Invariance Properties of Binary Linear Block Codes over a Memoryless Channels with Discrete Input

Ali Abedi and Amir K. Khandani

Coding & Signal Transmission Laboratory(www.cst.uwaterloo.ca)

Department of Electrical and Computer Engineering,

University of Waterloo, Waterloo, ON, Canada, N2L 3G1

e-mail: {ali, khandani}@cst.uwaterloo.ca, Tel: (519)8851211, Fax: (519)8884338

Abstract

This work studies certain properties of the Probability Density Function (pdf) of the bit Log-Likelihood-Ratio (LLR) for binary linear block codes over a memoryless channel with discrete input and discrete or continuous output. We prove that under a set of mild conditions on the channel, the pdf of the bit LLR of a specific bit position is independent of the transmitted code-word. It is also shown that the pdf of a given bit LLR when the corresponding bit takes the values of zero and one are symmetric with respect to each other (reflection of one another with respect to the vertical axis). For the case of channels with binary input, a sufficient condition for two bit positions to have the same pdf is presented.

Index Terms

Bit Decoding, Block Codes, Geometrically Uniform, Log-Likelihood Ratio, Probability Density Function, Regular Channel, Symmetric Channel.

This work is financially supported by Natural Sciences and Engineering Research Council of Canada (NSERC) and by Communications and Information Technology Ontario (CITO). An earlier version [1] of this work has been presented at CISS 2002.

I. Introduction

In the application of channel codes, one of the most important problems is to develop an efficient decoding algorithm for a given code. The class of Maximum Likelihood (ML) decoding algorithms are designed to find a valid code-word with the maximum likelihood value. The ML algorithms minimize the probability of the Frame Error Rate (FER) under the mild condition that the code-words occur with equal probability.

The use of linear binary codes to label the points of signal constellations has been the subject of many investigations. The distance invariance property in such schemes guarantees that the (frame) error probability is independent of the transmitted code-word. This significantly simplifies their design and performance evaluation. For this reason, the study of distance invariance property of codes and signal constellations has been the subject of numerous research works (e.g. see [2]–[9] and their references). More recently, reference [10] studies the invariance property of the error probability at bit level, still relying on ML decoding.

Another class of decoding algorithms, known as bit decoding, compute the probability of the individual bits and decide on the corresponding bit values independent of each other. The straightforward approach to bit decoding is based on summing up the probabilities of different code-words according to the value of their component in a given bit position of interest. Reference [11] provides an efficient method (known as BCJR) to compute the bit probabilities of a given code using its trellis diagram. The main simplification of BCJR has been the SOVA (Soft Output Viterbi Algorithm) [12] which is a sub-optimum solution. A reduced-search BCJR algorithm is also proposed in [13]. There are some special methods for bit decoding based on coset decomposition principle [14], sectionalized trellis diagrams [15], and using the dual code [16], [17].

Maximum Likelihood decoding algorithms have been the subject of numerous research activities while bit decoding algorithms have received much less attention in the past. More recently, bit decoding algorithms have received increasing attention, mainly because they deliver bit reliability information. This reliability information has been effectively used in a variety of applications including Turbo decoding. There has been also an increasing interest to study the use of Turbo-like codes in conjunction with non-binary constellations (e.g. see [18] and its references). Noting the significance of the invariance property (as explained in [2]–[10] in the context of ML decoding) and the growing interest in bit decoding algorithms, it is of interest to define and study invariance properties in the case of bit decoding algorithms in conjunction with binary or non-binary constellations. Such a study sheds light on the structure of codes equipped with bit decoding and simplifies their design and performance evaluation. For example, in [19], such a distance invariance property is exploited to simplify the performance evaluation of the code.

Asymptotic performance analysis of codes in the absence of distance invariance property has received attention in the recent years [18], [20], [21]. Reference [20] studies the limits of performance of Low-Density-Parity-Check (LDPC) codes (under sum-product message-passing decoding) over binary Inter-Symbol-Interference (ISI) channels and proves certain concentration theorems. In [20], an ensemble of coset codes are used to handle the complication caused by the channel memory. Reference [21] applies the concept of LDPC coset codes to multilevel coding and bit-interleaved coded modulation and provides similar concentration theorems. Reference [18] studies asymptotic performance of LDPC codes (under ML decoding) for transmission over non-binary discrete memoryless channels, again by defining an appropriate ensemble of coset codes. Although these works present an effective method (based on applying common randomness at the transmitter and receiver through using an ensemble of coset codes) to overcome the complication caused by the lack of distance invariance property, their proposed approach is limited to asymptotic performance analysis. It would be of interest to study such distance invariance properties for finite block lengths which is the motivation behind current work. In addition, the current article provides some guidelines to distinguish when the underlying code structure already possesses such distance invariance properties as elaborated below.

In Multi-Level Coding (MLC), each bit in the constellation labels is protected by a different binary code. Accordingly, in Multi-Stage Decoding (MSD), these component codes are successively decoded based on the channel output and the decisions from lower levels. It is well known that the combination of MLC and MSD can achieve capacity if the code rates at each level are properly chosen. With this view point, the transmission of multiple label bits can be separated into the parallel transmissions over equivalent binary-input component channels provided that bits from lower levels are known. This provides an effective tool for the analysis and design of these schemes [22]. Reference [21] argues that the application of density evolution and concentration theorem for schemes based MLC and MSD is complicated because for the binary-input component channels the decoding analysis of the all-zeros codeword alone will not necessarily suffice. Reference [21] also shows that for a Gray labeled Amplitude Shift Keying (ASK) constellation, the equivalent binary-input component channels do not posses such invariance property. This motives the authors to use LDPC coset codes to overcome this problem. We will later show (refer to Example 5) that for Natural labeling[1] and ASK modulation, these binary-input component channels indeed satisfy the derived necessary conditions and consequently all-zero codeword can be used to analyze their performance for MLC and MSD. Some more recent studies for the application of MLC and MSD using Turbo and LDPC component codes are reported in [23] and [24], respectively.

Probability density function (pdf) of the bit Log-Likelihood-Ratio (LLR) can be used as a tool for analysis of bit decoding algorithms. It is shown in [25] that for a binary input, *output-symmetric* channel as defined in [27] (assuming that the all zero code-word is transmitted), the pdf of the bit LLR at each node of the code graph, say $f(.)$, possesses a special symmetry defined as $f(x) = f(-x)e^x$ and this symmetry is preserved under belief propagation decoding. Note that the definition of "symmetry" in the current article is different from [25]. Reference [26] on analysis of Sum-Product decoding of LDPC codes over binary input channels with Additive White Gaussian Noise (AWGN) uses a Gaussian approximation for message densities and improves accuracy by enforcing symmetry in the sense of [25]. In [27], it is shown that for a binary input, output-symmetric channel, the conditional probability of error is independent of the transmitted code-word.

[1] Note that the channel capacity is not affected by the method of labeling.

This paper is organized as follows. In section II, the model used to analyze the problem is presented. All notations and assumptions are given in this section. Some theorems are proved on bit decoding algorithms in section III. We conclude in section IV. This work is a continuation of [28] in which the case of AWGN channel with Binary Phase Shift Keying (BPSK) modulation is considered. Throughout the paper, higher indices represent the elements of the sets (for example different code-words), vectors are shown in bold-face, and lower indices represent subsequent components of a sequence (for example sequence of bits within a code-word).

II. MODELING

Assume that a binary linear code $\mathcal{C}$ with code-words of length N is given. Notation $\mathbf{c}^i = (c_1^i, c_2^i, \ldots, c_N^i)$ is used to refer to the i'th code-word and its components. We partition the code into a sub-code C_k^0 and its coset C_k^1 according to the value of the k'th bit position of its code-words, i.e.,

$$C_k^i = \{\mathbf{c} \in \mathcal{C} : c_k = i\}, \quad i = 0, 1. \tag{1}$$

We denote bit wise binary addition of two code-words on the code book as $\mathbf{c}^i \oplus \mathbf{c}^j$. Note that the sub-code C_k^0 is closed under binary addition. Each code-word will be partitioned into L blocks of m bits, assuming $N = mL$, to be transmitted over a channel with a discrete input alphabet set composed of 2^m elements. Notation $\mathbf{I}_j^i$, $i = 1, \ldots, |\mathcal{C}|$, $j = 1, \ldots, L$, is used for these blocks which will be called m-blocks hereafter. For example, code-word $\mathbf{c}^i$ is composed of $(\mathbf{I}_1^i, \mathbf{I}_2^i, \ldots, \mathbf{I}_L^i)$. We assume that there exists a one to one correspondence between the 2^m possible m-blocks and the input symbols of the channel. The set of m-blocks referred as $\mathcal{I}$ forms a group under binary addition.

The channel has 2^m discrete input and discrete or continuous output as shown in Figure 1. For channels with a discrete output alphabet set composed of v elements, $\mathcal{O} = \{\mathbf{x}^1, \ldots, \mathbf{x}^v\}$ and $p(.)$ stands for the probability mass function (pmf). For channels with continuous output $\mathcal{O} \subset \Re^n$ where $\Re$ is the set of real numbers, n is the size of vector $\mathbf{x}$ and $p(.)$ denotes the pdf. In addition, we consider the following two classes of channels: (i) channels with a geometrical representation (motivated by scenarios like a discrete signal constellation over an additive noise channel), and (ii) channels without a geometrical representation. In practice, all considered cases of channels with continuous output fall under the category of channels with a geometrical representation. For channels without a geometrical representation, "channel input" always refers to the corresponding binary label (m-block). For channels with a geometrical representation, depending on the context, the "channel input" may refer to the corresponding binary label or to the actual signal point (these cases will be distinguished by using different notations). In all cases, the channel model considered is *memoryless*. For channels with a finite memory, a known approach to reduce the channel to memoryless is based on a periodic transmission of a known symbol (with a length that is equal to the memory length of the channel) between sub-blocks of data. This results in a memoryless channel, operating over the sub-blocks. An example in this category, based on using a Turbo-like code over an interference channel, is presented in [20].

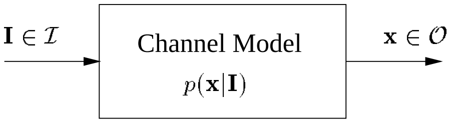
$\mathbf{I} \in \mathcal{I}$
Channel Model
$p(\mathbf{x}|\mathbf{I})$
$\mathbf{x} \in \mathcal{O}$

Fig. 1. Channel Model

Consider the situation of sending a code-word $\tilde{\mathbf{c}} = (\tilde{\mathbf{I}}_1, \ldots, \tilde{\mathbf{I}}_L)$ through the channel. Each m-block $\tilde{\mathbf{I}}_j$, $j = 1, \ldots, L$, will be transmitted and a symbol $\mathbf{x}_j \in \mathcal{O}$, $j = 1, \ldots, L$, will be received at the channel output. A common tool to express the bit probabilities in bit decoding algorithms is based on using the so-called Log-Likelihood-Ratio (LLR). The LLR of the k'th bit position is defined by the following equation,

$$LLR_{\tilde{\mathbf{c}}}(k) = \log \frac{P(\tilde{c}_k = 1|\mathbf{x}_1, \ldots, \mathbf{x}_L)}{P(\tilde{c}_k = 0|\mathbf{x}_1, \ldots, \mathbf{x}_L)}, \tag{2}$$

where $\tilde{c}_k$ is the value of the k'th bit in the transmitted code-word and $\log$ stands for natural logarithm.

Assuming,

$$P(\tilde{c}_k = 0) = P(\tilde{c}_k = 1) = \frac{1}{2}, \tag{3}$$

for a memoryless channel we have,

$$LLR_{\tilde{\mathbf{c}}}(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\mathbf{c}^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\mathbf{c}^i)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)}. \tag{4}$$

We are interested in studying the probabilistic behavior of the LLR.

Assuming a linear code, we derive a set of conditions on the channel for which the choice of $\tilde{\mathbf{c}}$ does not have any impact on the pdf of $LLR_{\tilde{\mathbf{c}}}(k)$ as long as the value of the k'th bit remains unchanged. It will be also shown that under the same set of conditions, the pdf of a given bit LLR when the corresponding bit takes the values of zero and one are symmetric with respect to each other (reflection of one another with respect to the vertical axis). For the case of channels with binary input, i.e., $m = 1$, a sufficient condition for the LLR of two bit positions to have the same pdf is presented. The following sufficient condition is required to carry out the proofs.

For any $\tilde{\mathbf{I}} \in \mathcal{I}$ and any $\mathbf{x} \in \mathcal{O}$, one can find $\mathbf{y} \in \mathcal{O}$ such that for all $\mathbf{I} \in \mathcal{I}$, we have $p(\mathbf{x}|\mathbf{I} \oplus \tilde{\mathbf{I}}) = p(\mathbf{y}|\mathbf{I})$, i.e.,

$$\forall \tilde{\mathbf{I}} \in \mathcal{I},\ \forall \mathbf{x} \in \mathcal{O},\ \exists \mathbf{y} \in \mathcal{O} : p(\mathbf{x}|\mathbf{I} \oplus \tilde{\mathbf{I}}) = p(\mathbf{y}|\mathbf{I}),\quad \forall \mathbf{I} \in \mathcal{I}. \tag{5}$$

This is equivalent to,

$$\forall \tilde{\mathbf{I}}^1, \tilde{\mathbf{I}}^2 \in \mathcal{I},\ \forall \mathbf{x} \in \mathcal{O},\ \exists \mathbf{y} \in \mathcal{O} : p(\mathbf{x}|\mathbf{I} \oplus \tilde{\mathbf{I}}^1 \oplus \tilde{\mathbf{I}}^2) = p(\mathbf{y}|\mathbf{I}),\quad \forall \mathbf{I} \in \mathcal{I}. \tag{6}$$

As we will see later, the form given in (6) is more convenient to use in the proofs of the theorems.

A. Channels with a Geometrical Representation

We use the notation $\mathbf{P}_{\mathbf{I}^i} \in \Re^n$ to refer to the channel input symbols representing $\mathbf{I}^i$. In this case, the m-blocks are just labels of the points in an Euclidean space. We assume that the signal set at the channel input is Geometrically Uniform (GU) [29]. This means that for any given pair of signal points, say $\mathbf{P}_{\tilde{\mathbf{I}}^1}$

and $\mathbf{P}_{\tilde{\mathbf{I}}^2}$, there exists an isometry which transforms $\mathbf{P}_{\tilde{\mathbf{I}}^1}$ to $\mathbf{P}_{\tilde{\mathbf{I}}^2}$ while leaving the signal set unchanged. In addition, we assume that the isometry that transforms $\mathbf{P}_{\tilde{\mathbf{I}}^1}$ to $\mathbf{P}_{\tilde{\mathbf{I}}^2}$ will also transform $P_{\mathbf{I}^i \oplus \tilde{\mathbf{I}}^1 \oplus \tilde{\mathbf{I}}^2}$ to $P_{\mathbf{I}^i}$ for all $\tilde{\mathbf{I}}^1$, $\tilde{\mathbf{I}}^2$ and $\mathbf{I}^i$.

Defining $\mathbf{x}$ and $\mathbf{y}$ following (6), it is easy to see that under the following conditions:

(i) $\mathbf{y}$ is selected as the image of $\mathbf{x}$ under the isometry $\mathbf{P}_{\tilde{\mathbf{I}}^1} \Longrightarrow \mathbf{P}_{\tilde{\mathbf{I}}^2}$

(ii) $p(\mathbf{x}|\mathbf{P}_{\tilde{\mathbf{I}}^1})$ is a function of $\|\mathbf{x} - \mathbf{P}_{\tilde{\mathbf{I}}^1}\|$, $\forall \mathbf{x}$, $\forall \mathbf{P}_{\tilde{\mathbf{I}}^1}$

the condition given in (6) will be satisfied. A schematic view illustrating this scenario is shown in Figure 2. A well known example for a channel satisfying condition (ii) above is the AWGN channel.

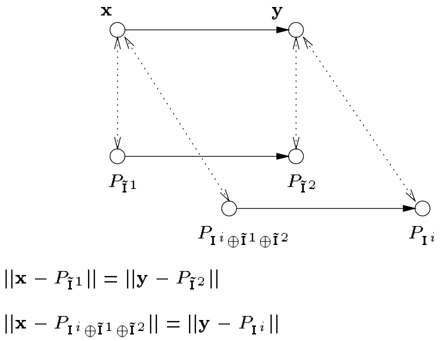
$\mathbf{x}$
$\mathbf{y}$
$P_{\tilde{\mathbf{I}}^1}$
$P_{\tilde{\mathbf{I}}^2}$
$P_{\mathbf{I}^i \oplus \tilde{\mathbf{I}}^1 \oplus \tilde{\mathbf{I}}^2}$
$P_{\mathbf{I}^i}$
$||\mathbf{x} - P_{\tilde{\mathbf{I}}^1}|| = ||\mathbf{y} - P_{\tilde{\mathbf{I}}^2}||$
$||\mathbf{x} - P_{\mathbf{I}^i \oplus \tilde{\mathbf{I}}^1 \oplus \tilde{\mathbf{I}}^2}|| = ||\mathbf{y} - P_{\mathbf{I}^i}||$

Fig. 2. Mapping of points with an isometry

B. *Channels without Geometrical Representation*

In this case, the channel is characterized by a matrix of transition probabilities $\mathbf{A}$ defined as,

$$\mathbf{A}_{u \times v} = [a_{ij}], \quad a_{ij} = p(\mathbf{x}^j|\mathbf{I}^i), \quad u = 2^m = |\mathcal{I}|, \ v = |\mathcal{O}|. \tag{7}$$

In this case, the condition in (6) will be satisfied if after permuting all input symbols by adding an arbitrary m-block $\mathbf{I}$ to them, for each column in $\mathbf{A}_{u \times v}$, there exists another column for which the probability values are permuted in the same order as the corresponding m-blocks.

Reference [30] defines the concept of the *Regular Channel* as follows. Assume that permutation $\psi_{\mathbf{I}}$ acts on the set $\mathcal{O}$ with the property,

$$\forall \ \mathbf{I}^1, \mathbf{I}^2 \in \mathcal{I}, \ \forall \ \mathbf{x}^j \in \mathcal{O} \quad \psi_{\mathbf{I}^1}(\psi_{\mathbf{I}^2}(\mathbf{x}^j)) = \psi_{\mathbf{I}^1 \oplus \mathbf{I}^2}(\mathbf{x}^j). \tag{8}$$

The channel is called a $RegularChannel$ if the probability $p(\mathbf{x}^j|\mathbf{I}^i)$ only depends on $\psi_{\mathbf{I}^i}(\mathbf{x}^j)$. Reference [31] is a more recent work involving regular channels and some of their properties. Using the language of [31], a channel is regular if the input alphabet can be identified with an Abelian group that acts on the output alphabet by permutation. It can be verified easily that a $Regular$ channels is always $Symmetric$ in sense of Gallager [32] where in [32] the symmetry condition only involves the channel symbols and not the underlying labeling. It turns out that our channel model is indeed equivalent to a regular channel[2].

Here are some examples for the discrete case.

Example 1: For the channel shown in Figure 3 we have,

$$\mathbf{A} = \left[\begin{array}{c|cccc} & \mathbf{x}^1 & \mathbf{x}^2 & \mathbf{x}^3 & \mathbf{x}^4 \\ \hline 0 & 1/2-\epsilon_1 & \epsilon_1 & 1/2-\epsilon_2 & \epsilon_2 \\ 1 & \epsilon_1 & 1/2-\epsilon_1 & \epsilon_2 & 1/2-\epsilon_2 \end{array}\right] \tag{9}$$

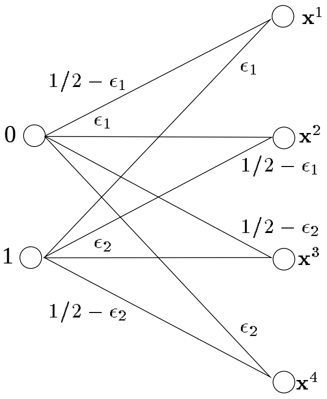

Fig. 3. Channel model for example 1.

Example 2: For the channel shown in Figure 4 we have,

[2] The authors would like to thank G. D. Forney for his invaluable comments on an earlier version of this article, including pointing out references [26], [30], [31].

$$\mathbf{A} = \left[\begin{array}{c|ccccc} & \mathbf{x}^1 & \mathbf{x}^2 & \mathbf{x}^3 & \mathbf{x}^4 & \mathbf{x}^5 \\ \hline 00 & e_0 & e_1 & \epsilon & 0 & 0 \\ 01 & e_1 & e_0 & \epsilon & 0 & 0 \\ 10 & 0 & 0 & \epsilon & e_0 & e_1 \\ 11 & 0 & 0 & \epsilon & e_1 & e_0 \end{array}\right] \tag{10}$$

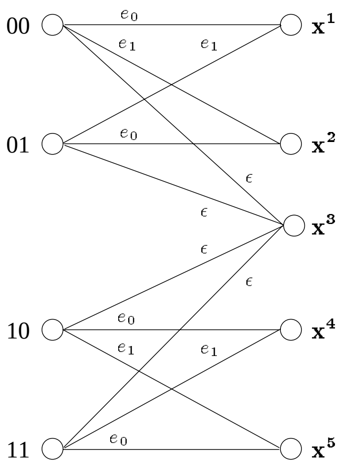

Fig. 4. Channel model for example 2.

where $e_0 + e_1 + \epsilon = 1$.

Example 3: For the channel shown in Figure 5 we have,

$$\mathbf{A} = \left[\begin{array}{c|cccccccc} & \mathbf{x}^1 & \mathbf{x}^2 & \mathbf{x}^3 & \mathbf{x}^4 & \mathbf{x}^5 & \mathbf{x}^6 & \mathbf{x}^7 & \mathbf{x}^8 \\ \hline 00 & e_0 & e_1 & e_2 & e_3 & e_4 & e_3 & e_2 & e_1 \\ 01 & e_2 & e_1 & e_0 & e_1 & e_2 & e_3 & e_4 & e_3 \\ 10 & e_4 & e_3 & e_2 & e_1 & e_0 & e_1 & e_2 & e_3 \\ 11 & e_2 & e_3 & e_4 & e_3 & e_2 & e_1 & e_0 & e_1 \end{array}\right] \tag{11}$$

where $e_0 + 2(e_1 + e_2 + e_3) + e_4 = 1$.

It is easy to see that the required condition for the columns of the probability matrix are satisfied in all of the above examples.

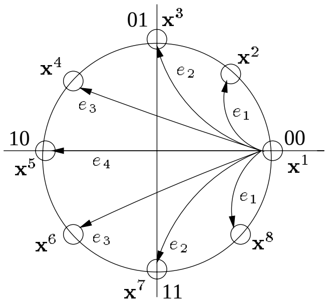

Fig. 5. Channel model for example 3: The values of error probabilities which are not shown follow the same pattern as the values specified on the figure.

Example 4: In this example, the invariance properties of the received signal set for a Code Division Multiple Access (CDMA) system with BPSK modulation (composed of signals $\mathbf{s} \in \{-1, 1\}$) in AWGN channel is studied. It is well known that in a CDMA system, the received signal for a given user is the sum of the transmitted signal plus an interference term $\mathbf{m}$; i.e.,

$$\mathbf{x} = \mathbf{s} + \mathbf{m} + \mathbf{n}, \tag{12}$$

where $\mathbf{m}$ is a linear combination of the modulated bits sent to the other users. Assuming that the bits sent to different users are independent of each other and zero/one occur with equal probability, it easily follows that the interference term $\mathbf{m}$ has a pdf which is symmetrical with respect to the vertical axis. Consequently, the equivalent additive noise term, namely $\mathbf{m} + \mathbf{n}$, has a probability density function which is symmetrical with respect to the vertical axis. It follows that the conditions given in Section II-A are satisfied.

Example 5: In this example, the invariance properties of the equivalent binary-input component channels in MLC and MSD (refer to [22] for definition) with ASK modulation and Natural labeling is investigated. We first note that for ASK modulation (assuming uniformly spaced points) and Natural labeling, the co-ordinates of the signal points can be expressed as linear combination of the underlying bit values. Using this property and following a reasoning similar to Example 4, it easily follows that the conditions expressed in Section II-A are satisfied.

III. Main Results

Using the above definitions and assuming that condition (6) is satisfied, we have the following theorems:

Theorem 1: The pdf of $LLR_{\tilde{\mathbf{c}}}(k)$ is not affected by the choice of the transmitted code-word $\tilde{\mathbf{c}}$ as long as the value of the k'th bit remains unchanged.

Proof: Consider two code-words $\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2$ which have the same value in their k'th bit position. Let us assume that $\tilde{\mathbf{c}}^1$ is transmitted through the channel and $(\mathbf{x}_1, \ldots, \mathbf{x}_L)$ is received. This results in a realization of random variable $LLR_{\tilde{\mathbf{c}}^1}(k)$ with a value of,

$$LLR_{\tilde{\mathbf{c}}^1}(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \mathbf{c}^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \mathbf{c}^i)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{x}_j | \mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{x}_j | \mathbf{I}_j^i)}, \tag{13}$$

that occurs with probability $p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \tilde{\mathbf{c}}^1)$. Noting the $\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2 \in C_k^0$, it is easy to show that,

$$LLR_{\tilde{\mathbf{c}}^1}(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \mathbf{c}^i \oplus \tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \mathbf{c}^i \oplus \tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{x}_j | \mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{x}_j | \mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2)}, \tag{14}$$

where $\tilde{\mathbf{I}}_j^1, \tilde{\mathbf{I}}_j^2, \mathbf{I}_j^i$ are the j'th m-blocks of the code-words $\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2, \mathbf{c}^i$, respectively.

If $\tilde{\mathbf{c}}^2$ is transmitted, assuming condition (6) is satisfied, there exists $(\mathbf{y}_1, \ldots, \mathbf{y}_L), \mathbf{y}_j \in \mathcal{O}, j = 1, \ldots, L$, such that,

$$p(\mathbf{y}_j | \tilde{\mathbf{I}}_j^2) = p(\mathbf{x}_j | \tilde{\mathbf{I}}_j^1). \tag{15}$$

Noting that the channel is memoryless, from (15), we conclude that,

$$\prod_{j=1}^{L} p(\mathbf{y}_j | \tilde{\mathbf{I}}_j^2) = \prod_{j=1}^{L} p(\mathbf{x}_j | \tilde{\mathbf{I}}_j^1), \tag{16}$$

$$p(\mathbf{y}_1, \ldots, \mathbf{y}_L | \tilde{\mathbf{c}}^2) = p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \tilde{\mathbf{c}}^1). \tag{17}$$

This $(\mathbf{y}_1, \ldots, \mathbf{y}_L)$ results in a realization of random variable $LLR_{\tilde{\mathbf{c}}^2}(k)$ with a value of,

$$LLR_{\tilde{\mathbf{c}}^2}(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{y}_1, \ldots, \mathbf{y}_L | \mathbf{c}^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{y}_1, \ldots, \mathbf{y}_L | \mathbf{c}^i)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{y}_j | \mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{y}_j | \mathbf{I}_j^i)}. \tag{18}$$

Using condition (6), we conclude that (14) and (18) are equal to each other. This means for each realization of the random variable $LLR_{\tilde{\mathbf{c}}^1}(k)$, there exists a realization of the random variable $LLR_{\tilde{\mathbf{c}}^2}(k)$ with the same value and occurring with the same probability, i.e., $p(\mathbf{y}_1, \ldots, \mathbf{y}_L|\tilde{\mathbf{c}}^2) = p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\tilde{\mathbf{c}}^1)$. This completes the proof that the random variables $LLR_{\tilde{\mathbf{c}}^1}(k)$ and $LLR_{\tilde{\mathbf{c}}^2}(k)$ have the same pdf. ■

The above result differs from the result derived in [27] in the following ways: (i) We deal with the *pdf* of the bit LLR (which is independent of the decoding algorithm) while [27] deals with the bit error probability in different iterations of a message passing algorithm and shows that it is independent of the transmitted code-word for a binary-input memoryless output-symmetric channel (refer to [27] for definition). (ii) Our channel model can handle non-binary inputs which is more general as compared to the binary-input channel considered in [27].

Theorem 2: The pdf's of $LLR_{\tilde{\mathbf{c}}}(k)$ for $\tilde{c}_k = 0$ and $\tilde{c}_k = 1$ are the reflections of each other with respect to the vertical axis, i.e., $f_{\tilde{c}_k=0}(.) = -f_{\tilde{c}_k=1}(.)$ where $f_{\tilde{c}_k=0}(.)$ and $f_{\tilde{c}_k=1}(.)$ are the pdf of $LLR_{\tilde{\mathbf{c}}}(k)$ for $\tilde{c}_k = 0$ and $\tilde{c}_k = 1$, respectively.

Proof: Consider two code-words $\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2$ which have different values in their k'th bit position. Let us assume that $\tilde{\mathbf{c}}^1$ is transmitted through the channel and $(\mathbf{x}_1, \ldots, \mathbf{x}_L)$ is received. This results in a realization of random variable $LLR_{\tilde{\mathbf{c}}^1}(k)$ with a value of,

$$LLR_{\tilde{\mathbf{c}}^1}(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\mathbf{c}^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\mathbf{c}^i)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)}, \tag{19}$$

that occurs with probability $p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\tilde{\mathbf{c}}^1)$. Noting the $\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2 \in C_k^1$, it is easy to show that,

$$LLR_{\tilde{\mathbf{c}}^1}(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2)} = -\log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2)}, \tag{20}$$

where $\tilde{\mathbf{I}}_j^1, \tilde{\mathbf{I}}_j^2, \mathbf{I}_j^i$ are the j'th m-blocks of the code-words $\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2, \mathbf{c}^i$, respectively.

Assuming condition (6) is satisfied and noting that the channel is memoryless, using the same approach as theorem 1, it is easy to show that if $\tilde{\mathbf{c}}^2$ is transmitted, there exists $(\mathbf{y}_1, \ldots, \mathbf{y}_L), \mathbf{y}_j \in \mathcal{O}, j = 1, \ldots, L$

occurring with probability $p(\mathbf{y}_1, \ldots, \mathbf{y}_L|\tilde{\mathbf{c}}^2) = p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\tilde{\mathbf{c}}^1)$ and resulting in a realization of random variable $LLR_{\tilde{\mathbf{c}}^2}(k)$ with a value of,

$$LLR_{\tilde{\mathbf{c}}^2}(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{y}_1, \ldots, \mathbf{y}_L|\mathbf{c}^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{y}_1, \ldots, \mathbf{y}_L|\mathbf{c}^i)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{y}_j|\mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{y}_j|\mathbf{I}_j^i)}. \tag{21}$$

Using condition (6), we conclude that (20) and (21) are only different in their signs. This means for each realization of the random variable $LLR_{\tilde{\mathbf{c}}^1}(k)$, there exists a realization of the random variable $LLR_{\tilde{\mathbf{c}}^2}(k)$ with the same magnitude and different sign which occurs with the same probability, i.e., $p(\mathbf{y}_1, \ldots, \mathbf{y}_L|\tilde{\mathbf{c}}^2) = p(\mathbf{x}_1, \ldots, \mathbf{x}_L|\tilde{\mathbf{c}}^1)$. This completes the proof that the *pdf* of random variables $LLR_{\tilde{\mathbf{c}}^1}(k)$ and $LLR_{\tilde{\mathbf{c}}^2}(k)$ are the reflections of each other with respect to the vertical axis. ■

Note that for the above two theorems, it is not necessary to partition the code-words into blocks of equal length. In other words, channels with different number of inputs can be used in subsequent block transmissions. The only condition is that the channels in different transmissions should be independent of each other.

Example 6: Reference [10] derives sufficient conditions such that the bit error probability in a constellation does not depend on the transmitted signal point (assuming ML decoding) and shows that an 8-PSK constellation with Gray labeling satisfies these conditions. Motivated by [10], in this example, the invariance properties of an 8-PSK constellation with Gray labeling are studied. In terms of our general framework, we consider a trivial binary code of length three with code-words mapped to the points of the 8-PSK constellation as shown in Fig. 6. In this case, $LLR_{\mathbf{c}}(k)$ refers to the LLR of the k'th bit position assuming that the constellation point with label $\mathbf{c}$ is transmitted. Figure 7 shows examples of various histograms of bit LLRs in this constellation (computed using computer simulation).

To refer to the action of an isometry M, we assume that a point with label $\mathbf{c}$ is mapped to a point with label $M(\mathbf{c})$. For GU constellations, it is easy to prove the following properties:

- The pdf of $LLR_{\tilde{\mathbf{c}}^1}(k)$ and $LLR_{\tilde{\mathbf{c}}^2}(l)$ are the same if there exists an isometry such that: (i) $M(\tilde{\mathbf{c}}^1) = \tilde{\mathbf{c}}^2$, and (ii) for all $\mathbf{c}$, the k'th bit position of $\mathbf{c}$ has the same value as the l'th bit position of $M(\mathbf{c})$.

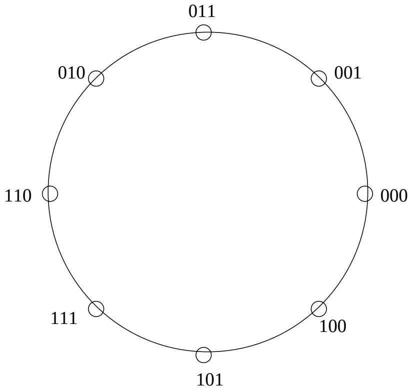

Fig. 6. 8-PSK with Gray labeling.

For example, in Fig. 8-(a), isometry M defined as the reflection with respect to the dashed line preserves the value of the first bit. This means the pdf of $LLR_{\tilde{\mathbf{c}}^1}(1)$ and $LLR_{\tilde{\mathbf{c}}^2}(1)$ are the same for $(\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2) = (001, 011), (000, 010), (100, 110), (101, 111)$. As another example, in Fig. 8-(c), isometry M defined as reflection with respect to the dashed line swaps the values of the 1st and 2nd bit positions. This means the pdf of $LLR_{\tilde{\mathbf{c}}^1}(1)$ and $LLR_{\tilde{\mathbf{c}}^2}(2)$ (similarly pdf of $LLR_{\tilde{\mathbf{c}}^1}(2)$ and $LLR_{\tilde{\mathbf{c}}^2}(1)$) are the same for $(\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2) = (111, 110), (101, 010), (100, 011), (000, 001)$.

- The pdf of $LLR_{\tilde{\mathbf{c}}^1}(k)$ and $LLR_{\tilde{\mathbf{c}}^2}(l)$ are the reflection of each other with respect to the vertical axis if there exists an isometry such that: (i) $M(\tilde{\mathbf{c}}^1) = \tilde{\mathbf{c}}^2$, and (ii) for all $\mathbf{c}$, the k'th bit position of $\mathbf{c}$ has the opposite value as the l'th bit position of $M(\mathbf{c})$.

  For example, in Fig. 8-(b), isometry M defined as the reflection with respect to the dashed line flips value of the 1st bit. This means the pdf of $LLR_{\tilde{\mathbf{c}}^1}(1)$ and $LLR_{\tilde{\mathbf{c}}^2}(1)$ are reflection of each other with respect to the vertical axis for $(\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2) = (100, 000), (101, 001), (111, 011), (110, 010)$. As another example, in Fig. 8-(d), isometry M defined as the reflection with respect to the dashed line swaps the value of the 1st bit with the value of the 2nd bit and then flips 1st and 2nd bits. This means the pdf of $LLR_{\tilde{\mathbf{c}}^1}(1)$ and $LLR_{\tilde{\mathbf{c}}^2}(2)$ (similarly pdf of $LLR_{\tilde{\mathbf{c}}^1}(2)$ and $LLR_{\tilde{\mathbf{c}}^2}(1)$) are reflection of each other with respect to the vertical axis for $(\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2) = (101, 100), (111, 000), (110, 001), (010, 011)$.

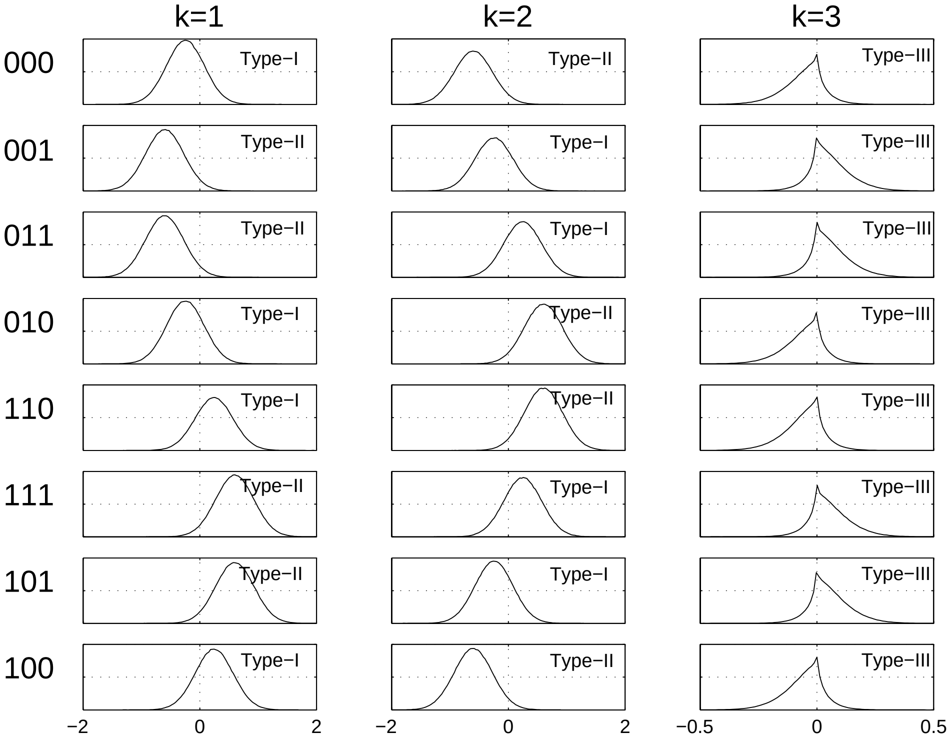

Fig. 7. Histograms of the bit LLR for various bit positions and for various transmitted signal points.

We conclude that the pdf of $LLR_{\mathbf{c}}(1)$ and $LLR_{\mathbf{c}}(2)$ for different $\mathbf{c}$ are either equal or reflection with respect to the vertical axis of the pdf of $LLR_{\mathbf{c}=000}(1)$ (denoted as Type I in Fig. 7) or $LLR_{\mathbf{c}=000}(2)$ (denoted as Type II in Fig. 7). The pdf of $LLR_{\mathbf{c}}(3)$ for different $\mathbf{c}$ is either equal or the reflection with respect to the vertical axis of the pdf of $LLR_{\mathbf{c}=000}(3)$ (denoted as Type III in Fig. 7).

We will now concentrate on the conditions for two bit positions to have the same pdf for their bit LLR. These conditions are presented for a memoryless channel with binary input, i.e., $m = 1$, $N = L$. Note that unlike the previous two theorems, here we require that the channel remains the same in subsequent transmissions.

Let $\mathcal{C}$ be a binary linear code of length N. Consider a permutation $\mathcal{P}$ which permutes the components of each code-word. The set of permutations which map the code-book $\mathcal{C}$ onto itself forms a group called Auto-morphism group of code $\mathcal{C}$.

Theorem 3: Consider two bit positions of a code-word a, b such that $1 \leq a, b \leq N$, $a \neq b$. The

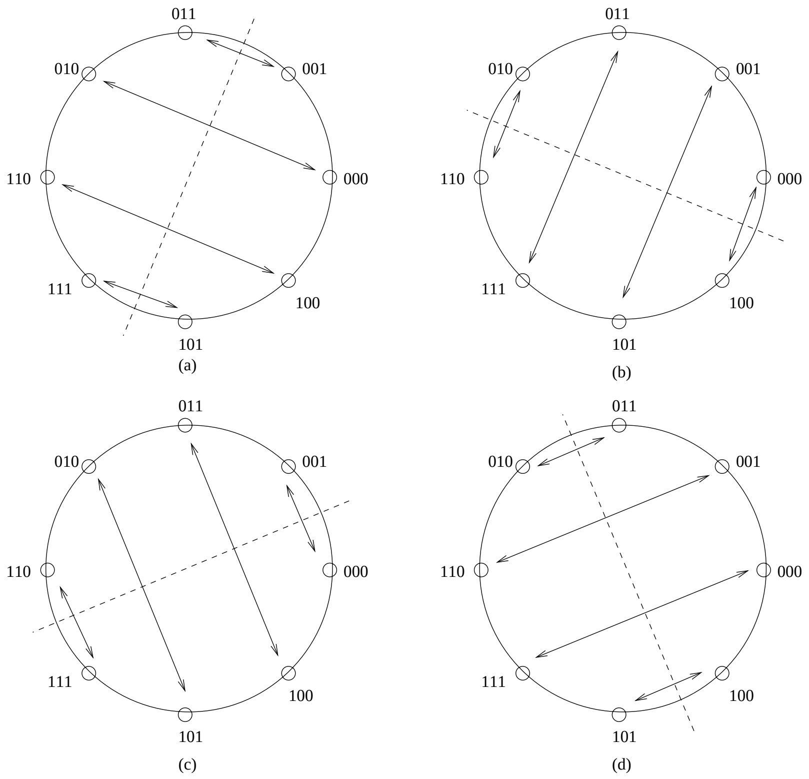

Fig. 8. (a) Isometry which preserves value of the 1st bit. (b) Isometry which flips value of the 1st bit. (c) Isometry which swaps the value of the 1st bit with the value of the 2nd bit. (d) Isometry which swaps the value of the 1st bit with the value of the 2nd bit and then flips 1st and 2nd bits.

channel model is assumed to be memoryless and time invariant with binary input, i.e., $m = 1$, $N = L$. Without loss of generality, assume that the all-zero code-word is transmitted. If there exists a permutation $\mathcal{P}$ within Auto-morphism group of code $\mathcal{C}$ which transfers bit position a to b, then,

$$f_{\tilde{c}_a}(y) = f_{\tilde{c}_b}((-1)^{\tilde{c}_a \oplus \tilde{c}_b} y), \tag{22}$$

where $f_{\tilde{c}_j}(y)$, $j = 1, \ldots, N$, denotes the pdf of random variable Y corresponding to $LLR_{\tilde{\mathbf{c}}}(j)$.

Proof: From theorem 1, we know that pdf of the bit LLR is independent of the transmitted code-word. For simplicity, we consider the situation of sending the all-zero code-word bit by bit and receiving $\mathbf{x}_j$ for bit $\tilde{\mathbf{I}}_j$ in the j'th transmission. This results in a realization of random variable $LLR_{\tilde{\mathbf{c}}=\mathbf{0}}(a)$ with a value of,

$$LLR_{\tilde{\mathbf{c}}=\mathbf{0}}(a) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_a^1} \prod\limits_{j=1}^{N} p(\mathbf{x}_j|\mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_a^0} \prod\limits_{j=1}^{N} p(\mathbf{x}_j|\mathbf{I}_j^i)}. \tag{23}$$

Note that in this theorem $m = 1$ which indicates that $\mathbf{I}_j^i$ are single bits. Permutation $\mathcal{P}$ acts on each code-word $\mathbf{c}^i$ as follows,

$$\mathcal{P} : C_a^0 \longrightarrow C_b^0, \tag{24}$$

$$\mathcal{P} : C_a^1 \longrightarrow C_b^1. \tag{25}$$

Assuming a memoryless time invariant channel, for each $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ we have,

$$P(\mathcal{P}(\mathbf{x}_1, \ldots, \mathbf{x}_N)|\mathcal{P}(\tilde{c}_1, \ldots, \tilde{c}_N)) = P(\mathbf{x}_1, \ldots, \mathbf{x}_N|\tilde{c}_1, \ldots, \tilde{c}_N),$$

where $\mathcal{P}(\tilde{c}_1, \ldots, \tilde{c}_N)$ and $\mathcal{P}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ are obtained by applying permutation $\mathcal{P}$ to $(\tilde{c}_1, \ldots, \tilde{c}_N)$ and $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$, respectively. Applying permutation $\mathcal{P}$ to the terms of summations in (23) and replacing $(\mathbf{y}_1, \ldots, \mathbf{y}_N) = \mathcal{P}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ reveals the one to one correspondence between terms within the summations in $LLR_{\tilde{\mathbf{c}}=\mathbf{0}}(a)$ and $LLR_{\tilde{\mathbf{c}}=\mathbf{0}}(b)$ as seen in (23) and (26),

$$LLR_{\tilde{\mathbf{c}}=\mathbf{0}}(b) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_b^1} \prod\limits_{j=1}^{N} p(\mathbf{y}_j|\mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_b^0} \prod\limits_{j=1}^{N} p(\mathbf{y}_j|\mathbf{I}_j^i)}. \tag{26}$$

The rest of the proof follows similar to the proof of theorem 1. This means for $\tilde{c}_a = \tilde{c}_b$, we have $f_{\tilde{c}_a}(y) = f_{\tilde{c}_b}(y)$. Using theorem 2, for the case of $\tilde{c}_a \neq \tilde{c}_b$, it easily follows that $f_{\tilde{c}_a}(y) = f_{\tilde{c}_b}(-y)$ which completes the proof. ■

We apply this result to the class of Quasi-Cyclic codes as an example for checking the existence of the desired permutation. A code is Quasi-Cyclic if for any cyclic shift of a codeword by ℓ positions, the resulting word is also a code-word ($\ell = 1$ corresponds to a Cyclic code). It easily follows that in this case transferring bit position a to b where $|a - b|$ is an integer multiple of ℓ is achievable by a cyclic shift by ℓ positions. Hence, such bit positions a and b satisfy the above sufficient condition. Cyclic and

Quasi-Cyclic codes allow for linear time encoding and have recently received significant attention in the context of LDPC codes (see [35]–[39] and their references).

On the other hand, the study of codes with Unequal Error Protection (UEP) property is a classical problem in coding theory [42]. Although UEP may be desirable in some applications, in most cases it is preferable to have a uniform level of error protection throughout the entire information sequence [40] (page 586). In situations that UEP is desirable, one usually needs several classes of error protection where the probability of error for bits within a given class are the same. It is known that codes with a pseudo-random construction inherently produce UEP for different bit positions (see [41] and its references). Our results above show that: (i) Quasi-Cyclic LDPC codes can be used to create several classes of error protection with identical performance within each class, and (ii) Cyclic LDPC codes are good candidates for situations that a uniform error protection is needed. This result is in contrast with [43] which presents a method to construct Cyclic codes with several classes of UEP under *hard decision* decoding.

IV. Summary

In this paper the probabilistic behavior of the bit LLR has been investigated over a general channel model with discrete input and discrete or continuous output. We proved that under certain symmetry conditions on the channel, the pdf of the bit LLR for a specific bit position is independent of the transmitted code-word if the value of that bit position remains unchanged. It is also shown that a change in the value of a bit position makes the pdf of that bit LLR reflect through the vertical axis. Finally, a sufficient condition for two bit positions to have the same pdf for their bit LLR is presented.

V. Acknowledgments

The authors would like to thank G. D. Forney, Jr., for his comments on an earlier version of this article, including his pointing out references [26], [30], [31]. We also thank R. Koetter and the anonymous reviewers for their technical input which substantially improved the quality of this article.

REFERENCES

[1] A. Abedi, A. K. Khandani, "Some Properties of Bit Decoding Algorithms Over A Generalized Channel Model," *Proceedings of Conference on Information Sciences and Systems (CISS 2002)*, Princeton, USA, pp. 112-117, March 2002.

[2] D. Slepian, "Group Codes for the Gaussian Channel," *Bell Syst. Tech. J.*, pp. 575-602, April 1968.

[3] G. D. Forney, Jr., "Geometrically Uniform Codes, *IEEE Transactions on Information Theory*, vol. 37, pp. 1241-1260, Sept. 1991.

[4] Y. C. Eldar, H. Bolcskei, "Geometrically Uniform Frames," *IEEE Transactions on Information Theory*, vol. 49, pp. 993-1006, April 2003.

[5] D. Cuong, T. Hashimoto, "A Systematic Approach to the Construction of Bandwidth-Efficient Multidimensional Trellis Codes," *IEEE Transactions on Communications*, vol. 48, pp. 1808-1817, Nov. 2000.

[6] Y. Levy, D. J. Costello Jr., "A Geometric Construction Procedure for Geometrically Uniform Trellis Codes," *IEEE Transactions on Information Theory*, vol. 42 , pp. 1498-1513, Sept. 1996.

[7] R. Garello, S. Benedetto, "Multilevel Construction of Block and Trellis Group Codes," *IEEE Transactions on Information Theory*, vol. 41, pp. 1257-1264, Sept. 1995.

[8] S. Benedetto, G. Montorsi, "A New Decoding Algorithm for Geometrically Uniform Trellis Codes," *IEEE Transactions on Communications*, vol. 44, pp. 581-590, May 1996.

[9] F. Daneshgaran and M. Mondin, "Simplified Viterbi Decoding of Geometrically Uniform TCM Codes," *IEEE Transactions on Communications*, vol. 44, pp. 930-937, Aug. 1996.

[10] R. Garello, G. Montorsi, S. Benedetto, D. Divsalar, F. Pollara, "Labelings and Encoders with the Uniform Bit Error Property with Applications to Serially Concatenated Trellis Codes," *IEEE Transactions on Information Theory*, vol. 48, pp. 123-136, January 2002.

[11] L. R. Bahl, J. Cocke, F. Jelinek, J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284-287, March 1974.

[12] J. Hagenauer, P. Hoeher, "A Viterbi Algorithm With Soft Decision Outputs and Its Applications," *Proceedings of IEEE GLOBECOM*, Dallas, USA, pp. 47.1.1-47.1.6., Nov. 1989.

[13] V. Franz, J. B. Anderson, "Concatenated Decoding With a Reduced-Search BCJR Algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 186-195, Feb. 1998.

[14] L. Ping, K. L. Yeung, "Symbol-by-Symbol Decoding of the Golay Code and Iterative Decoding of Concatenated Golay Codes," *IEEE Transactions on Information Theory*, vol. 45, pp. 2558-2562, Nov. 1999.

[15] Y. Liu, S. Lin, M. P. C. Fossorier, "MAP Algorithms for Decoding Linear Block Codes Based on Sectionalized Trellis Diagrams," *IEEE Transactions on Communications*, vol. 48, pp. 577-586, April 2000.

[16] S. Riedel, "Symbol-by-Symbol MAP Decoding Algorithm For High-Rate Convolutional Codes That Use Reciprocal Dual Codes," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 175-185, Feb. 1998.

[17] C. R. P. Hartmann, L. D. Rudolph, "An Optimum Symbol-by-Symbol Decoding Rule for Linear Codes," *IEEE Transactions on Information Theory*, vol. 22, pp. 514-517, Sept. 1976.

[18] A. Bennatan and D. Burshtein, "On the Application of LDPC Codes to Arbitrary Discrete-Memoryless Channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 417-438, March 2004.

[19] A. Abedi and A. K. Khandani, "An Analytical Method for Approximate Performance Evaluation of Binary Linear Block Codes," *IEEE Transactions on Communications*, vol. 52, pp. 228-235, Feb. 2004.

[20] A. Kavcic, X. Ma, and M. Mitzenmacher, "Binary Inter-symbol Interference Channels: Gallager Codes, Density Evolution and Code Performance Bounds," *IEEE Transactions on Information Theory*, vol. 49, pp. 1636-1652, July 2003.

[21] J. Hou, P. H. Siegel, L. B. Milstein and H. D. Pfister, "Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," *IEEE Transactions on Information Theory*, pp. 2141-2155, Sept. 2003.

[22] U. Wachsmann, R.F.H. Fischer and J.B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Transactions on Information Theory*, vol. 45, pp. 1361-1391, July 1999.

[23] E. Eleftheriou, S. Olcer and H. Sadjadpour, "Application of capacity approaching coding techniques to digital subscriber lines," *IEEE Communications Magazine*, vol. 42, pp. 88-94, April 2004.

[24] P. Limpaphayom and K. A. Winick, "Power- and bandwidth-efficient communications using LDPC codes," *IEEE Transactions on Communications*, vol. 52, pp. 350-354, March 2004.

[25] T. J. Richardson, M. A. Shokrollahi, R. L. Urbanke, "Design of Capacity Approaching Irregular Low-Density-Parity-Check Codes," *IEEE Transactions on Information Theory*, vol. 47, pp. 619-637, Feb. 2001.

[26] S. Y. Chung, T. J. Richardson, and R. L. Urbanke, "Analysis of Sum-Product Decoding of Low-Density-Parity-Check Codes Using Gaussian Approximation," *IEEE Transactions on Information Theory*, vol. 47, pp. 657-670, Feb. 2001.

[27] T. J. Richardson, R. L. Urbanke, "The Capacity of Low-Density-Parity-Check Codes Under Message Passing Decoding," *IEEE Transactions on Information Theory*, vol. 47, pp. 599-618, Feb. 2001.

[28] A. Abedi, P. Chaudhari, A. K. Khandani, "On Some Properties of Bit Decoding Algorithms," *Proceedings of the Canadian Workshop on Information Theory (CWIT 2001)*, Vancouver, Canada, pp. 106-109, June 2001. (available from www.cst.uwaterloo.ca)

[29] G. D. Forney Jr., "Geometrically Uniform Codes," *IEEE Transactions on Information Theory*, vol. 37, pp. 1241-1260, Sept. 1991.

[30] P. Delsarte, P. Piret, "Algebraic Constructions of Shannon Codes for Regular Channels," *IEEE Transactions on Information Theory*, vol. 28, pp. 593-599, July 1982.

[31] G. D. Forney Jr., M. D. Trott, S. Y. Chung, "Sphere-Bound-Achieving Coset Codes and Multilevel Coset Codes," *IEEE Transactions on Information Theory*, vol. 46, pp. 820-850, May 2000.

[32] R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.

[33] U. Erez and R. Zamir, "Error Exponent of Modulo-Additive Noise Channels with Side Information at the Transmitter," *IEEE Transactions on Information Theory*, vol. 47, pp. 210-218, Jan. 2001.

[34] P. Grillet, Algebra, Wiley, May 1999.

[35] H. Tang, J. Xu, Y. Kou, S. Lin, and K. Abdel-Ghaffar, "On Algebraic Construction of Gallager and Circulant Low-Density Parity-Check Codes," *IEEE Transactions on Information Theory*, vol.50, pp. 1269-1279, pp. 1269-1279, June 2004.

[36] B. Ammar, B. Honary, Y. Kou, J. Xu, S. Lin, “Construction of Low-Density Parity-Check Codes Based on Balanced Incomplete Block Designs,” *IEEE Transactions on Information Theory*, vol. 50, pp. 1257-1268, June 2004.

[37] B. Vasic and O. Milenkovic, “Combinatorial Constructions of Low-Density Parity-Check Codes for Iterative Decoding,” *IEEE Transactions on Information Theory*, vol. 50, pp. 1156-1176, June 2004.

[38] M. Fossorier, “Quasi-Cyclic Low-Density Parity-Check Codes From Circulant Permutation Matrices,” *IEEE Transactions on Information Theory*, vol. 50, pp. 1788-1793, Aug. 2004.

[39] Y. Kou, S. Lin, M. Fossorier, “Low-density parity-check codes based on finite geometries: a rediscovery and new results,” *IEEE Transactions on Information Theory*, vol. 47, pp. 2711-2736, Nov. 2001.

[40] Error Control Coding, S. Lin, D. J. Costello, Prentice Hall; 2nd edition, April 2004

[41] A. Huebner, T. Nuecker, D.J. Costello, and K.Sh. Zigangirov, “On Joint Permutor Design and Unequal Error Protection for Multiple Turbo Codes,” *5th International ITG Conference on Source and Channel Coding, Erlangen, Germany*, January 2004.

[42] B. Masnick, J. Wolf, “On linear unequal error protection codes,” *IEEE Transactions on Information Theory*, vol. 13, pp. 600-607, Oct. 1967.

[43] M. Lin, and S. Lin, “Cyclic unequal error protection codes constructed from cyclic codes of composite length,” *IEEE Transactions on Information Theory*, vol. 34, pp. 867-871, July 1988.