodulator level, namely here $n = \{12, 24, 32, 48, 100\}$. Based on the average time complexity curves on Figure 4, the complexity of the JED ($M = 4$) receiver is of the form $\forall \epsilon \in (0, 1), \exists a \in \mathbb{R}_{>0}$ tel que $T(n) = an^{1+\epsilon} = \mathcal{O}(n^{1+\epsilon})$.

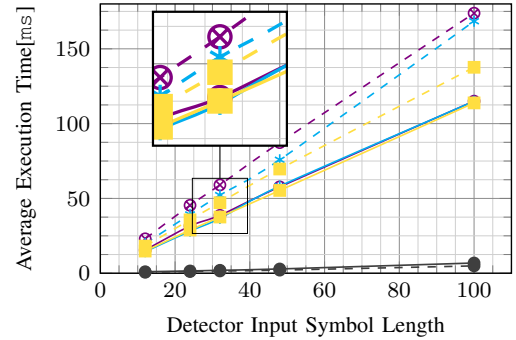
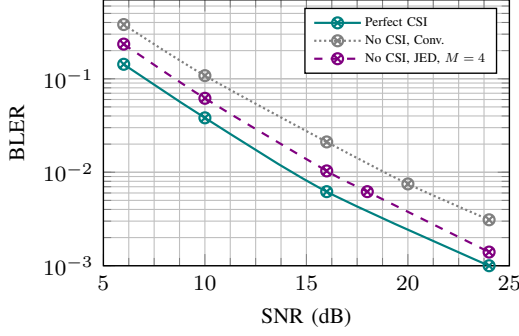
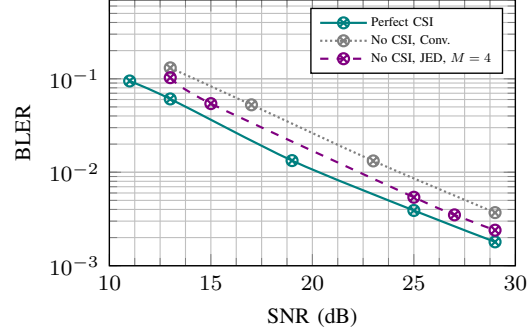


Fig. 4. Average Time Complexity of the proposed SIMO JED metrics vs Conventional metrics.

At $n = 32$ for example, as in our simulations, the conventional receiver metric is around $39\times$ faster in terms of detection than the JED ($M = 4$) metric over pure LOS condition ($\alpha = 0$) in log domain versus $25\times$ against its max-log domain version. This makes the JED metric in max-log domain about $1.5\times$ faster than the log-domain version. It should also be emphasized that by also increasing M (e.g., $M = 8$), we'll expect the complexity to escalate, making the JED approach difficult to handle.



(a) MIMO Rayleigh Block Fading channel, 5G Polar BICM.



(b) MIMO Rayleigh block Fading channel, 5G LDPC BICM.

Fig. 5. Block Error Rate, 48 bits(TBs+CRC), NR POLAR BICM (CRC-aided successive-cancellation list decoder, List length=8), NR LDPC BICM (belief propagation decoder, iteration=30) QPSK modulation, 1 OFDM symbol, 4 PRBs, 48 REs (32 data, 16 DMRS), (4×4) MIMO.

C. Current Limitations and Future Work

This proposed joint detection approach has inherent limitations. It becomes computationally prohibitive for higher-order modulation schemes like 16-QAM (where $|\mathcal{X}| = 2^4$). With a joint detection block size of $M = 4$, this would require processing $|\mathcal{X}|^M = 65,536$ possible symbol vector combinations, which is computationally intractable. In such scenarios, reducing the detection block size (e.g., to $M = 2$) becomes necessary to maintain feasible complexity.

V. CONCLUSIONS

This paper has introduced enhanced BICM receiver metrics for joint estimation–detection in short blocklength transmissions, specifically targeting scenarios with unknown channel state information and limited or sparse training resources. Through block-wise joint estimation–detection, we demonstrate significant improvements in performance and sensitivity compared to conventional receivers. Our analysis, conducted using full 5G transmitter and receiver chains with both Polar and LDPC coded transmissions under QPSK modulation, shows that even when reference signals are interleaved with coded data over a small number of OFDM symbols, precluding near-perfect channel estimation. Unlike conventional symbol-by-symbol detection in BICM systems, where the observation for a given coded bit is confined to the symbol in which it is conveyed, the proposed method performs block-wise joint detection over a sliding window of adjacent symbols to fundamentally leverages their statistical dependencies; that is, the LLR for each coded bit incorporates information from multiple symbols rather than being confined to its host symbol. Performance evaluation in SIMO and SU-MIMO configurations confirms the effectiveness of this strategy in realistic base station receiver scenarios. Notably, detection windows spanning approximately four modulated symbols allow the proposed receivers to substantially outperform conventional ones, achieving detection performance approaching that of coherent receivers with perfect CSI. Overall, the presented joint estimation–detection strategy built upon an enhanced likelihood metrics offers a significant advancement in receiver design for 5G/6G, especially in highly dynamic or resource-constrained environments.

APPENDIX

A. Metric derivation for non-coherent fading channel

Assuming that θ_r is unknown and randomly distributed over $[0, 2\pi)$, the conditional probability density function is defined as

$$q(\mathbf{x}, \mathbf{y}_r) := \frac{1}{2\pi \det \Phi} \int_0^{2\pi} \exp \left(-\frac{1}{2} (\mathbf{y}_r - \mu\{\mathbf{x}, \theta_r\})^\dagger \Phi^{-1} (\mathbf{y}_r - \mu\{\mathbf{x}, \theta_r\}) d\theta_r. \quad (28)$$

Saying $\mu\{\mathbf{x}, \theta_r\} := \sqrt{\alpha} e^{j\theta_r} \mathbf{x}$, then

$$q(\mathbf{x}, \mathbf{y}_r) = \frac{1}{2\pi \det \Phi} \int_0^{2\pi} \exp \left(-\frac{1}{2} (\mathbf{y}_r - \sqrt{\alpha} e^{j\theta_r} \mathbf{x})^\dagger \Phi^{-1} (\mathbf{y}_r - \sqrt{\alpha} e^{j\theta_r} \mathbf{x}) d\theta_r. \quad (29)$$

Covariance Matrix:

Knowing that $\mathbf{y}_r - \sqrt{\alpha} e^{j\theta_r} \mathbf{x} = \sqrt{1 - \alpha} \mathbf{h}_{i,f} \mathbf{x} + \mathbf{z}_i$, then

$$\begin{aligned} \Phi &= \frac{1}{2} \mathbb{E} \left[\left(\sqrt{1 - \alpha} \mathbf{h}_r^{(f)} \mathbf{x} + \mathbf{z}_i \right) \left(\sqrt{1 - \alpha} \mathbf{h}_r^{(f)} \mathbf{x} + \mathbf{z}_i \right)^\dagger \right], \\ &= (1 - \alpha) \mathbf{x} \mathbf{x}^\dagger \sigma_h^2 + \sigma_z^2 \mathbf{I}, \text{ where } \sigma_h^2 = 1, \end{aligned} \quad (30)$$

Determinant:

$$\begin{aligned} \det \Phi &= \det ((1 - \alpha) \mathbf{x} \mathbf{x}^\dagger + \sigma_z^2 \mathbf{I}), \\ &= \det (\sigma_z^2 \mathbf{I} + (1 - \alpha) \mathbf{x} \mathbf{x}^\dagger), \end{aligned} \quad (31)$$

Lemma 1. [36, *Sylvester's determinant identity*].

when we have a matrix of the form: $\mathbf{A} = \mathbf{B} + \mathbf{u} \mathbf{v}^\dagger$, where $\mathbf{u}$ and $\mathbf{v}^\dagger$ are column vectors, we can apply the determinant formula for a matrix perturbed by a rank-1 term (i.e., the *rank-1 matrix* determinant formula or Sylvester's determinant Lemma), where $\mathbf{B}$ is an invertible square matrix of dimension $N \times N$, $\mathbf{u}$ and $\mathbf{v}^\dagger$ are complex column vectors of dimension $N \times 1$, and $\mathbf{u} \mathbf{v}^\dagger$ is thus a rank-1 matrix. The determinant of $\mathbf{A}$ is given by:

$$\det(\mathbf{A}) := \det(\mathbf{B}) \cdot (1 + \mathbf{v}^\dagger \mathbf{B}^{-1} \mathbf{u}). \quad (32)$$

Regarding the Sylvester's determinant Lemma, (31) then becomes,

$$\begin{aligned}
\det \Phi &= \det (\sigma_z^2 \mathbf{I} + (1 - \alpha) \mathbf{x} \mathbf{x}^\dagger) = \sigma_z^{2N} \left(1 + \frac{(1 - \alpha) \mathbf{x}^\dagger \mathbf{x}}{\sigma_z^2} \right), \\
&= \sigma_z^{2N} \left(1 + \frac{(1 - \alpha) \|\mathbf{x}\|^2}{\sigma_z^2} \right), \\
&= \left(\frac{N_0}{2} \right)^N \left(1 + \frac{(1 - \alpha) \|\mathbf{x}\|^2}{\frac{N_0}{2}} \right), \\
&= \left(\frac{N_0}{2} \right)^{N-1} \left(\frac{N_0}{2} + (1 - \alpha) \|\mathbf{x}\|^2 \right), \\
&= \frac{1}{2} \left(\frac{N_0}{2} \right)^{N-1} (N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2), \\
&= c (N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2), \text{ where } c = \frac{1}{2} \left(\frac{N_0}{2} \right)^{N-1}, \\
&\propto N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2.
\end{aligned} \tag{33}$$

In what follows, we will pose $\mathbf{L}_x = N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2$.

Inverse of Φ :

The covariance matrix Φ involves the addition of two matrices, which is amenable to consider the use of matrix inversion lemmas, Sherman-Morrison-Woodbury formula, or simply the Woodbury Matrix identity.

Lemma 2. [34, *The Woodbury Matrix identity*].

$$(\mathbf{A} + \mathbf{UCV})^{-1} := \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{C}^{-1} + \mathbf{VA}^{-1} \mathbf{U})^{-1} \mathbf{VA}^{-1}, \tag{34}$$

where $\mathbf{A}$, $\mathbf{U}$, $\mathbf{C}$, and $\mathbf{V}$ are matrices with comfortable dimensions: $\mathbf{A}$ is a $n \times n$ matrix, $\mathbf{C}$ is a $k \times k$ matrix, $\mathbf{U}$ is a $n \times k$ matrix, and $\mathbf{V}$ is a $k \times n$ matrix.

Note that, here, we have a special case where $\mathbf{V}$, $\mathbf{U}$ are vectors, consequently $\text{rank}\{\mathbf{x}^\dagger \mathbf{x}\} = 1$.

$$\text{Saying : } \mathbf{A} = \sigma_z^2 \mathbf{I} \quad \mathbf{C} = (1 - \alpha) \mathbf{I} \quad \mathbf{U} = \mathbf{x} \quad \mathbf{V} = \mathbf{x}^\dagger. \tag{35}$$

$$\begin{aligned}
\Phi^{-1} &:= (\mathbf{A} + \mathbf{UCV})^{-1}, \\
&= \frac{2}{N_0} \mathbf{I} - \frac{2}{N_0} \mathbf{x} \left(\frac{2(1 - \alpha)}{N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2} \right) \mathbf{x}^\dagger.
\end{aligned} \tag{36}$$

Saying $\beta_x = \frac{2(1 - \alpha)}{N_0(N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2)}$, then $\Phi^{-1} = \frac{2}{N_0} \mathbf{I} - 2\beta_x \mathbf{x} \mathbf{x}^\dagger$.

$$\begin{aligned}
&\textbf{Likelihood function :} \\
q(\mathbf{x}, \mathbf{y}_r) &= \frac{1}{2\pi \det \Phi} \int_0^{2\pi} \exp \left(-\frac{1}{2} (\mathbf{y}_r - \sqrt{\alpha} e^{j\theta_r} \mathbf{x})^\dagger \left(\frac{2}{N_0} \mathbf{I} - 2\beta_x \mathbf{x} \mathbf{x}^\dagger \right) (\mathbf{y}_r - \sqrt{\alpha} e^{j\theta_r} \mathbf{x}) \right) d\theta_r, \\
&= \frac{1}{2\pi \det \Phi} \int_0^{2\pi} \exp \left(-\frac{1}{N_0} \|\mathbf{y}_r - \sqrt{\alpha} e^{j\theta_r} \mathbf{x}\|^2 + \beta_x \left\| \left(\mathbf{y}_r^\dagger - \sqrt{\alpha} e^{-j\theta_r} \mathbf{x}^\dagger \right) \mathbf{x} \right\|^2 \right) d\theta_r.
\end{aligned} \tag{37}$$

Extending the terms into the exponential, ignoring those that are independent of $\mathbf{x}$, the likelihood function is equivalent to

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}_r) &= \frac{1}{2\pi \det \Phi} \exp \left(-\alpha \|\mathbf{x}\|^2 \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) \right. \\
&\quad \left. + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 \right) \int_0^{2\pi} \exp \left(2\sqrt{\alpha} \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) |\mathbf{x}^\dagger \mathbf{y}_r| \cos(\phi_i + \theta_r) \right) d\theta_r.
\end{aligned} \tag{38}$$

knowing that $\frac{1}{\pi} \int_{\varphi=0}^{\pi} \exp(z \cos(\varphi)) d\varphi = I_0(z)$ [37].

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}_r) &= \frac{1}{2\pi \det \Phi} \exp \left(-\alpha \|\mathbf{x}\|^2 \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) \right. \\
&\quad \left. + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 \right) \times 2\pi \times I_0 \left(2\sqrt{\alpha} \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) |\mathbf{x}^\dagger \mathbf{y}_r| \right), \\
&= \frac{1}{\det \Phi} \exp \left(-\alpha \|\mathbf{x}\|^2 \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 \right) \\
&\quad \times I_0 \left(2\sqrt{\alpha} \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) |\mathbf{x}^\dagger \mathbf{y}_r| \right).
\end{aligned} \tag{39}$$

Then after ignoring multiplicative term that are independent of $\mathbf{x}$, it comes

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}_r) &\propto \frac{1}{\mathbf{L}_x} \exp \left(-\alpha \|\mathbf{x}\|^2 \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) \right. \\
&\quad \left. + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 \right) \times I_0 \left(2\sqrt{\alpha} \left(\frac{1}{N_0} - \beta_x \|\mathbf{x}\|^2 \right) |\mathbf{x}^\dagger \mathbf{y}_r| \right).
\end{aligned} \tag{40}$$

Expressing β_x w.r.t. $\mathbf{L}_x$, we have the relation

$$\begin{aligned}
\beta_x &= \frac{2(1 - \alpha)}{N_0(N_0 + 2(1 - \alpha) \|\mathbf{x}\|^2)} = \frac{2(1 - \alpha)}{N_0 \mathbf{L}_x} = \frac{\mathbf{L}_x - N_0}{\|\mathbf{x}\|^2 N_0 \mathbf{L}_x}, \\
&= \frac{\mathbf{L}_x}{\|\mathbf{x}\|^2 N_0 \mathbf{L}_x} - \frac{N_0}{\|\mathbf{x}\|^2 N_0 \mathbf{L}_x} = \frac{1}{\|\mathbf{x}\|^2 N_0} - \frac{1}{\|\mathbf{x}\|^2 \mathbf{L}_x}.
\end{aligned} \tag{41}$$

$$\begin{aligned}
&q(\mathbf{x}, \mathbf{y}_r) \\
&= \frac{1}{\mathbf{L}_x} \exp \left(-\alpha \|\mathbf{x}\|^2 \left(\frac{1}{N_0} - \left(\frac{1}{\|\mathbf{x}\|^2 N_0} - \frac{1}{\|\mathbf{x}\|^2 \mathbf{L}_x} \right) \|\mathbf{x}\|^2 \right) \right. \\
&\quad \left. + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 \right) I_0 \left(2\sqrt{\alpha} \left(\frac{1}{N_0} - \left(\frac{1}{\|\mathbf{x}\|^2 N_0} - \frac{1}{\|\mathbf{x}\|^2 \mathbf{L}_x} \right) \|\mathbf{x}\|^2 \right) |\mathbf{x}^\dagger \mathbf{y}_r| \right).
\end{aligned} \tag{42}$$

Thus the likelihood function is simply:

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}_r) &= \frac{1}{\mathbf{L}_x} \exp \left(-\frac{\alpha \|\mathbf{x}\|^2}{\mathbf{L}_x} + \beta_x |\mathbf{x}^\dagger \mathbf{y}_r|^2 \right) \\
&\quad \times I_0 \left(\frac{2\sqrt{\alpha}}{\mathbf{L}_x} |\mathbf{x}^\dagger \mathbf{y}_r| \right).
\end{aligned} \tag{43}$$

B. Finit-Blocklength Bounds

Herein, we present the finite-blocklength information theory tools. An outer bound, derived from the metaconverse theorem (cf. [14, Th. 28]), is introduced, while an inner bound is established using the RCUs bound [17, Th. 1].

Theorem 1. [17, RCU bound, Th. 1].

Let denote the random vectors via $\mathbf{X} = [X_1, X_2, \dots, X_N]$, and their vector realizations via $\mathbf{x} = [x_1, x_2, \dots, x_N]$.

To set the achievability bound, let define the generalized information density as

$$\iota_s(\mathbf{x}, \mathbf{y}) := \ln \frac{q(\mathbf{x}, \mathbf{y})^s}{\mathbb{E}[q(\mathbf{X}', \mathbf{y})^s]}. \quad (44)$$

In the domain of information theory, the Gallager exponent, represented by $s > 0$, characterizes a pivotal factor. The expectation relates to the random vector $\mathbf{X}'$ having N -dimensional i.i.d. components. Over a memoryless channel, the decoding metric $q(\mathbf{x}, \mathbf{y}) = \prod_{n=1}^N q(x_n, y_n)$. The random coding unions (RCUs) posits that, for a specified rate R , the upper bound on the average error probability is defined as:

$$\epsilon \leq \inf_{s>0} \mathbb{E} \left[e^{-[\iota_s(\mathbf{x}, \mathbf{y}) - \ln(2^{R \times N} - 1)]^+} \right], \quad (45)$$

where $[u]^+ := \max(0, u)$.

The maximum likelihood decoding metric is shown to be $q(\mathbf{x}, \mathbf{y}) = p_{\mathbf{Y}|\mathbf{X}, \mathbf{H}}(\mathbf{y} | \mathbf{x}, \mathbf{h})$. Thus, the underlying decoding metric is expressed as $q(\mathbf{x}, \mathbf{y}) \propto \exp\left(-\frac{1}{N_0} \|\mathbf{y} - \mathbf{h}\mathbf{x}\|^2\right)$.

Theorem 2. [14, Metaconverse (Verdú–Han) bound, Th. 28]. To set the converse bound, consider :

$$j_s(\mathbf{x}, \mathbf{y}) := \ln \frac{q(\mathbf{x}, \mathbf{y})^s}{\mathbb{E}[q(\mathbf{X}', \mathbf{y})^s]^{1/s}} = \frac{1}{s} \iota_s(\mathbf{x}, \mathbf{y}) \quad (46)$$

Subsequently, for a given rate R , the lower bound on the average error probability is given as follows.

$$\epsilon \geq \sup_{s>0} \max_{\lambda} \mathbb{P}[j_s(\mathbf{x}, \mathbf{y}) \leq \lambda] - e^{\lambda - R \times N}. \quad (47)$$

REFERENCES

- [1] 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.
- [2] G. Caire, G. Taricco and E. Biglieri, "Bit-Interleaved Coded Modulation," *IEEE Transactions on Information Theory*, vol. 44, pp. 927–946, May 1998.
- [3] B. Lee, S. Park, D. J. Love, H. Ji and B. Shim, "Packet Structure and Receiver Design for Low Latency Wireless Communications With Ultra-Short Packets," in *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 796–807, Feb. 2018.
- [4] M. Sy, R. Knopp, "Enhanced Low-Complexity Receiver Design for Short Block Transmission Systems," 34th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2023), Toronto, ON, Canada, Sept. 2023.
- [5] M. Sy, R. Knopp, "Novel Joint Estimation and Decoding Metrics for Short-Blocklength Transmission Systems," 2023 IEEE Conference on Standards for Communications and Networking (CSCN), Munich, Germany, Nov. 2023.
- [6] M. Sy, R. Knopp, "Receiver Metric Design for Short-Block Channels : A Perspective for Reliable 6G Signalling Scenarios", 2024 European Conference on Networks and Communications & 6G Summit (EuCNC & 6G Summit), Antwerp, Belgium, June 3–6, 2024.
- [7] N. Doan, "Low-complexity decoding of short linear block codes with machine learning," PhD Dissertation, McGill University, May 2022.
- [8] C. Yue, V. Miloslavskaya, M. Shirvanimoghaddam, B. Vucetic and Y. Li, "Efficient Decoders for Short Block Length Codes in 6G URLLC," *IEEE Communications Magazine*, Vol. 61, no. 4, pp 84–90, April 2023.
- [9] H. Gamage, V. Ransinghe, N. Rajatheva and M. Latva-aho, "Low Latency Decoder for Short Blocklength Polar Codes," 2020 European Conference on Networks and Communications (EuCNC), Dubrovnik, Croatia, 2020.
- [10] C. Zheng, F. Zheng, J. Luo, P. Zhu, X. You and D. Feng, "Mini-Slot-Assisted Short Packet URLLC: Differential or Coherent Detection?", preprint on ArXiv, <https://arxiv.org/abs/2408.14089>, August 26, 2024.
- [11] X. Mestre, D. Gregoratti and P. Zhang, "Optimum Performance of Short Block Length Codes Under Multivariate Stationary Rayleigh Fading," in *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2122–2136, March 2020.
- [12] J. Cao et al., "Independent Pilots Versus Shared Pilots: Short Frame Structure Optimization for Heterogeneous-Traffic URLLC Networks," in *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 5755–5769, Aug. 2022.
- [13] M. Xhemrishi, M. C. Coşkun, G. Liva, J. Östman 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.
- [14] Y. Polyanskiy, H. V. Poor and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," in *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 3GPP TS 38.212 V16.2.0, "Technical Specification Group Radio Access Network, Multiplexing and channel coding", July 2020.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] K. Niu and K. Chen, "CRC-Aided Decoding of Polar Codes", *IEEE Communications Letters*, vol. 16, No. 10, pp. 1668–167, October 2015.
- [24] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes," 2008 IEEE International Symposium on Information Theory, Toronto, ON, Canada, 2008, pp. 1173–1177.
- [25] U. U. Fayyaz and J. R. Barry, "Polar codes for partial response channels," 2013 IEEE International Conference on Communications (ICC), Budapest, Hungary, 2013, pp. 4337–4341.
- [26] R. G. Gallager, "Low Density Parity Check Codes", Cambridge, USA: MIT Press, July 1963.
- [27] D. E. Hocevar, "A reduced complexity decoder architecture via layered decoding of LDPC codes," *IEEE Workshop on Signal Processing Systems*, 2004. SIPS 2004., Austin, TX, USA, 2004, pp. 107–112.
- [28] J. Zhang, Y. Lei, and Y. Jin, "Check-node lazy scheduling approach for layered belief propagation decoding algorithm," *Electron. Lett.*, vol. 50, no. 4, pp. 278–279, Feb. 2014.
- [29] LIFANG WANG "Implementation of Low-Density Parity-Check codes for 5G NR shared channels," master Thesis, August 30, 2021.
- [30] M. Sy, "Demystifying 5G Polar and LDPC Codes: A Comprehensive Review and Foundations", preprint on ArXiv, <https://doi.org/10.48550/arXiv.2502.11053>, February 15, 2025.
- [31] Robert G. Gallager, "Circularly-Symmetric Gaussian random vectors," *Massachusetts Institute of Technology*, January 1, 2008.
- [32] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," in *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.
- [33] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," in *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 543–564, March 2000.
- [34] Woodbury, Max A., "Inverting modified matrices", *Statistical Research Group, Memo. Rep. no. 42*, Princeton University, Princeton, N. J., 1950. pp.4.
- [35] 3GPP TR 38 901 v16.1.0, "Technical Report, Study on channel model for frequencies from 0.5 to 100 GHz", Nov 2020.
- [36] Sylvester, James J., "On the relation between the minor determinants of linearly equivalent quadratic functions". *Philosophical Magazine* 1, pp. 295–305, 1851.
- [37] I. S. Gradshteyn and I. M. Ryzhik, "Table of Integrals, Series and Products" Academic Press, Seventh edition, 2007.