Symbol-Level GRAND for High-Order Modulation over Block Fading Channels

Ioannis Chatzigeorgiou, *Senior Member, IEEE,* and Francisco A. Monteiro, *Member, IEEE*

***Abstract*—Guessing random additive noise decoding (GRAND) is a noise-centric decoding method, which is suitable for low-latency communications, as it supports error correction codes that generate short codewords. GRAND estimates transmitted codewords by guessing the error patterns that altered them during transmission. The guessing process requires the testing of error patterns that are arranged in increasing order of Hamming weight. This approach is fitting for binary transmission over additive white Gaussian noise channels. This letter considers transmission of coded and modulated data over block fading channels and proposes a more computationally efficient variant of GRAND, which leverages information on the modulation scheme and the fading channel. In the core of the proposed variant, referred to as symbol-level GRAND, is an expression that approximately computes the probability of occurrence of an error pattern and determines the order with which error patterns are tested. Analysis and simulation results demonstrate that symbol-level GRAND produces estimates of the transmitted codewords faster than the original GRAND at the cost of a small increase in memory requirements.**

***Index Terms*—Random linear codes, GRAND, hard detection, block fading, QAM, short-packet communication, URLLC.**

I. Introduction

The requirement for ultra-reliable low latency communication (URLLC) was introduced in fifth generation (5G) networks in order to support services that have stringent requirements for extremely low latency (e.g., 1 ms) and high reliability (e.g., 99.999%). Examples of applications that rely on URLLC include machine-type communications for the industrial internet of things (IIoT), virtual reality and driverless vehicles [1], [2]. The low latency goal implies the use of error correction codes with short codewords [3]. However, in an effort to achieve Shannon's capacity, emphasis in pre-5G systems was placed on the construction of codes that map information words onto long codewords. Codes that were invented in the 1960s, e.g., BCH codes and Reed-Solomon codes, have sparked renewed interest [4] for URLLC. This is because they can have short codewords, although they support a limited number of code rates. In contrast, random linear codes (RLCs) can support any code rate but decoding RLCs is NP-hard [5] and is therefore considered impractical.

The recently proposed guessing random additive noise decoding (GRAND) [6] enables universal decoding, i.e., decoding of any binary or nonbinary linear code, including RLCs. GRAND leverages the fact that noise entropy decreases as the channel conditions improve; thus, the list of all possible error patterns that could alter a transmitted codeword reduces in size. Attempting to guess the most likely error pattern becomes more efficient than searching for the most likely transmitted codeword in the code-book, and yields maximum-likelihood performance. A code-book membership test is required to verify whether the combination of a likely error pattern with the received word corresponds to a valid codeword. GRAND considers error patterns in descending order of likelihood and returns the first error pattern that passes the test.

Error correction codes are combined with high-order modulation in wireless systems to improve spectral efficiency. At the receiver, a demodulator converts the sequence of received modulated symbols into a stream of bits. GRAND attempts to estimate the transmitted codeword from the input stream of bits but knowledge of the modulation type is not exploited in the search for error patterns that satisfy the conditions for code-book membership. Although the exploitation of knowledge at the symbol level was envisaged in [6], a mechanism to use this knowledge in the generation of error patterns was not developed. The objective of this letter is to develop a more computationally efficient variant of GRAND that treats the input stream of bits as an equivalent sequence of hard-detected modulated symbols corrupted by block fading and additive noise. We refer to the proposed decoder as *symbol-level* GRAND. The modulation scheme considered in this work is M-ary quadrature phase modulation (M-QAM).

A modification of GRAND that leverages information on the adopted modulation scheme was recently proposed by An *et al.* [3]. Their paper considers the position of M-QAM symbols on the constellation, obtains the probability of a symbol transitioning to one of its nearest neighbors in the special case of channels with memory, and identifies rules that place constraints on the generation of error patterns. In contrast to [3], we derive a closed-form expression for the probability that the input stream of bits is a sequence of a particular combination of binary strings that correspond to different constellation symbols. These symbols have, thus, different numbers of nearest and next-nearest neighbors. This probability expression is used in the generation of error patterns in order of likelihood, when transmission is over a block fading channel. In essence, both our letter and [3] consider the modulation scheme in the decoding process but the estimation and ordering of the error patterns follow different rules because they have been tailored to different channel models. Abbas *et al.* [7] also look into GRAND for flat fading channels but do not consider high-order modulation; their work focuses on the design of a variant of GRAND that exploits receive diversity.

I. Chatzigeorgiou is with the School of Computing & Communications, Lancaster University, UK, e-mail: i.chatzigeorgiou@lancaster.ac.uk.

F. A. Monteiro is with Instituto de Telecommunicações, and ISCTE-Instituto Universitário de Lisboa, Portugal, e-mail: francisco.monteiro@lx.it.pt.

II. System Model

Consider a coding and modulation scheme that employs an $[n,k]$ binary linear block code followed by square Gray-coded M-QAM. The $[n,k]$ block code uses a $k \times n$ generator matrix $\mathbf{G}$ to encode a k-bit input information word, represented by $\mathbf{u} \in \{0,1\}^k$, into an n-bit codeword $\mathbf{x} = \mathbf{uG}$ for $n > k$. The set of all 2^k valid codewords is known as the code-book $\mathcal{C} \subseteq \{0,1\}^n$ of the $[n,k]$ code. The n bits of codeword $\mathbf{x}$ are divided into L strings of $\log_2 M$ bits, i.e., $L = n/\log_2 M$. If E_b denotes the energy per information bit, then $(k/n)E_b$ is the energy per codeword bit, and $(\log_2 M)(k/n)E_b$ is the energy per string of $\log_2 M$ bits. M-QAM maps the L strings onto L symbols of a complex set $\mathcal{S} \subset \mathbb{C}$, which has cardinality M and average energy per symbol $(\log_2 M)(k/n)E_b$. The sequence of L symbols, denoted by $\mathbf{s} \in \mathcal{S}^L$, is transmitted to a receiver over a single-input single-output (SISO) channel.

The channel is impaired by additive white Gaussian noise (AWGN) and block fading. The relationship between the input and output of the discrete-time equivalent channel is [8], [9]

$$\mathbf{r} = h\mathbf{s} + \mathbf{z}, \tag{1}$$

where $\mathbf{r} \in \mathbb{C}^L$ is the sequence of noisy M-QAM symbols at the output of the channel, and $\mathbf{z} \in \mathbb{C}^L$ is a sequence of L zero-mean complex Gaussian random variables with variance N_0. The fading coefficient h is a zero-mean complex random variable with variance $\mathrm{E}\big[|h|^2\big] = 1$. The value of h remains constant during the transmission of sequence $\mathbf{s}$ but changes independently from one sequence to the next. If $\gamma \triangleq (\log_2 M)(k/n)(E_b/N_0)$, the *instantaneous* signal-to-noise ratio (SNR) at the receiver is given by $\texttt{SNR} = |h|^2\gamma$, while the *average* SNR takes the form $\mathrm{E}\big[\texttt{SNR}\big] = \mathrm{E}\big[|h|^2\big]\gamma = \gamma$.

We assume that channel state information is available at the receiver, i.e., the value of h is known for every received sequence, and coherent detection is possible. The received sequence $\mathbf{r}$ is multiplied by $h^{-1} \triangleq h^*/|h|^2$, where h^* and $|h|$ denote the complex conjugate and magnitude, respectively, of h. The product $h^{-1}\,\mathbf{r}$ is input to a hard-detection demodulator, which maps each received symbol to the nearest symbol of the M-QAM constellation and generates $\hat{\mathbf{s}} \in \mathcal{S}^L$. Given that every symbol in $\hat{\mathbf{s}}$ conveys $\log_2 M$ bits, the output of the demodulator is a sequence of $n = L\log_2 M$ bits, denoted by $\mathbf{y} \in \{0,1\}^n$.

The n-bit sequence $\mathbf{y}$ is a potentially erroneous copy of codeword $\mathbf{x}$, and can be expressed as $\mathbf{y} = \mathbf{x} \oplus \mathbf{e}$, where $\oplus$ denotes modulo-2 addition and $\mathbf{e} \in \{0,1\}^n$, referred to as the *error vector*, represents additive noise. GRAND attempts to determine $\mathbf{x}$ by estimating $\mathbf{e}$ [6]. Prospective error vectors, also known as *error patterns*, are ordered in increasing Hamming weight. Each error pattern is subtracted from sequence $\mathbf{y}$ and the result is the estimated codeword if it belongs to code-book $\mathcal{C}$. In other words, if $\hat{\mathbf{e}}$ is the first error pattern in the ordered list of error patterns that satisfies $\mathbf{y} \oplus \hat{\mathbf{e}} \in \mathcal{C}$, then $\hat{\mathbf{x}} = \mathbf{y} \oplus \hat{\mathbf{e}}$. The $(n-k) \times n$ parity-check matrix $\mathbf{H}$ of the $[n,k]$ block code can be used to verify membership of $\mathbf{y} \oplus \hat{\mathbf{e}}$ in $\mathcal{C}$. Given that $\mathbf{G}\mathbf{H}^{\mathrm{T}} = \mathbf{0}$, where $\mathbf{H}^{\mathrm{T}}$ is the transpose of $\mathbf{H}$, the sequence $\mathbf{y} \oplus \hat{\mathbf{e}}$ is a codeword in $\mathcal{C}$ if and only if $\big(\mathbf{y} \oplus \hat{\mathbf{e}}\big)\mathbf{H}^{\mathrm{T}} = \mathbf{0}$ [10]. To curtail computational complexity without inordinately compromising the error correction capability of GRAND, the Hamming weight of the error patterns in the ordered list could be constrained by an upper limit, termed the *abandonment threshold* w_{th} [6]. If w_{th} is set to no more than half the minimum Hamming distance of the $[n,k]$ block code, the error correction capability of GRAND will be similar to or better than that of existing code-specific decoders [11].

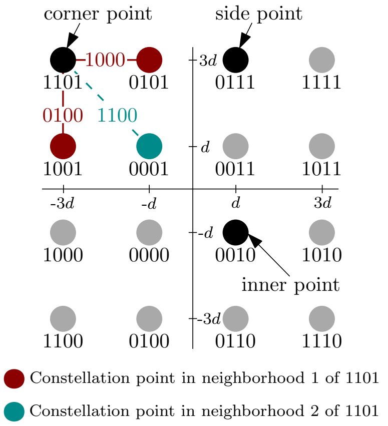

Fig. 1. Constellation diagram of 16-QAM. Representatives of corner, side and inner points are shown in black. The nearest neighbors of 1101 are shown in dark red (neighborhood 1), while the next-nearest neighbor of 1101 is shown in dark cyan (neighborhood 2). The same color-coding scheme has been used to depict the error strings between 1101 and each neighboring point.

The concatenation of M-QAM demodulation and GRAND leads to a hard detection and decoding scheme that operates at the bit level; GRAND, in its current form, neither exploits the structure of the M-QAM constellation nor leverages channel state information. Ordering of the error patterns should be guided by the Euclidean and Hamming distances of the hard-detected symbols to neighboring symbols in the constellation, the channel noise and fading. In the remainder of the letter, knowledge of the constellation and channel state information are integrated in the ordering process of error patterns, and *symbol-level* GRAND is developed and contrasted to the original GRAND [6], henceforth referred to as *bit-level* GRAND.

III. Symbol-Level GRAND

The n-bit codeword $\mathbf{x}$ defined in Section II can be expressed as $\mathbf{x} = (\mathbf{x}_i)_{i=1}^{L}$, where $\mathbf{x}_i$ is the i-th string of $\log_2 M$ bits mapped onto symbol $s_i \in \mathcal{S}$ of M-QAM. The correspondence between all possible values of $\mathbf{x}_i$ and the symbols of M-QAM, when Gray coding is used, is depicted in Fig. 1. The M symbols compose a two-dimensional constellation of equally spaced points, labeled by the mapped values from $\{0,1\}^{\log_2 M}$. The Euclidean distance between two adjacent points along one dimension is $2d$, where d is given by [12], [13]

$$d = \sqrt{\frac{3(\log_2 M)(k/n)E_b}{2(M-1)}}. \tag{2}$$

This distance ensures that the average squared distance of a point from the origin or, equivalently, the average energy per M-QAM symbol is $(\log_2 M)(k/n)E_b$, as stated in Section II.

Let $\mathbf{x}_i$ be mapped onto a point of the constellation diagram. We define two neighborhoods around the point with label $\mathbf{x}_i$; *neighborhood 1* encloses points located at Euclidean distance

TABLE I
EXPRESSIONS FOR THE PROBABILITY TERMS IN (3). FUNCTION $Q(z) \triangleq (1/\sqrt{2\pi}) \int_z^\infty \exp(-t^2/2)\, dt$ IS THE TAIL DISTRIBUTION OF THE STANDARD NORMAL DISTRIBUTION. VARIABLE d' IS GIVEN BY $d' = \sqrt{3\,\mathtt{SNR}/(M-1)}$.

$p_{\mathrm{c},0} = (1/M)(1-Q(d'))^2$
$p_{\mathrm{s},0} = (1/M)(1-Q(d'))(1-2Q(d'))$
$p_{\mathrm{i},0} = (1/M)(1-2Q(d'))^2$
$p_{\mathrm{c},\mathrm{e}_1} = 2(1/M)(1-Q(d'))Q(d')$
$p_{\mathrm{s},\mathrm{e}_1} \approx (1/M)[2(1-Q(d'))Q(d') + (1-2Q(d'))Q(d')]$
$p_{\mathrm{i},\mathrm{e}_1} \approx 4(1/M)(1-2Q(d'))Q(d')$
$p_{\mathrm{c},\mathrm{e}_2} = (1/M)Q^2(d')$
$p_{\mathrm{s},\mathrm{e}_2} \approx 2(1/M)Q^2(d')$
$p_{\mathrm{i},\mathrm{e}_2} \approx 4(1/M)Q^2(d')$

$2d$ from $\mathbf{x}_i$, while *neighborhood 2* includes points located at Euclidean distance $2\sqrt{2}d$ from $\mathbf{x}_i$. We denote by $\mathcal{E}_1(\mathbf{x}_i)$ and $\mathcal{E}_2(\mathbf{x}_i)$ the sets of all *error strings* obtained by modulo-2 addition of $\mathbf{x}_i$ with the members of neighborhoods 1 and 2, respectively. For example, as shown in Fig. 1, neighborhood 1 of $\mathbf{x}_i = 1101$ consists of the points with labels 0101 and 1001; these two points generate the error strings $1101 \oplus 0101 = 1000$ and $1101 \oplus 1001 = 0100$, therefore $\mathcal{E}_1(1101) = \{1000, 0100\}$. Neighborhood 2 of 1101 comprises only one point with label 0001, hence $\mathcal{E}_2(1101) = \{1100\}$. As a result of Gray coding, the Hamming weight of all members of $\mathcal{E}_1(\mathbf{x}_i)$, for any value of $\mathbf{x}_i$, is 1. Similarly, the Hamming weight of all members of $\mathcal{E}_2(\mathbf{x}_i)$, for any $\mathbf{x}_i$, is 2. The cardinalities of $\mathcal{E}_1(\mathbf{x}_i)$ and $\mathcal{E}_2(\mathbf{x}_i)$ depend on the location of the point with label $\mathbf{x}_i$ in the constellation. As illustrated in Fig. 1, the square M-QAM constellation is composed of *corner* points, *side* points and *inner* points. If $\mathbf{x}_i$ is mapped onto a corner, side or inner point, the number of elements in $\mathcal{E}_1(\mathbf{x}_i)$ is 2, 3 or 4, respectively, while the size of $\mathcal{E}_2(\mathbf{x}_i)$ is 1, 2 or 4, respectively.

At the receiver, the n-bit sequence $\mathbf{y}$ at the output of the demodulator can also be written as a sequence of L strings, that is, $\mathbf{y} = (\mathbf{y}_i)_{i=1}^L$, where $\mathbf{y}_i$ is a string of $\log_2 M$ bits that corresponds to a point in the M-QAM constellation diagram. *Bit-level* GRAND generates and tests error patterns in order of likelihood until an error pattern $\hat{\mathbf{e}}$ that satisfies $\mathbf{y} \oplus \hat{\mathbf{e}} \in \mathcal{C}$ is found. The likelihood of an error pattern is taken to be a monotonically decreasing function of its Hamming weight. In the proposed *symbol-level* GRAND, the requirement for $\mathbf{y} \oplus \hat{\mathbf{e}} \in \mathcal{C}$ remains in place but is expressed as $(\mathbf{y}_i \oplus \hat{\mathbf{e}}_i)_{i=1}^L \in \mathcal{C}$, where $\hat{\mathbf{e}}_i$ is the i-th error string of the error pattern $\hat{\mathbf{e}}$, and $\mathbf{y}_i \oplus \hat{\mathbf{e}}_i$ is a string that corresponds to a point in the constellation diagram. Given the structure of the M-QAM constellation, $\hat{\mathbf{e}}_i$ will be a member of either $\mathcal{E}_1(\mathbf{y}_i)$ or $\mathcal{E}_2(\mathbf{y}_i)$ with high probability, or will be a string of $\log_2 M$ zeros, denoted by $\mathbf{0}$. In contrast to bit-level GRAND, symbol-level GRAND does not need to generate and test every realization of $\hat{\mathbf{e}}$ for an increasing Hamming weight; instead, it generates only realizations of $\hat{\mathbf{e}}$ that are composed of error strings, which belong to $\mathcal{E}_1$ and $\mathcal{E}_2$ of the respective strings in $\hat{\mathbf{y}}$, or are equal to $\mathbf{0}$, that is, $(\mathbf{y}_i \oplus \hat{\mathbf{e}}_i)_{i=1}^L \in \mathcal{C}$ for $\hat{\mathbf{e}}_i \in \{\mathbf{0}\} \cup \mathcal{E}_1(\mathbf{y}_i) \cup \mathcal{E}_2(\mathbf{y}_i)$.

Henceforth, for simplicity, we say that an error string $\hat{\mathbf{e}}_i$ is *of type $\mathcal{E}_j$* if $\hat{\mathbf{e}}_i \in \mathcal{E}_j(\mathbf{y}_i)$ for $j = 1, 2$. Let $P(L_1, L_2)$ denote the probability of an error pattern $\hat{\mathbf{e}}$ being composed of L_1 error strings of type $\mathcal{E}_1$, L_2 error strings of type $\mathcal{E}_2$, and $L - L_1 - L_2$ error strings that contain zeros, for a given fading coefficient h and noise variance N_0. The following propositions derive a tight approximation of $P(L_1, L_2)$ and obtain an estimate of the worst-case complexity of symbol-level GRAND.

Proposition 1: An error pattern is a sequence of L error strings, each of length $\log_2 M$ bits. The probability that $L_1 + L_2$ of the error strings are non-zero, when L_1 of them are of type $\mathcal{E}_1$ and L_2 are of type $\mathcal{E}_2$, can be approximated by:

$$\begin{aligned} &P(L_1, L_2) \approx \\ &\sum_{L_\mathrm{c}+L_\mathrm{s}+L_\mathrm{i}=L} \binom{L}{L_\mathrm{c}, L_\mathrm{s}, L_\mathrm{i}} 4^{L_\mathrm{c}+L_\mathrm{s}} \left(\sqrt{M}-2\right)^{L_\mathrm{s}+2L_\mathrm{i}} \\ &\times \sum_{\substack{L_{\mathrm{c,e}}+L_{\mathrm{s,e}}+L_{\mathrm{i,e}}=L_1+L_2 \\ L_{\mathrm{c,e}} \le L_\mathrm{c} \\ L_{\mathrm{s,e}} \le L_\mathrm{s} \\ L_{\mathrm{i,e}} \le L_\mathrm{i}}} \prod_{\substack{\forall \ell \in \mathcal{L} \\ \mathcal{L}=\{\mathrm{c,s,i}\}}} \binom{L_\ell}{L_{\ell,\mathrm{e}}} p_{\ell,0}^{L_\ell - L_{\ell,\mathrm{e}}} \\ &\times \sum_{\substack{L_{\mathrm{c,e}_1}+L_{\mathrm{s,e}_1}+L_{\mathrm{i,e}_1}=L_1 \\ L_{\mathrm{c,e}_1} \le L_{\mathrm{c,e}} \\ L_{\mathrm{s,e}_1} \le L_{\mathrm{s,e}} \\ L_{\mathrm{i,e}_1} \le L_{\mathrm{i,e}}}} \prod_{\forall \ell \in \mathcal{L}} \binom{L_{\ell,\mathrm{e}}}{L_{\ell,\mathrm{e}_1}} p_{\ell,\mathrm{e}_1}^{L_{\ell,\mathrm{e}_1}} p_{\ell,\mathrm{e}_2}^{L_{\ell,\mathrm{e}}-L_{\ell,\mathrm{e}_1}} \end{aligned} \tag{3}$$

where $\mathcal{L} = \{\mathrm{c}, \mathrm{s}, \mathrm{i}\}$ is a set of indices, which signify the possible locations of a constellation point, i.e., corner (c), side (s) or inner (i) point, as explained in Fig. 1. For a given constellation point $\ell \in \mathcal{L}$, an error string will contain zeros, be of type $\mathcal{E}_1$ or be of type $\mathcal{E}_2$ with probability $p_{\ell,0}$, p_{ℓ,e_1} and p_{ℓ,e_2}, respectively. Expressions for the three probabilities for $\ell \in \mathcal{L}$ are provided in Table I. All probabilities in Table I are functions of the distance between a constellation point and the nearest decision boundary, defined as $d' \triangleq d|h|/\sqrt{N_0/2}$ and depicted in Fig. 2. Notice that d' encapsulates the effect of the channel parameters $|h|$ and N_0 on d. Using (2), we obtain $d' = \sqrt{3\,\mathtt{SNR}/(M-1)}$.

Proof: To compute $P(L_1, L_2)$, we need to consider the structure of sequences of L error strings. A sequence will contain L_c, L_s and L_i error strings that are specific to the neighborhoods of corner, side and inner points, respectively. The M-QAM constellation consists of 4 corner points, $4(\sqrt{M}-2)$ side points and $(\sqrt{M}-2)^2$ inner points. Hence, there exist $4^{L_\mathrm{c}}(4(\sqrt{M}-2))^{L_\mathrm{s}}(\sqrt{M}-2)^{2L_\mathrm{i}}$ unique sequences for fixed values of L_c, L_s and L_i. If we take the sum over all possible values of L_c, L_s and L_i, provided that $L_\mathrm{c} + L_\mathrm{s} + L_\mathrm{i} = L$, we obtain the second line of (3). Let us now focus on the L_i error strings in the sequence, which are associated with inner points ($\ell = \mathrm{i}$). Of them, $L_{\mathrm{i,e}}$ error strings will correspond to symbols received in error and will be non-zero; $L_{\mathrm{i,e}_1}$ non-zero error strings will be of type $\mathcal{E}_1$ with probability $p_{\mathrm{i,e}_1}^{L_{\mathrm{i,e}_1}}$, and the remaining $L_{\mathrm{i,e}} - L_{\mathrm{i,e}_1}$ non-zero error strings will be of type $\mathcal{E}_2$ with probability $p_{\mathrm{i,e}_2}^{L_{\mathrm{i,e}}-L_{\mathrm{i,e}_1}}$. The sequence will also contain $L_\mathrm{i} - L_{\mathrm{i,e}}$ error strings that are composed of zeros only, as they correspond to symbols that have been received correctly with probability $p_{\mathrm{i},0}^{L_\mathrm{i}-L_{\mathrm{i,e}}}$. Expressions for $p_{\mathrm{i},0}$, $p_{\mathrm{i,e}_1}$ and $p_{\mathrm{i,e}_2}$ are provided in Table I. We focused on inner points but the same reasoning can be applied to error strings that are associated with corner points ($\ell = \mathrm{c}$) and side points ($\ell = \mathrm{s}$). If we take the sum over all possible values of $L_{\ell,\mathrm{e}}$ and L_{ℓ,e_1}, for every

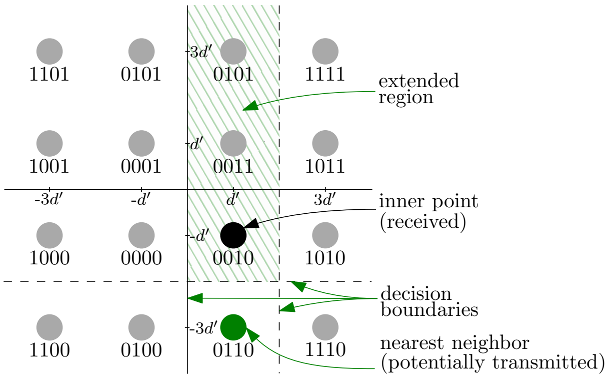

Fig. 2. Example of the extended region considered in the calculation of the probability that 0110 was transmitted given that 0010 is received.

$\ell \in \mathcal{L}$, provided that the total number of non-zero error strings is $L_1 + L_2$ and the total number of non-zero error strings of type $\mathcal{E}_1$ is L_1, we obtain the third and fourth lines of (3). ■

Some expressions in Table I consider extended regions instead of individual constellation points and are, therefore, approximations. For instance, assume that 0010 is received, as shown in Fig. 2. The probability that 0110 was transmitted and a type-$\mathcal{E}_1$ error string occurred considers points that lie above the top-side boundary and between the left-hand and right-hand boundaries of 0110. These constraints do not single out 0010 but define a region (striped area in Fig. 2), which contains three points, including 0010. Owing to the approximations in Table I, expression (3) is also an approximation that becomes tighter as the channel conditions improve.

Proposition 2: The number of error patterns that are tested, until a transmitted codeword is estimated, is a measure of complexity. Given that received codewords of n bits are composed of L strings of $\log_2 M$ bits when M-QAM is used, the worst-case number of error patterns that symbol-level GRAND tests is $9^{\frac{n}{\log_2 M}}$.

Proof: Symbol-level GRAND determines the most likely value of $\hat{\mathbf{e}}_i \in \{\mathbf{0}\} \cup \mathcal{E}_1(\mathbf{y}_i) \cup \mathcal{E}_2(\mathbf{y}_i)$ for $i = 1, \ldots, L$, such that $(\mathbf{y}_i \oplus \hat{\mathbf{e}}_i)_{i=1}^{L} \in \mathcal{C}$, where $\mathbf{y}_i$ corresponds to a constellation point. The number of error strings that are tested, for any value of i, is maximized when $\mathbf{y}_i$ is mapped to a point that has the most neighbors, i.e., an inner point. In that case, $\hat{\mathbf{e}}_i$ will take values from a set that contains nine values: $\mathbf{0}$, four values from $\mathcal{E}_1(\mathbf{y}_i)$ and four values from $\mathcal{E}_2(\mathbf{y}_i)$. The worst-case scenario is the receipt of L strings that all correspond to inner points, which will lead to $9^L = 9^{\frac{n}{\log_2 M}}$ possible error patterns. Using the same reasoning, bit-level GRAND would consider all error patterns of length n bits, that is, $2^n = 2^{L \log_2 M} = M^L$. However, as proven in [6], fewer than 2^{n-k} error patterns are tested, on average, before bit-level GRAND selects an error pattern that produces a valid codeword (either the transmitted codeword or an incorrect codeword). ■

Symbol-level GRAND uses (3) to evaluate $P(L_1, L_2)$ for $L_1 = 0, \ldots, L$ and $L_2 = 0, \ldots, L - L_1$, where $L_1 + L_2 > 0$, and arranges the obtained probability values in descending order. If $P(L_1^*, L_2^*)$ is the i-th probability value in the ordered list, L_1^* and L_2^* are assigned to the entries of the i-th row of an $(L(L+3)/2) \times 2$ lookup table $\mathbf{A} = [a_{i,j}]$, that is, $a_{i,1} = L_1^*$ and $a_{i,2} = L_2^*$. Consequently, the relationship between the entries of consecutive rows of $\mathbf{A}$ is $P(a_{i,1}, a_{i,2}) \geq P(a_{i+1,1}, a_{i+1,2})$. For every row i of $\mathbf{A}$, symbol-level GRAND generates all error patterns that consist of $a_{i,1}$ error strings of type $\mathcal{E}_1$ and $a_{i,2}$ error strings of type $\mathcal{E}_2$. As the algorithm moves from the top row toward the bottom row of $\mathbf{A}$, the likelihood of the generated error patterns reduces. Symbol-level GRAND terminates when an error pattern $\hat{\mathbf{e}}$ that meets the requirement for $\mathbf{y} \oplus \hat{\mathbf{e}} \in \mathcal{C}$ is identified.

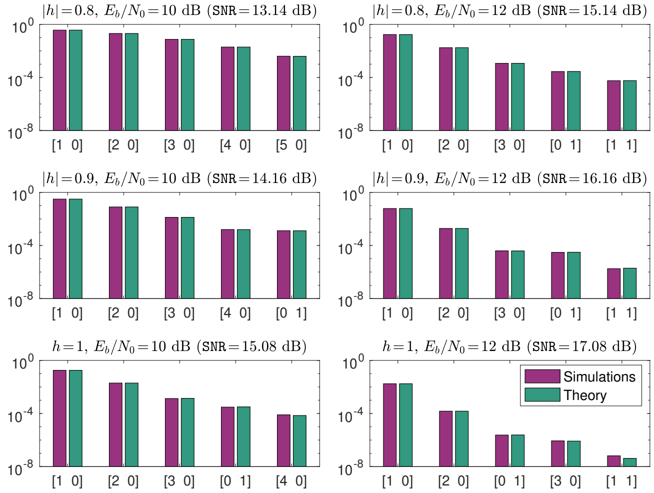

Fig. 3. Structures of error patterns arranged in order of likelihood for different values of E_b/N_0 and $|h|$. A structure $[L_1\ L_2]$ on the horizontal axis of any subplot represents error patterns containing L_1 type-$\mathcal{E}_1$ and L_2 type-$\mathcal{E}_2$ error strings, which occur with probability $P(L_1, L_2)$ shown on the vertical axis.

IV. Results and Discussion

To verify the tightness of (3), we ran simulations for $E_b/N_0 = 10$ dB and $E_b/N_0 = 12$ dB, and observed the error patterns at the output of a 16-QAM demodulator for received sequences impaired by Rayleigh fading or by AWGN only. In the former case, fading coefficients with magnitudes fixed at $|h| = 0.8$ and $|h| = 0.9$ were considered. In the case of AWGN, the fading coefficient was set to $h = 1$. Since coding does not influence $P(L_1, L_2)$ in (3), we focused on uncoded 16-QAM, where $k = n = 128$. As per the definition of SNR in Section II, the two values of E_b/N_0 and the three values of $|h|$ produce six SNR values. For each SNR value, expression (3) was used to build a lookup table $\mathbf{A}$, as described in Section III.

Fig. 3 compares the occurrence probabilities of the five most likely structures of error patterns, which were measured through simulations for each SNR value, with predictions obtained using (3). The structure of an error pattern composed of L_1 type-$\mathcal{E}_1$ and L_2 type-$\mathcal{E}_2$ error strings has been expressed as $[L_1\ L_2]$ on the horizontal axis of each subplot in Fig. 3. The Hamming weight of an error pattern with structure $[L_1\ L_2]$ is $L_1 + 2L_2$ for Gray-coded QAM. The vertical axis of each subplot displays the theoretical and measured values of $P(L_1, L_2)$ for each pattern structure. Fig. 3 demonstrates that predictions match simulation results, thus (3) is a tight approximation of $P(L_1, L_2)$. Although error patterns that contain only type-$\mathcal{E}_1$ error strings, i.e., with structure $[L_1\ 0]$, become dominant at high E_b/N_0 values, structures that incorporate type-$\mathcal{E}_2$ error strings continue to appear among the most likely structures. Fig. 3 also confirms that the structure of error patterns is more pivotal to their likelihood than their Hamming weight.

To compare symbol-level GRAND with bit-level GRAND, in terms of performance and complexity, RLC with $k = 103$ and $n = 128$ was combined with 16-QAM at the transmitter. At the receiver, the decoding algorithms consider error patterns of Hamming weight up to w_{th}, where $w_{\text{th}} \in \{2, 3\}$. This means that symbol-level GRAND examines each row of table $\mathbf{A}$, from top to bottom, and considers a structure $[L_1\ L_2]$ in a row if $L_1 + 2L_2 \leq w_{\text{th}}$; otherwise, it moves to the next row. Given that the estimated codeword $\hat{\mathbf{x}}$ at the output of the decoder is referred to as a *block*, the block error rate (BLER) is used as a measure of performance. The number of error patterns that are tested, on average, until the decoder estimates the transmitted codeword, is adopted as a measure of complexity. Simulation results[1] for a Rayleigh fading channel are presented in Fig. 4. Observe that symbol-level GRAND perfectly matches the BLER of bit-level GRAND and converges to a solution faster than bit-level GRAND. For example, a switch from bit-level to symbol-level GRAND reduces the average number of tests by $\sim 40\%$ for $w_{\text{th}} = 2$, and by $\sim 56\%$ for $w_{\text{th}} = 3$ when $E_b/N_0 \in [20, 34]$, as shown in Fig. 4. The complexity gain of symbol-level GRAND over bit-level GRAND was also proven in Proposition 2, which concluded that, for fixed n, the worst-case complexity of symbol-level GRAND diminishes as M increases, whereas the worst-case complexity of bit-level GRAND is independent of M.

The complexity advantage of symbol-level GRAND comes with a small increase in memory requirements. A two-dimensional (2D) lookup table $\mathbf{A}$ contains the ordered structures of error patterns for a desired value of $\texttt{SNR}$. Different 2D lookup tables can be constructed for a range of $\texttt{SNR}$ values and then stacked together to form a 3D lookup table. If structures of the form $[L_1\ L_2]$ that satisfy $0 < L_1 + 2L_2 \leq w_{\text{th}}$ are only considered, then L_1 and L_2 take values in the ranges $0 \leq L_1 \leq w_{\text{th}}$ and $0 \leq L_2 \leq \lfloor w_{\text{th}}/2 \rfloor$, where $\lfloor t \rfloor$ denotes the integer part of t. Thus, storage of a structure $[L_1\ L_2]$ requires

$$\lambda = \lceil \log_2(w_{\text{th}} + 1) \rceil + \lceil \log_2(\lfloor w_{\text{th}}/2 \rfloor + 1) \rceil \text{ bits} \qquad (4)$$

where $\lceil t \rceil$ = $\lfloor t \rfloor + 1$ if $t > \lfloor t \rfloor$ else $\lceil t \rceil = t$. If only the υ most likely structures for each $\texttt{SNR}$ value are stored, and τ values of $\texttt{SNR}$ are required, the memory size of the 3D lookup table will be $\lambda \upsilon \tau$ bits. For example, the 3D lookup table that was used to obtain the simulation results of symbol-level GRAND for $w_{\text{th}} = 3$ in Fig. 4 contained 1995 bits. This is because $\lambda = 3$ bits per structure of error patterns are needed, according to (4), the $\upsilon = 5$ most likely structures per $\texttt{SNR}$ value were stored, and $\tau = 133$ evenly spaced $\texttt{SNR}$ values, between 5 dB and 38 dB, were considered. The size of the 3D lookup table could be further reduced if fewer and unevenly spaced $\texttt{SNR}$ values were recorded to account for the fact that changes in the ordering of the υ structures occur less frequently as $\texttt{SNR}$ increases.

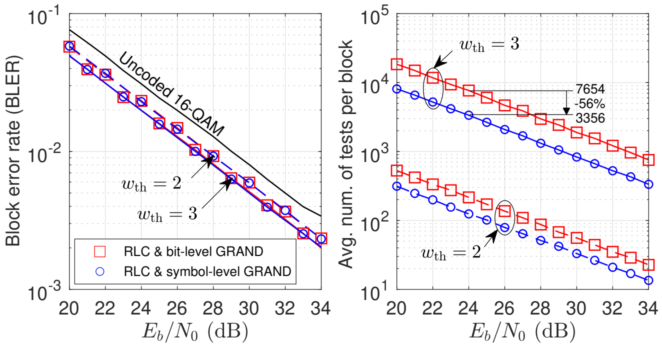

Fig. 4. BLER and average number of tests per block, as functions of E_b/N_0, for bit-level GRAND and symbol-level GRAND, when RLC$[128, 103]$ is used with 16-QAM on a Rayleigh fading channel. The abandonment threshold is set to $w_{\text{th}} = 2$ and $w_{\text{th}} = 3$.

V. Conclusion

This letter introduced symbol-level GRAND, a variant of bit-level GRAND that takes into consideration the structure of the constellation of the adopted M-QAM scheme and the channel conditions in the search of error patterns that are used in the estimation of transmitted codewords. Analysis and simulation results established that symbol-level GRAND requires a marginal increase in memory allocation but offers a significant computational complexity advantage over bit-level GRAND without compromising its error correction capability.

References

[1] P. Nouri, H. Alves, M. A. Uusitalo, O. L. A. López, and M. Latva-aho, "Machine-type wireless communications enablers for beyond 5G: Enabling URLLC via diversity under hard deadlines," *Computer Networks*, vol. 174, no. 3, pp. 1–9, Jun. 2020.

[2] F. Monteiro, O. L. A. López, and H. Alves, "Massive wireless energy transfer with statistical CSI beamforming," *IEEE J. Sel. Topics Sign. Proc.*, vol. 15, no. 5, pp. 1169–1184, Aug. 2021.

[3] W. An, M. Médard, and K. R. Duffy, "Keep the bursts and ditch the interleavers," *IEEE Trans. Commun.*, vol. 70, no. 6, pp.3655–3667, Jun.2022.

[4] P. Pfeifer and H. T. Vierhaus, "Forward error correction in wireless communication systems for industrial applications," in *Proc. Sig. Proc. Alg. Arch. and App. (SPA)*, Poznan, Poland, Sep. 2017.

[5] A. Becker, A. Joux, A. May, and A. Meurer, "Decoding random binary linear codes in $2^{n/20}$: How $1 + 1 = 0$ improves information set decoding," in *Proc. 31st Int. Conf. Theory and App. of Crypt. Techn. (EUROCRYPT)*, Cambridge, United Kingdom, Apr. 2012.

[6] K. R. Duffy, J. Li, and M. Médard, "Capacity-achieving guessing random additive noise decoding," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4023–4040, Jul. 2019.

[7] S. M. Abbas, M. Jalaleddine, and W. J. Gross, "GRAND for Rayleigh fading channels," in *Proc. IEEE Globecom Workshops*, Rio de Janeiro, Brazil, Dec. 2022.

[8] S. Benedetto and E. Biglieri, *Principles of digital transmission: With wireless applications*. Kluwer Academic Publishers, USA, 1999.

[9] D. Tse and P. Viswanath, "Point-to-point communication: Detection, diversity and channel uncertainty," in *Fundamentals in Wireless Communication*. Cambridge Univeristy Press, UK, 2005.

[10] R. J. McEliece, *The Theory of Information and Coding*. Cambridge University Press, 2004.

[11] W. An, M. Médard, and K. R. Duffy, "CRC codes as error correction codes," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Montreal, Canada, Jun. 2021.

[12] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, Jul. 2002.

[13] J. Lu, K. B. Letaief, J. C.-I. Chuang, and M. L. Liou, "M-PSK and M-QAM BER computation using signal-space concepts," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 181–184, Feb. 1999.

[1] Openly available at https://github.com/IoannisChatzigeorgiou/sGRAND